MODELING IMAGE SEQUENCES, WITH PARTICULAR APPLICATION TO FMRI DATA

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF STATISTICS
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

By
Neil J. Crellin
December 1996

I certify that I have read this dissertation and that in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Trevor J. Hastie
(Principal Advisor)

I certify that I have read this dissertation and that in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Iain M. Johnstone

I certify that I have read this dissertation and that in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Jerome H. Friedman

Approved for the University Committee on Graduate Studies:

Abstract

Functional magnetic resonance imaging (fMRI) has made it possible to conduct sophisticated human brain mapping neuroscience experiments. Such experiments commonly consist of human subjects being exposed to a designed temporal sequence of stimulus conditions while repeated MRI scans of the brain region of interest are taken. The images produce a view of recently activated neural areas by detecting the subsequent blood flow replenishing the sites.

The data collected in fMRI human brain mapping experiments consists of sequences of images, which may also be thought of as a spatially organized collection of time series collected at each pixel or voxel location of the brain. The dimensionality of such data sets is very high, usually containing thousands of pixels with short time series from a few to around a hundred scans or time points. The data are correlated, both in space and in time. The data are very noisy and the magnitude of the signal increase due to activation can be low.

A wide variety of statistical methodologies have been proposed for the analysis of such data for purposes of detecting localized and inter-connected regions of activation. This dissertation includes an extensive literature review of statistical techniques as applied in the analysis of fMRI human brain mapping experiments.

In this dissertation, we examine time-course models for fMRI human brain mapping data at each pixel location. The convolution model proposed by Friston, Jezzard and Turner [31] and generalized by Lange and Zeger [44] is shown to have intrinsic structural inadequacies leading to fits which do not appear to capture the patterns in the data. More flexible data-driven models based on spline basis expansions are compared. Using bootstrap resampling, we are able to show that a Poisson-based convolution model for the data performs no better than fitting a sinusoid at the activation frequency. Using cubic splines is shown by the resampling methodology to afford a non-trivial improvement in capturing the underlying pattern of the data. Use of the model fits as inputs to a clustering procedure shows an interesting and informative brain map.

The dissertation also proposes use of a restricted singular value decomposition for fMRI data to exploit the known temporal and spatial patterns in the data such as periodicity and smoothness. The restrictions find contrasts from within a basis expansion family with the required properties which orthogonally maximize variability in the data and provide new informative data views.

Regularization of image sequence regression models is also considered, leading to an analogous wavelet shrinkage procedure, and finally interactive software tools for exploratory data visualization in fMRI image sequence models is discussed.

Acknowledgments

Thanks are due to many colleagues with whom I worked at the CSIRO Division of Mathematics and Statistics in Australia for encouraging me to travel abroad to undertake my doctoral research. Peter Diggle, Adrian Baddeley, Terry Speed and Nick Fisher all deserve special mention in this regard.

Many other graduate students with whom I have passed through Stanford's Statistics doctoral program deserve acknowledgment for their friendship and support. Among these I must include John Overdeck, Steve Stern, Josee Dupuis, and Charles Roosen, by no means a complete list. I particularly thank John Overdeck for convincing me to stay, and Charles Roosen for convincing me to finish.

I am also grateful to Steve Stern, Michael Martin and Tom DiCiccio, for making the environment of Sequoia Hall as lively as they did, and for entertaining conversations and helpful advice when they were most needed.

For their roles in helping me along the path to completion, including advice, suggestions, and projects, I am extremely grateful to Tom Cover, Richard Olshen, Jerry Friedman, Brad Efron, Tom DiCiccio, Michael Martin and Iain Johnstone.

I wish to specially thank Geoff Boynton, Steve Engel, Brian Wandell, David Heeger of Psychology, and Gary Glover of Radiology for providing access to their data, numerous extremely stimulating discussions on the analysis of fMRI data, and their friendly support and encouragement.

I particularly thank Judi Davis for her friendship and assistance.

I wish to express my great appreciation to Trevor Hastie for being an excellent and encouraging advisor. From proposing the topic of this dissertation, to being accessible and providing intellectual stimulation and input, Trevor's guidance has been truly outstanding.

Most of all, I thank Kim, for her patience, support, encouragement and love.

Contents

List of Figures

Chapter 1

Introduction

Functional magnetic resonance imaging (fMRI) has made it possible to conduct sophisticated human brain mapping neuroscience experiments. Such experiments commonly consist of human subjects being exposed to a designed temporal sequence of stimulus conditions while repeated MRI scans of the brain region of interest are taken. A common experimental design alternates equal length periods of stimulus and rest for a number of cycles. The magnetic characteristics of hemoglobin in the blood are detectably changed by oxygenation. Recently activated sites of neural activity are replenished with oxygenated blood, allowing identification of brain activity using MRI scans designed to detect such changes in magnetic susceptibility. Neural activation occurs on a millisecond time-scale. The detectable blood flow replenishing the activation sites depends however on local vasculature and can occur as long as several seconds after activation.

The data collected in fMRI human brain mapping experiments consists of sequences of images, which may also be thought of as a spatially organized collection of time series collected at each pixel or voxel location of the brain. The dimensionality of such data sets is very high, usually containing thousands of pixels with short time series from a few to around a hundred scans or time points. The data are correlated, both in space and in time. The data are very noisy and the magnitude of the signal increase due to activation can be low.

A wide variety of statistical methodologies have been proposed for the analysis of such data for purposes of detecting localized and inter-connected regions of activation. This dissertation includes an extensive literature review of statistical techniques as applied in the analysis of fMRI human brain mapping experiments. The literature review is divided between two chapters. The literature reviewed in Chapter 2 considers statistical methodologies for analysis, estimation and modeling of fMRI data, and Chapter 7 describes the literature on statistical inference which may be applied to the results of statistical analyses such as those described in both the literature and also to the analyses proposed in Chapters 3, 4, and 5.

The original contributions contained in the dissertation are in Chapters 3, 4, 5, and 6. Chapter 3 evaluates several proposals for modeling pixel time-courses. Chapter 4 considers a modified form of the singular value decomposition to exploit spatial and temporal relationships known to exist in fMRI data sets. Chapter 5 considers regularized image sequence regression, and Chapter 6 describes interactive visualization tools developed for data exploration in conjunction with the dissertation research.

The dissertation is organized as follows:

- Chapter 2 contains background and literature review. In it can be found an overview description of how human brain mapping experiments are conducted with functional magnetic resonance imaging, with an illustrative example of such data. This chapter also has a substantial literature review of the state of the art in statistical methodology from the fMRI journals. The emphasis in this review is on analysis, estimation and modeling, with a detailed description of inferential issues deferred to Chapter 7.

- Chapter 3 considers time-course models for fMRI data on a pixel by pixel basis. The dominant model based on convolution of a hemodynamic response function with the treatment profile is examined closely and specific inadequacies highlighted. Alternative flexible data-driven models are proposed and shown to offer superior faithfulness to the data while preserving interpretability. Model fits in neighboring pixels showed physiologically interesting patterns of similarity when used as inputs to a clustering procedure.

- Chapter 4 investigates singular value decompositions of fMRI data restricted to have singular vectors with particular features, such as temporal periodicity or spatial smoothness.

- Chapter 5 looks at regularized image regression models and compares a basis expansion restriction shrinkage method to wavelet shrinkage.

- Chapter 6 describes methods of visualizing large fMRI data sets, focusing on interactive, exploratory data visualization methods.

- Further literature review detailing the sophisticated inferential procedures for fMRI data analysis in the literature are detailed in Chapter 7.

Chapter 2

Background and Literature Review

This chapter provides background information describing the nature of data collected in fMRI human brain mapping experiments and a substantial review of the statistical methodology extant in the fMRI literature. The emphasis in the review is on techniques of analysis, estimation and modeling. A separate literature review of inference procedures is contained in Chapter 7. The findings in this chapter are summarized in Section 2.3.6 on page 35.

2.1 Functional Magnetic Resonance Imaging (fMRI)

Functional magnetic resonance imaging (fMRI) is a methodology by which brain areas of cognitive and sensory function can be identified non-invasively using sequences of very fast brain scans.

Magnetic resonance imaging (MRI) is accomplished by observing differences in response between various types of tissue when observed in a strong magnetic field and subjected to radio frequency (RF) radiation pulses. In MRI, the subject is placed in an extremely strong magnetic field, usually produced by a superconducting magnet with field strength between 1 and 4 Tesla. Magnetic resonance imaging and spectroscopy can be performed to target particular isotopes of many elements, but typically hydrogen protons are considered, here exploiting their ubiquity and high abundance in human tissues. In the presence of the strong magnetic field, some fraction of the proton spins align with the field, precessing around the field direction at a frequency called the Larmor frequency, which is directly proportional to the field strength. This precessional spin frequency, the *resonance* frequency, lies in the radio frequency range. At a quantum mechanical level, the hydrogen protons ordinarily consist of equal populations of the two spin orientations $+\frac{1}{2}$ and $-\frac{1}{2}$. In the presence of the strong magnetic field, one of these orientiations $(-\frac{1}{2})$ is a lower energy state, since it has a magnetic characteristic more aligned with the external field. By introducing a momentary pulse of RF radiation at the resonance frequency, the protons will rise to a higher energy state back into probabilistic equilibrium and spin alignment. Local inhomogeneities in the magnetic nature of surrounding tissue will cause the gradual relaxation of these higher energy state protons, releasing the energy as an RF signal which is detected on receiver coils around the subject. Protons in different types of molecules, and hence tissues, relax at different rates based on the differences in their magnetic character. By recording the RF signal *echoes* over time and observing the signal decay rates for different regions, a *contrast* image can be computed based on these recorded signals. Different modalities of MRI image acquisition are based on different contrast calculations, called T_1, T_2 and T_2^*. The time delay between the RF pulse and signal collection is called TE or *time-to-echo* and is set to enhance contrast between different types of tissue.

Functional magnetic resonance imaging is possible because of an effect called **BOLD** contrast, standing for **B**lood **O**xygenation-**L**evel-**D**ependent contrast. The BOLD contrast mechanism was first described, so-named, and observed *in vivo* in Ogawa et al. [54]. Kwong et al. [42] are also among the first to describe the phenomenon in the context of fMRI experiments on primary sensory stimulation. The effect occurs because neural activation is followed by a localized increase in regional cerebral blood flow (rCBF) delivering oxygenated blood to the activated area. Deoxygenated hemoglobin is paramagnetic, while oxygenated hemoglobin is not, and paramagnetism more quickly destabilizes the spin phasing of the protons produced by the RF pulse. It follows that regions of neural activation have a greater concentration of oxygenated blood shortly after neuronal activity, and the protons in that oxygenated blood dephase more slowly, leading to an RF signal in that region which decays more slowly as well. Consequently, activated brain regions appear brighter in MR images when based on contrasts which are able to detect this phenomenon.

Brain mapping experiments using fMRI are typically of the treatment versus control variety, although more sophisticated experimental designs are certainly used in the literature, ranging from multiple treatment designs to fully parametric studies. For paradigms of the latter, see for example Engel et al. [13] and Sereno et al. [56]. The more common treatment versus control experiments are conducted by repeatedly scanning the brain area of interest. Periods of rest and stimulus are alternated, with multiple replications within each regime and several cycles of alternation. The data collected are thus an image sequence of MRI brain scans, together with a covariate encoding whether the scan was taking during a period of rest or stimulation.

2.2 Image Sequence Data

The data we are considering consists of sequences or time series of image data. Let y_i denote the i^{th} image in the sequence, for $i = 1, \ldots, n$, where each image is viewed as a vector of $M = m_1 \times m_2$ pixels, where m_1 and m_2 are the image dimensions. As a motivating example, we consider image sequences obtained from functional MRI scans intended to identify regions of the brain activated by particular (in our example below, visual) stimuli. In these fMRI experiments, the experimental subject is scanned while being exposed to alternating periods of rest and stimuli. By contrasting images taken during the rest period with those taken during stimulation, identification of the brain region associated with the functional neural activity is possible[1]. Encoding whether an image is scanned during a period of rest or stimulation, we can create a covariate vector which might be used in regression style data analysis. In general, we could have associated with each image a vector of p covariates x_i which may contain a general mix of quantitative variables such as time or amount of drug and qualitative variables such as our encoding of rest versus stimulus state.

We represent our image data as an $n \times M$ matrix Y, with each row being the pixels of an image in the sequence, and each column being the time-course at individual pixel locations. The covariate data is represented as an $n \times p$ matrix with the p covariate values corresponding to each image in the sequence as rows. We can model such sequences using a linear regression model of the form

$$Y_{n\times M} = X_{n\times p}\beta_{p\times M} + \epsilon$$

The rows of the coefficient matrix β represent regression coefficient values at each pixel location, and as such can be viewed as coefficient images. In the fMRI example, high values of regression coefficients associated with covariate terms which alternate at the same frequency as the stimulus could be used to identify the regions of neural activation, and presentation of the regression coefficient images could visually display these regions in an immediately interpretable anatomical framework.

The quantity of data collected in an fMRI experiment can be substantial. Over one hundred scans will typically be taken at a resolution which produces hundreds or thousands of pixels per scan. To visually inspect such data as either an image

[1]Oxygenation of hemoglobin changes its magnetic characteristics, and the vascular flow to recently activated sites of neural activity allows identification of brain activity using MRI scans. Note that it is the post-activation blood flow to the region which is detected, not neural activity per se.

sequence or as pixelwise time courses usefully may require careful data reduction or presentation methods. Visualization will be elaborated in greater detail in later sections. Some data from an experiment are presented here to assist in understanding the nature of the data in preparation for discussion of analytical techniques in the next section.

Figure (2.1) shows an anatomical MRI scan of the human brain. This is an oblique slice taken as part of an investigation into activation of primary visual cortex, which is located along the calcarine sulcus. The area of activation for this particular experiment is expected to lie within the boxed 16 by 16 pixel subregion.

Figure 2.1: *An anatomic MRI oblique scan localized around the calcarine sulcus. The 16 by 16 subregion indicated contains the area of primary visual cortex being stimulated.*

The time series for each pixel are displayed in Figure (2.2), both as raw data and mean-corrected series. Clearly some organization or visualization procedures are needed to make sense of this quantity of such data.

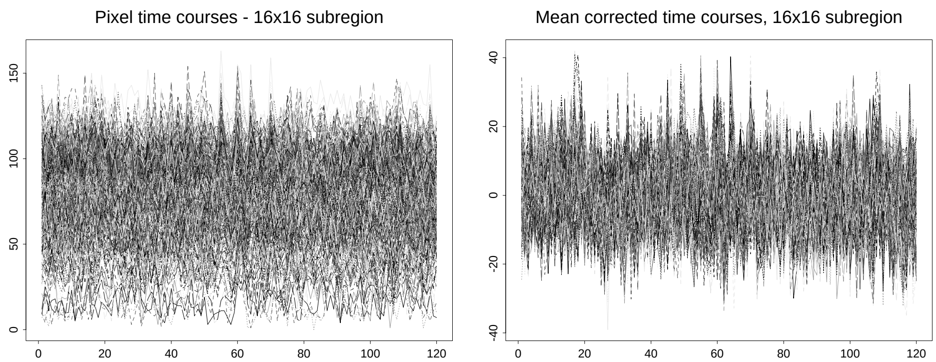

Figure 2.2: *On the left are shown the raw data of the 256 pixelwise time courses in the 16 by 16 subregion identified in Figure (2.1). On the right are the same data adjusted to each have mean zero.*

Organizing the time series from contiguous pixels provides display of the spatial arrangement of the responses and allows viewing of the individual time series at pixels. Figure (2.3) offers such a display, showing a tessellation into subregions and a representative time course from each subtessellations pixels. In this example, a pixel closest to the subtile center is chosen from each tile. Similar displays using tile averages or all series contained in a tiles pixels are similarly informative.

The stimulus-rest regime for this series of functional scans consists of stimulus for 15 scans then rest for 15 scans repeated for 4 cycles, a total of 120 functional scans. A scan was taken every 1.5 seconds. Figure (2.4) below shows the time series produced at a single pixel, with the stimulus regime superimposed at the top of the figure. Note there is a delay between the change in treatment and the change in signal. Note also the appearance of an overall linear trend, a common artifact from the imaging modality, usually subtracted from the raw data before analysis.

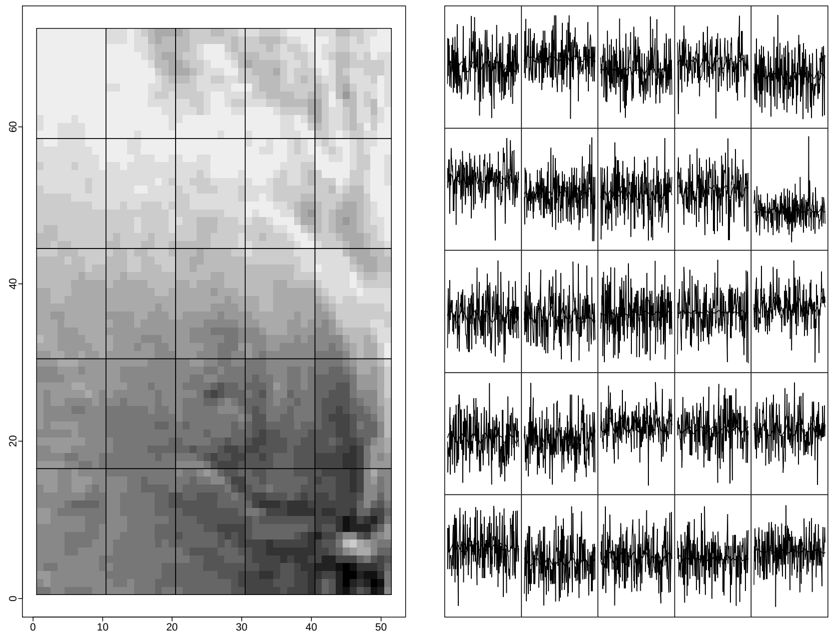

Figure 2.3: *An anatomical brain scan is divided into subregions and a representative time course from each subregion is presented on the right.*

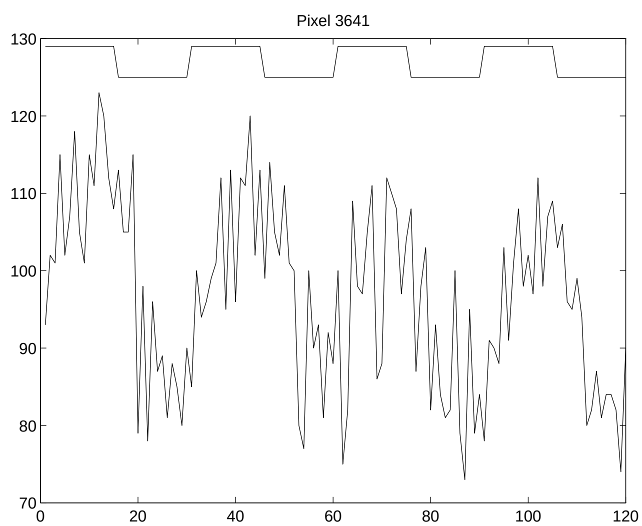

Figure 2.4: *The time course from a single pixel is displayed. Superimposed is the stimulus regime, with the higher level indicating stimulus and the lower rest.*

2.3 Statistical techniques used for fMRI data to date

Statistical methodology in the fMRI literature to date is summarized in the following list. Many of the methodologies follow from earlier work done with positron emission tomography (PET).

- ▷ Subtraction-based methods such as taking the difference between the average of all stimulation state images and the average of all rest state images effectively amount to performing pixel-wise t-tests.
- ▷ Correlation analysis with fixed frequency sinusoids matching stimulation period (Bandettini et al [3], Lee et al [46], Engel et al [13]), arbitrary phase shifts for each pixel location, Fourier analysis of pixel-wise time series, ranging from solely considering the fundamental stimulus frequency to considering all harmonics to using the entire power spectrum and complex phase information.

Figure 2.5: *Alternating periods of photic stimulation and rest at a fixed frequency stimulate the primary visual cortex. Identified here are pixels whose time-course behavior matched with a sinusoid at the stimulation frequency have correlation coefficients higher than 0.5*

▷ Nonlinear modeling of hemodynamic response function (using a parametric kernel convolution with the stimulation square wave) by Friston et al [31] and by Lange and Zeger [44].

▷ Friston's statistical parametric maps (SPM) and general linear model, including multiple comparison corrected z- and T-testing of image volumes.

▷ Use of singular value decomposition and other multivariate statistical techniques.

2.3.1 Subtraction methods

Considering Figure (2.4) we see that for each pixel in our example data set we have a time series and know whether a value was recorded during stimulus or rest. The simplest of statistical analyses is simply to compute an average of all values recorded during the stimulus state and an average of all values recorded during the rest state, take their difference, and perform a t-test to determine whether they are significantly different. Computed for each pixel in the image, we can compute what Friston [26, 30, 23, 25] has termed a *statistical parametric map*, or SPM, of t-values from each pixel value, and the areas of the SPM which exceed a certain threshold can be considered to be the areas of brain significantly activated by the stimulus. Figure (2.6) shows such a t statistic image computed for the data set of Figure (2.1).

For an early example of the use of subtraction methods, see Fox et al [17] who use subtraction images in a PET study to map human primary visual cortex. In their experiments, they injected radioactive oxygen-15 isotope tracers and performed paired 40 second PET scans under both treatment and control stimuli for three differently eccentric visual stimuli. They write: "Images of regional CBF change ... were formed by paired subtraction of control-state and stimulated-state CBF measurements." They achieve super-resolution localization of primary visual cortex due to

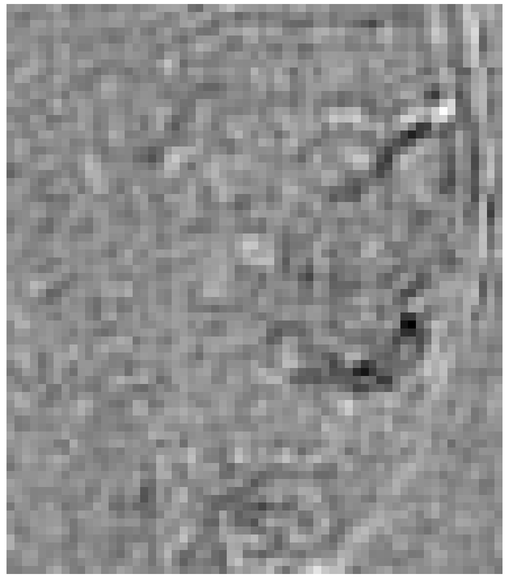

Figure 2.6: *An SPM{T} statistic image from the experimental data of Figure (2.1). The image is formed by taking the difference between the average image during stimulation and the average image during rest. Note the darker regions in the same locations identified in Figure (2.5). The bright pixel corresponds to a blood vessel.*

the focal nature of the activation. They analyze 17 treatment-control pairs obtained from 6 different subjects for three visual eccentricities (5 macular, 6 perimacular and 6 peripheral pairs). The paper does not make clear whether the ANOVA performed assumed independence between trials or was adjusted for repeated measurements within subjects.

In Fox et al [16], a similar PET study into human visual retinotopy is described, with each subject undergoing the same number of identically designed treatment regimes of stimulus and control. In this study, their statistical methodology is described as follows: "Comparisons between 2 conditions were made with a t test. Comparisons among more than 2 conditions were made using a 1-way, repeated-measures analysis of variance (ANOVA)". They use Bonferroni correction for multiple comparisons and a repeated measures Newman-Keuls procedure on statistically significant ANOVA results. Fox et al [15] describe identical statistical methods used in a PET study of human somatosensory cortex, with a similar experimental design of controls and multiple locations of cutaneous vibration, but missing data from one treatment for one subject.

Fox et al [18] discuss in some detail the advantages of utilizing replication in PET experiments. In this paper they discuss inter-subject averaging of PET images to improve the signal to noise ratio in functional brain mapping. They utilize a method of stereotactic anatomical localization to co-register PET images from multiple subjects as described in Fox et al [19]. They use 28 stimulus-control pairs and systematically assess the effects of averaging on groups of 2 through 5, 10, 15 and 20 images in a bootstrap-like procedure, taking 10 groups (to compute 10 average images) at each grouping size. They note that the signal loss they ascribe to residual anatomic variability "was much less than the noise suppression achieved, giving a marked rise in signal:noise ratio". This paper [18] also discusses local maximum sampling for localization of activation, use of control-control pairs to assess noise variability and characterize the noise distribution, and use of the Snedecor and Cochran ([58]) gamma-1 and gamma-2 statistics as measures of skew and kurtosis as test statistics for detecting significant activation. Their motivation for the use of gamma-2 ($\gamma_2 = \frac{\sum (x-\bar{x})^4}{n\hat{\sigma}^4} - 3$) as a test statistic was based on their observation that their noise distribution appeared platykurtic and their physiological responses induced leptokurtosis, and they cite Grubbs [37] on outlier detection as a similar application.

Belliveau et al [5], in a early fMRI paper mapping human visual cortex and comparing the mapping to those earlier PET studies, follow a similar methodology. They used a susceptibility-based contrast agent (a gadolinium-based compound) injected into the subject and took 60 images in 45 seconds, before, during and after contrast agent injection. They converted each voxel time course into a concentration-time curve of the contrast agent, and computed the area under those curves, corrected for recirculation, as a measure known to be proportional to local cerebral blood volume

(CBV). Having arrived at a value for CBV at each voxel location, and running such a sequence for both a rest and photic stimulation case on each of seven subjects, they obtain seven rest versus photic stimulation subtraction images. They use a paired t-test to note a $32 \pm 10\%$ average increase in CBV in the anatomically defined primary visual cortex, and use the subtraction images and a threshold of two standard deviations to estimate the extent of the activated cortex, ignoring activated extrastriate regions. No mention of use of a multiple comparison correction appears in this paper.

Belliveau et al [6] describe a similar (possibly the same) study in which they state: "Activated cortical areas are shown by subtraction of averaged baseline (resting) images from all subsequent images. Cine display of subtraction images (activated minus baseline) directly demonstrates activity-induced change in brain MRI signal." They also report the results of the non-invasive Ogawa et al [54] study and further report on a single subject experiment in which they perform a hemifield experiment to produce activation in left and right visual fields. In this latter experiment they present a checkerboard semi-annular stimulus flickering at 8Hz alternating between left and right visual hemifields every 20 seconds for several complete cycles, with an image acquired every 2 seconds. They describe their results as follows: "In this subject there was a 2% to 3% change in signal intensity with a 1% variation of the left and right ROIs during a single, unaveraged sampling interval. The observed significance for activation in the selected regions within V1 is $P = .03$ and the corresponding power is greater than .9 from an analysis of the activation map computed from an average of two time samples. With this fairly robust activation task, only two intra-subject averages (4 seconds of data) were required to obtain significance." Whether these two averages were randomly selected or chosen to maximize the difference is not reported.

The earliest papers describing the BOLD contrast in the Proceedings of the National Academy of Science both use subtraction methods. Kwong et al [42] state that

"Focal regions of cortical activation were revealed by magnitude subtraction of averaged baseline (resting) images from all subsequent images." They alternate through 20 scans of rest state, 20 of stimulus and repeat for a total of 80 scans. It is not clearly stated, but it appears that the baseline average to which they refer is based only on the first 20 scans for each subject. They perform paired t-tests on the responses from within the anatomically defined primary visual cortex region, and make no mention of multiple comparison corrections. Ogawa et al [54] describe that "When the sum of eight images acquired during a visual stimulation period are subtracted from the the sum of eight images without stimulation, a localized change is seen in the region corresponding to human visual cortex." They took a scan every 10 seconds for a total of 85 images, with visual stimulation present between images 20 through 30 and 50 through 65. It is not described how they choose the eight images for each average.

Frahm et al [20] describe a series of experiments investigating how changes in MRI parameters (voxel size and echo time) affect functional investigation of visual cortex. Their functional activation maps "were obtained by subtracting the average of 4 to 12 images in the dark state from a corresponding average of images in the activated state". Again, it is not clearly stated how the selection of images for these averages was made.

Menon et al [50] describe an fMRI experiment mapping primary visual cortex. Their results are presented as "difference images ... found by subtracting the baseline image from images obtained during the stimulation" which "represent the functional brain maps." They compute the baseline rest image apparently by averaging all dark (rest-state) images excluding "the very first few at the beginning of the study."

Bandettini et al [4] describe a set of experiments to show fMRI can detect human brain function during task activation. Experimental subjects were "instructed to touch each finger to thumb in a sequential, self-paced, and repetitive manner", and areas of primary sensory and motor cortex were identified. They computed brain activity maps by computing the correlation coefficient of each voxel time-course with what they called "an ideal response to the task activation." Specifically, they construct a vector containing +1 for task activation scan times and −1 for rest condition scan times, and compute an image of the correlation of each voxel time-course with that vector. Clearly this is a scaled subtraction method akin to computing a t-statistic SPM for the difference between stimulus and rest states using all images in the scan sequence.

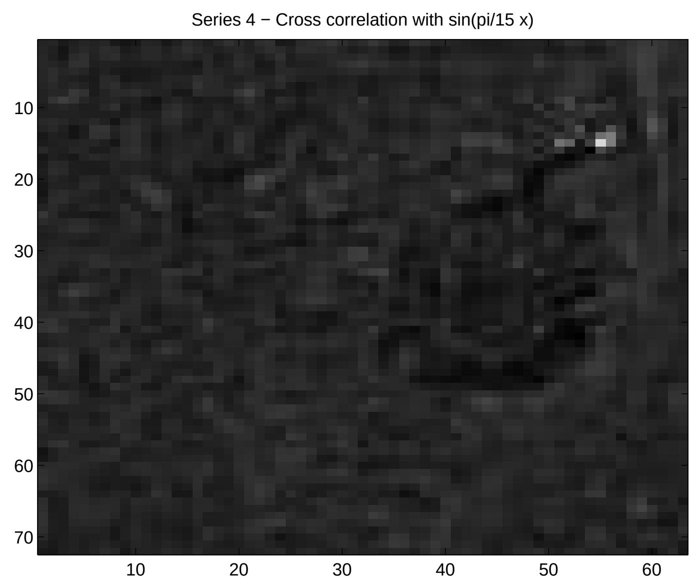

Figure 2.7: *The correlation coefficients between each pixel time course and a reference sinusoid at the experimental forcing frequency represented as an image for the experimental data of Figure (2.1). Note the darker regions in the same locations identified in Figure (2.5). The bright pixel corresponds to a blood vessel.*

2.3.2 Correlation and Fourier methods

Bandettini et al [3] criticize subtraction based techniques, pointing out that "Simple image subtraction neglects useful information contained in the time response of

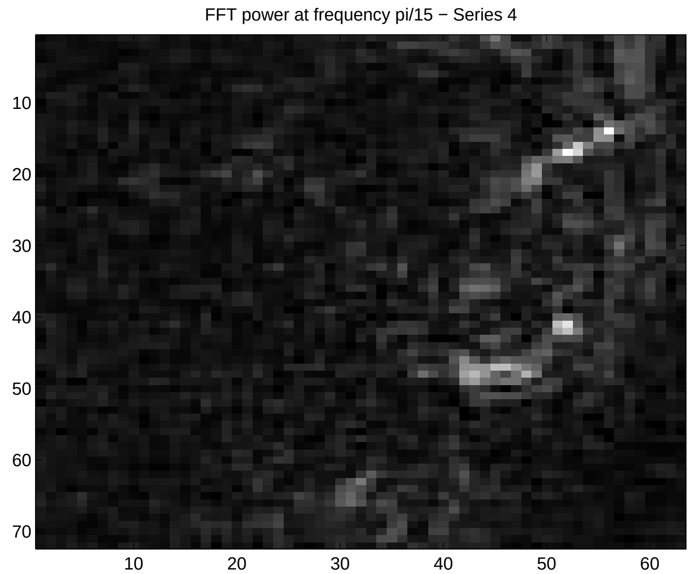

Figure 2.8: *The power of the pixelwise Fourier transforms at the experimental forcing frequency represented as an image for the experimental data of Figure (2.1). Higher spectral power in this figure corresponds to lighter shades. Observe the highest power is located in the regions identified in Figure (2.5).*

the activation-induced signal change, and is, therefore, an ineffective means to differentiate artifactual from activation-induced signal enhancement." They propose processing strategies based on thresholding correlation-derived images and compare image subtraction, temporal cross-correlation and Fourier methods on a motor cortex activation example. They describe what they call a Gram-Schmidt orthogonalization, which is basically least-squares, to remove linear drift as part of their data processing methodology.

In an experiment similar to that described in [4], the subject performs "self-paced sequential tapping of fingers to the thumb." A sequence of 128 images are taken every 2 seconds, with the subject switching between right and left thumb every 8 images. They then discuss the following processing strategies.

Subtraction. Selecting images 54 and 62 as "obtained during periods of peak signal enhancement for the right and left motor cortices, respectively", they form a subtraction image by taking their difference. They note "artifactual signal enhancement in the sagittal sinus region" which they attribute to CSF or blood flow.

Averaging with Subtraction. They subtract the average of images 52 through 57 from the average of images 60 through 65, and observe improved image quality but note "the artifactual signal enhancement in the sagittal sinus region is still present".

Cross-Correlation. By cross-correlation, Bandettini et al refer to the dot product of the pixelwise time-course with a reference vector. In effect, they describe it as the correlation coefficient between the time-course and the reference vector "unnormalized" to reflect amplitude information. As reference waveforms, they discuss use of a box-car "ideal" vector, an actual pixel vector, and a time-averaged response vector. As noted above, cross-correlation with the box-car waveform "is essentially the same as averaging all images during the interleaved "on" periods and subtracting the average of all images during the interleaved "off" periods." Use of a selected pixels waveform as the reference avoids the over-simplification of the "ideal" model the box-car represents, enhancing the image quality "because the reference vector more closely approximates the actual response vectors and the dot products are higher." The disadvantages are that artifacts may be present in the selected pixel and the issue of selection bias. The time-averaged response vector reference is constructed by averaging a pixels response across stimulation cycles. Again, selection bias and the potential presence of artifacts correlated with activation are of concern.

Fourier Transform. The magnitudes of the Fourier transforms of the time responses at the stimulation frequency display similar activation patterns to those produced using cross-correlation, including sagittal sinus signal enhancement.

Imaging Formation with Thresholding. Working with either time-domain or frequency-domain representations of the pixel waveforms, Bandettini et al observe that ignoring amplitude information and working with correlation coefficient images, spurious artifacts presumably associated with venous or cardio-vascular flow show minimal correlation. They propose thresholding of the correlation coefficient image. Threshold selection is based on an i.i.d. Gaussian assumption for the noise and takes no account of multiple comparisons. Formation of the cross-correlation image at pixel locations not excluded by correlation coefficient thresholding is their recommendation, providing "a high degree of contrast-to-noise."

McCarthy et al [48] provide an example of frequency domain Fourier analysis of an fMRI experiment designed to localize brain activation in response to moving visual stimuli. Rather than using an activation-rest paradigm, they compare two stimuli states they call Motion and Blink. "Stimuli were 100 small white dots randomly arranged on a visual display. During the Motion condition, the dots moved along random non-coherent linear trajectories at different velocities. During the Blink condition, the dots remained stationary but blinked on and off every 500 ms. The Motion and Blink conditions continuously alternated with 10 cycles per run and 6-8 runs per experiment. In half of the runs, the starting stimulus condition was Motion, while in the remaining runs it was Blink." A series of 128 images, taken every 1.5 seconds, of multiple brain slices were acquired for each run.

For each voxel time-series, the Fast Fourier Transform was computed, and only the frequency and phase at the stimulus alternation frequency were considered. Voxels were retained for analysis if they satisfied both a spectral power and phase threshold. The average frequency spectrum was considered to have a significant peak at the frequency of condition alternation if the spectral power at that frequency exceeded the mean spectral power of the fourteen neighboring frequencies (seven frequency points on each side of the stimulus alternation frequency) by more than 2 standard deviations for both stimulus orders (Motion-Blink and Blink-Motion). The stimulus alternation frequency was selected in such a way as to exactly correspond to a frequency of the FFT. The relative phases at the condition alternation frequency for the two stimulus orders were also required to be within ±15 degrees of 180 degrees out of phase in order to be retained. Voxels satisfying both conditions were considered to have been activated.

Binder and Rao [7], in an overview chapter on human brain mapping with fMRI, compare subtraction, Fourier and cross-correlation methods of analysis. They list several drawbacks of using Fourier analysis, including artifacts caused when "large magnitude signal changes create multiple peaks in the spectral density profile" and the lack of a "readily apparent" "acceptable method for testing the statistical significance of signal enhancement". They promote the cross-correlation procedure of Bandettini et al [3] using a sinusoidal reference waveform and make reference to multiple comparison correction in setting a correlation threshold to detect activation. A Bonferroni adjustment based on the number of pixels per brain slice and the total number of images in the series is used. Noting that "the normalized correlation coefficient contains no information concerning the magnitude of the signal change", they produce functional brain maps overlaying signal amplitude information for the pixels exceeding the correlation threshold over the anatomical scans.

Lee, Glover and Meyer [46] use correlation and phase lag information to discriminate venous responses from gray matter neural activity in an experiment stimulating primary visual cortex. For each voxel, they first subtract a "trend" line from the time course, and then compute the correlation coefficient between the detrended time course and both a sine and cosine function at the stimulus frequency. They combine these two correlation coefficients, effectively treating them as the real and imaginary parts of a complex correlation coefficient, to compute magnitude and phase components of the correlation. They compare the magnitude and phase images so computed with anatomical scans and a percentage signal deviation above baseline image computed using the Fourier power at the stimulus frequency and a baseline image. The temporal phase lags in their experiments correspond directly to hemodynamic time delays, with each 10 seconds of phase corresponding to a one second delay. They observe that shorter delays appear to correspond to gray matter activation, with longer delays attributable to larger blood vessels showing higher percentage signal deviations, often for low baseline levels of MR signal. They offer a variety of detailed physiological explanations for this behavior.

They find that for their data, "all the high percentage signal changes with $90 <$ phase < 110 degrees are almost certainly due to large veins" but that not all large veins are identified "due to the range of signal deviation at each phase." They also find pixels roughly 180 degrees out of phase with the hemodynamic delay ("indicating a signal decrease with stimulation") "anatomically correlated with large vessels." They quantify the uncertainty in temporal phase via a Fisher z-transformation and normal approximation, and use this to produce confidence band images of activations based on temporal phase (of the correlation).

They also compute similar correlation quantities from the NMR signal phase image. (MR images are typically based on signal strength contrasts or magnitude images, but the Fourier nature of image reconstruction can also provide a phase angle image of the scan region.) They observe that "regions with a high value of r_ϕ appear to be anatomically correlation with large vessels." They propose a technique for discriminating large veins based on comparison of both the phase and magnitude correlation responses. Both proposed methods of discrimination identified similar regions where larger blood vessels were anatomically expected, with the latter technique considered to be more sensitive to partial volume effects.

Engel et al [13] employ what is sometimes called a phase-delay encoding or phase-tagging experimental design in an fMRI experiment of human primary visual cortex. By virtue of the spatial organization of the primary visual cortex along the calcarine sulcus, whereby "as a stimulus moves from the fovea to the periphery the locus of responding neurons varies from posterior to anterior portions of the calcarine sulcus", they were able to map the retinotopic organization of the visual cortex with regard to increasing visual angle or eccentricity. Their stimulus was a pair of concentric expanding annuli displaying high-contrast flickering checkerboard patterns, with the rate of expansion constant and a new annuli emerging from the center as the outer annulus passed an experimental upper limit of visual field. Four cycles of the stimulus were viewed during a 192 second session, with an image taken every 1.5 seconds. They observe periodic response behavior at the expansion frequency (one cycle being the time taken for an annulus to expand from the center to the peripheral limit) along the calcarine with phase delay increasing from anterior to posterior areas. They use the average and standard deviations of the phase delays estimated for a best-fitting least squares sinusoid fit from each of the four stimulus cycles to compute an estimate of the cortical distance for which reliable separation of signal is possible using their technique. The retinotopic mapping between eccentricity and cortical distance produced is compared with previous studies and seen to provide superior detail and resolution.

De Yoe et al [10] describe an experimental analysis using cross-correlation techniques where the correlation threshold is determined by comparing the distribution of correlation values for both a blank series of scans (no stimulation) and the experimental scans, including a multiple comparison correction. They use a reference waveform based on ten representative pixels they consider to have been activated. One study described maps visual field eccentricity along the sulcus by performing a set of experiments with each run having a stimulus at a fixed eccentricity. They also report on a phase-tagged experiment identifying a circulating retinotopic pattern corresponding to visual field angle. The stimulus consisted of a flashing hemifield placed so as to rotate 18 degrees of visual field angle every two seconds. The mapping was computed "by performing a cross-correlation analysis with a set of reference wave forms spanning a complete range of phase shifts (0–360 degrees = 0–40 seconds)."

Sereno et al [56] produce a detailed mapping of the retinotopic organization for both polar angle and eccentricity using stimulus phase-encoding and Fourier analysis. Stimulus regimes of both expanding and contracting rings mapped eccentricity, while both clockwise and counter-clockwise rotating hemifields mapped polar angles. They describe their analysis as follows: "The linear trend and baseline offset for each pixel time series was first removed by subtracting off a line fitted through the data by least squares. ...The phase angle of the response at the stimulus rotation (or dilation-contraction) frequency at a voxel is related to the polar angle (or eccentricity) represented there. ...The phase angles were mapped to different hues ...". They note that their "basic procedure is closely related to correlation of the signal with a sinusoid function."

In summary, correlation methods appear attractive by providing a single number summary for each pixel time course, compared against a comparison reference curve. Selection of the reference curve remains a somewhat arbitrary choice. The box-car curve of the stimulus regime leads to a scaled t-statistic image. Correlation with smooth reference curves ignores high frequency noise in the time-course. Use of a reference sinusoid at the stimulation frequency or the Fourier spectral power at that frequency exemplify this approach. Use of the correlation coefficient without regard to the amplitude of the signal risks identification of spurious low-amplitude artifacts correlated with stimulation as regions of activation. Choice of a reference waveform based on selected pixels introduces the statistical issue of selection bias. Determination of the appropriate threshold above which correlation indicates activation is typically either done in an ad hoc fashion, with questionable distributional assumptions or without regard to the need for multiple comparisons correction. The temporal phase of the best-fitting sinusoid to a time course has a direct interpretation as a time-delay of activation.

2.3.3 Models for hemodynamics

Bandettini et al [3] write: "The activation-induced signal change has a magnitude and time delay determined by an imperfectly understood transfer function of the brain." Several authors have attempted to improve upon the understanding of that transfer function.

Kwong et al [42] model the rise time in recorded MRI signal intensity of activation using an exponential model: "$A(1 - e^{-kt})$ where t is time and A is a constant", reporting a mean time constant, k, of 4.4±2.2 seconds for gradient echo imaging and 8.9±2.8 seconds for inversion recovery imaging.

Friston, Jezzard and Turner [31] introduce the concept of the *hemodynamic response function* to describe the delay and dispersion that blood flow dynamics must contribute to the detection of brain function using fMRI. Their paper, one of the most influential in the fMRI literature, also introduces the use of a normalized cross-covariance SPM for determining regional activation, explores stationary Gaussian random field theory approximations for inference, and considers eigenimages or singular value decompositions of fMRI data. They propose the following model: "$x(t) = \Lambda(\epsilon(v) + z(t))$ where $x(t)$ is the observed signal at a particular time (t), $\Lambda(.)$ is a convolution operator modeling the effect of the response function, $\epsilon(v)$ is the evoked neuronal transient as a function of time from the stimulus onset (v) and $z(t)$ is an *uncorrelated* neuronal process representing fast intrinsic neuronal dynamics." In a form more familiar to statisticians, the operator $\Lambda(.)$ can be written as the convolution of its argument with a function h, the *hemodynamic response function.* The Friston, Jezzard and Turner model can then be re-expressed as $x(t) = (h * (\epsilon + z))(t)$. They then propose use of a Poisson form for the hemodynamic response function (to be convolved, for example, with the temporal treatment profile), with a single parameter λ globally describing delay, dispersion, and hemodynamic response function shape. This choice is based on mathematical convenience and parsimony. Note that $x(t) = (h*(\epsilon+z))(t) = (h*\epsilon)(t)+(h*z)(t)$ by linearity of convolution, and the Friston model is thus making a very strong statement about the nature of any additive noise in their model.

Their test statistic for detecting activation is a statistical parametric map of the normalized cross-covariance between the observed time course at each voxel and the convolution model fit based on the globally estimated λ. They heuristically argue that this forms a SPM{Z}, or Gaussian SPM, and use stationary Gaussian random field theory to propose inference methods. These are described in more detail in Chapter 7.

Lange and Zeger [44] consider experiments based on alternating periods of stimulus and rest of equal length repeating for several cycles. They propose use of all the harmonic frequencies of the stimulus square wave and a Poisson or Gamma model for the hemodynamic response function in a Fourier-domain method for iteratively estimating model parameters. They allow for spatially varying hemodynamic response parameters at each voxel. By transforming into Fourier space, the convolution of the hemodynamic response with the stimulus signal can be modeled as a (complex-valued) multiplication of the corresponding Fourier transforms of each. The Fourier transform of a square wave has power only at the fundamental frequency and odd-numbered harmonics of that frequency.

The Lange-Zeger model assumes the observed MRI signal time-courses at each voxel are formed by the convolution of their parametric hemodynamic response functions with the stimulus signal plus a noise term consisting of a stationary Gaussian process at each pixel. The parameters of the hemodynamic response function are iteratively fit using complex-valued least-squares in the Fourier domain on a pixel by pixel basis. Under the stationary Gaussian process assumption, the temporal auto-covariances translate into a diagonal variance-covariance matrix in the frequency domain with unequal variances. Having fit the parameters pixel-wise, inference is accomplished assuming a spatio-temporal autocovariance structure on the residuals and using multivariate Gaussian theory. Empirical spatial autocorrelation functions as a function of intervoxel spacing at the fundamental and harmonic frequencies were computed. These were approximated by exponential or Gaussian functions to construct a patterned variance-covariance estimate upon which to base inference. Isotropy was assumed.

Figure (2.9) displays at left the mean-corrected time-series observed at each of the 23 high-correlation pixels identified in Figure (2.5) on page 11 in the 16 by 16 pixel sub-region of Figure (2.1). On the right are the parametric reconvolution fits obtained by the Lange and Zeger procedure, using Fourier analysis and convolving a parametric Poisson kernel with a box-car waveform at the stimulation frequency. Only a single cycle of the fit is shown. Figure (2.10) offers the same display but shows the time-courses and fits for every pixel in the 16 by 16 subregion of Figure (2.1) on page 8.

Boynton, Engel and Heeger [8] conduct experiments on the response of human visual cortex to varying contrast stimuli. They quantify both visual cortex contrast response and hemodynamic response functions by means of linear systems analysis. They model the hemodynamic response function in a convolution model using a Gamma model.

These convolution-based hemodynamic response models are considered along with more data-driven modeling approaches in Chapter 3.

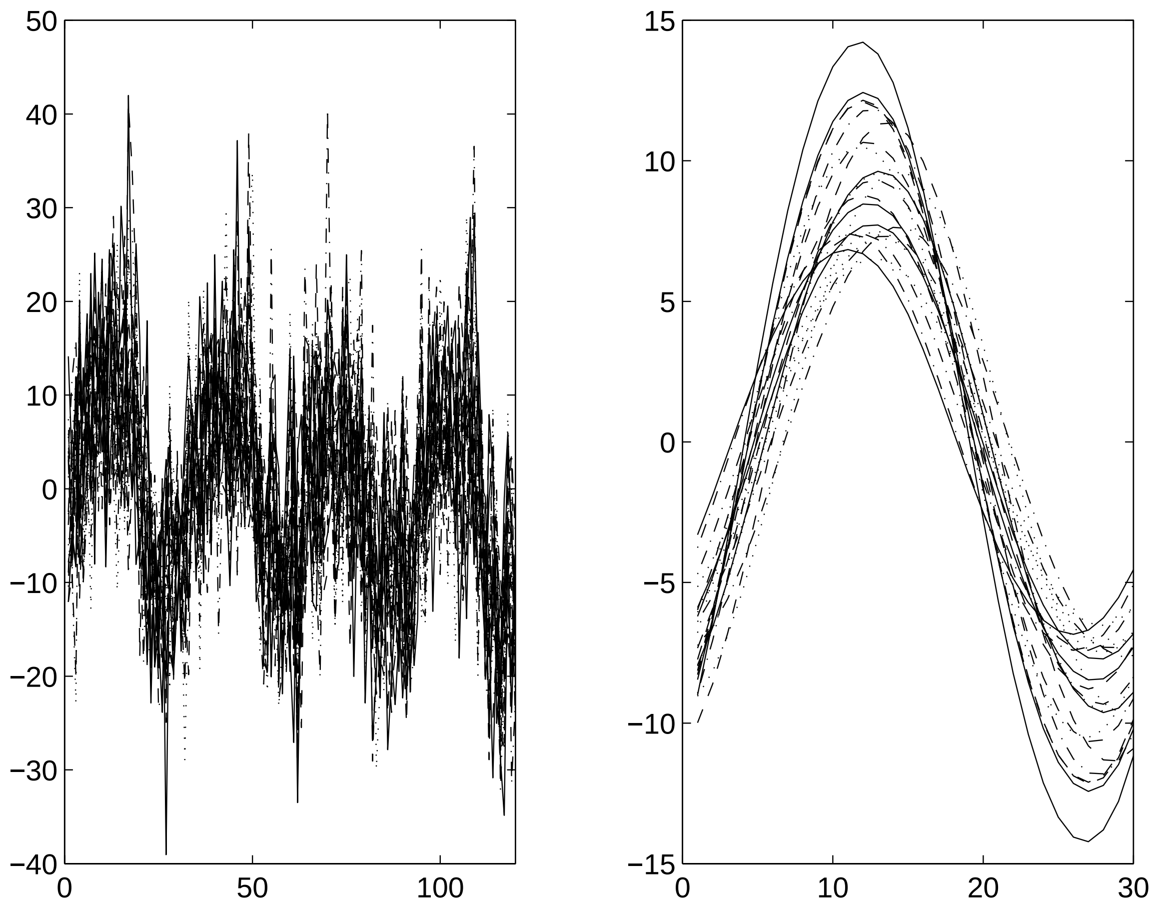

Figure 2.9: *Time-courses and one cycle of Lange-Zeger model fit for 23 high-correlation pixels of the subregion identified in Figure (2.1). Note the left and right panels are on different scales.*

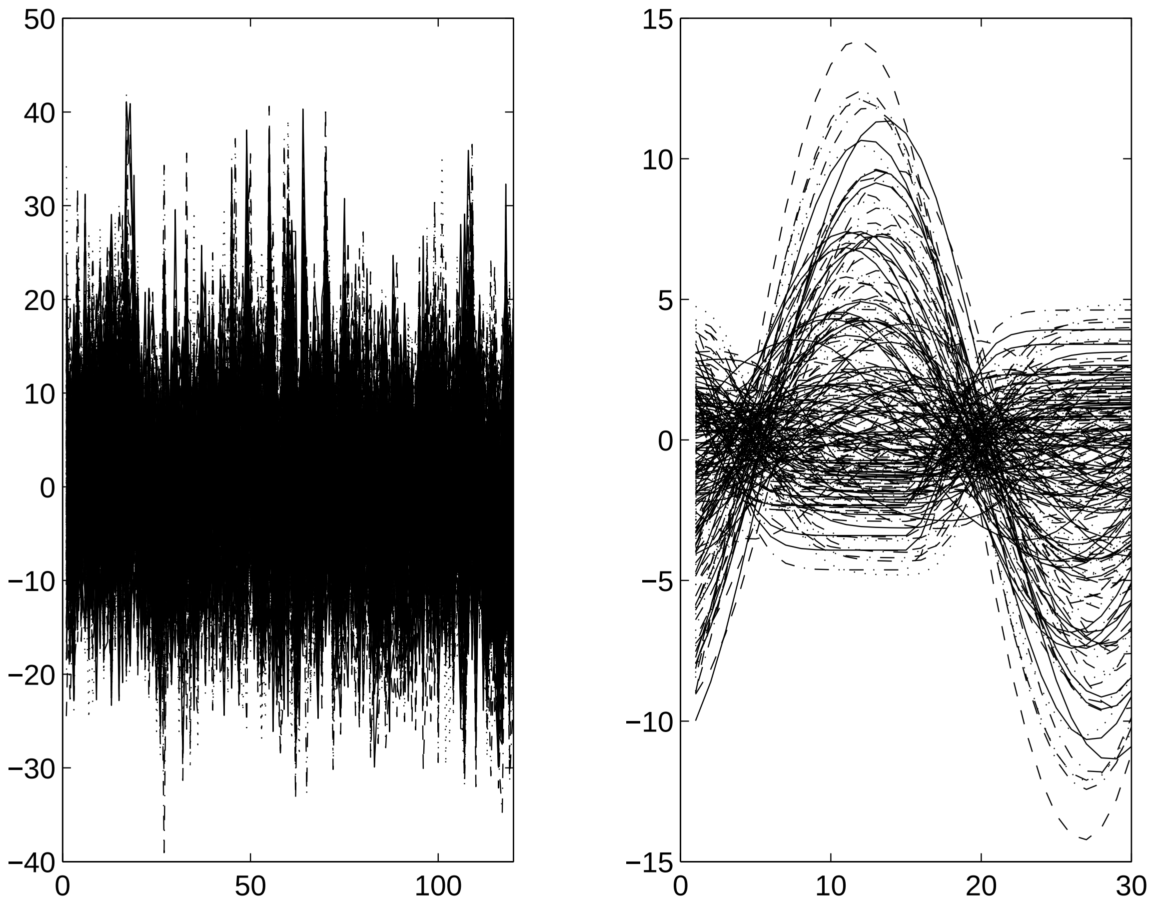

Figure 2.10: *Time-courses and one cycle of Lange-Zeger model fit for all pixels of the 16 by 16 subregion identified in Figure (2.1). Note the left and right panels are on different scales.*

2.3.4 Statistical Parametric Maps and the General Linear Model

The statistical parametric maps referred to above in the context of subtraction-based t-statistics at each pixel location can be generally applied to any pixel-wise or voxel-wise summary statistic, including the correlation coefficient. Friston et al [25] pioneered their use in PET data analysis with an analysis of covariance or ANCOVA model for activation accounting for global and local changes in rCBF flow.

In two papers by Friston and others [30, 29], of which the latter was subsequently corrected in Worsley and Friston [67], analysis of statistical parametric maps from functional imaging experiments is presented in the context of a general linear model. Not to be confused with generalized linear models or GLMs, the model proposed is simply a linear model for the expectation of a response variable. They observe that this presents "a framework that can accommodate any form of parametric statistical test ranging from correlation with a single reference vector, in a single subject design, to mixed multiple regression/ANCOVA models in many subjects."

They propose that the (mean-corrected) time series at any single pixel be modeled as $\mathbf{X} = \mathbf{G}\boldsymbol{\beta} + \mathbf{e}$. In Friston et al [30], the residuals $\mathbf{e}$ are said to be "a matrix of normally distributed error terms" and are treated as if independent for inferential purposes. Later papers such as Friston et al [29] and Worsley and Friston [67] impose a strict structure on them, arising from their hemodynamic model. Friston et al [29] hold as a central tenet of their work "that a more powerful test obtains if the data are smoothed in time" and apply smoothing with a Gaussian kernel as a preprocessing step before analysis. The width of the Gaussian kernel is determined by comparison to their expectations of the dispersion of the hemodynamic response function. To accommodate this temporal smoothing, they pre-multiply the linear model equation with a matrix representing the Gaussian kernel of specified smoothness. If $\mathbf{K}$ is that matrix, the new model is

$$\mathbf{K} \cdot \mathbf{X} = (\mathbf{K} \cdot \mathbf{G})\boldsymbol{\beta} + \mathbf{K} \cdot \mathbf{e} = \mathbf{G}^*\boldsymbol{\beta} + \mathbf{K} \cdot \mathbf{e}$$

The residuals from the fitted linear models are assumed to be independent and identically distributed normal variates with mean zero and variance σ^2, smoothed when multiplied by $\mathbf{K}$. They assume that after smoothing, "the correlations that manifest are stationary and (almost) completely accounted for by the convolution (or smoothing operation)." They claim that "the aim of this work is to extend the general linear model so that it can be applied to data with a stationary and known autocorrelation." Worsley and Friston [67] observe that "there is a large literature on this topic" and cite Watson [60] as one of the best early references.

Worsley and Friston [67] correct many of the results given in Friston et al [29], mostly providing mathematical rigor where heuristic arguments had been used and correcting errors. Friston et al [29] claim their parameter estimator for β optimally "maximizes variance in the signal frequencies relative to other frequencies". Worsley and Friston [67] note that the optimal estimator of β in that sense "is obtained by deconvoluting or unsmoothing the data by multiplying by $\mathbf{K}^{-1}$ and applying least-squares to the uncorrelated data $\mathbf{X}$" by the Gauss-Markov theorem. Instead, Friston et al [29] apply least squares to the smoothed observations, with an estimator

$$\mathbf{b} = (\mathbf{G}^{*\mathbf{T}}\mathbf{G}^*)^{-1}\mathbf{G}^{*\mathbf{T}}\mathbf{K}\mathbf{X}$$

where $\mathbf{G}^* = \mathbf{K}\mathbf{G}$. Worsley and Friston note this estimator "is unbiased and the loss of efficiency is more than offset by the gain in robustness". The variance-covariance matrix of the parameter estimates is given by

$$Var(\mathbf{b}) = \sigma^2(\mathbf{G}^{*\mathbf{T}}\mathbf{G}^*)^{-1}\mathbf{G}^{*\mathbf{T}}\mathbf{V}\mathbf{G}(\mathbf{G}^{*\mathbf{T}}\mathbf{G}^*)^{-1}$$

where $\mathbf{V} = \mathbf{K}\mathbf{K}^{\mathbf{T}}$.

Hypothesis testing in Friston et al [29] is framed in terms of single degree of freedom linear contrasts $\mathbf{c}$ of weights summing to zero. In their notation $\mathbf{c}$ is a row vector. The test statistic for a particular linear compound $\mathbf{c}$ is given by Worsley and Friston as

$$T = \mathbf{c}\mathbf{b}/(\mathbf{c}\boldsymbol{\epsilon}^{\mathbf{2}}(\mathbf{G}^{*\mathbf{T}}\mathbf{G}^{*})^{-\mathbf{1}}\mathbf{G}^{*\mathbf{T}}\mathbf{V}\mathbf{G}(\mathbf{G}^{*\mathbf{T}}\mathbf{G}^{*})^{-\mathbf{1}}\mathbf{c}^{\mathbf{T}})^{\frac{\mathbf{1}}{\mathbf{2}}}$$

where ϵ^2 is an unbiased estimator of σ^2. Worsley and Friston point out the estimator ϵ^2 suggested in Friston et al [29] is biased, and give the formula for the correct unbiased estimator as

$$\boldsymbol{\epsilon}^{\mathbf{2}} = \mathbf{r}^{\mathbf{T}}\mathbf{r}/\text{trace}(\mathbf{R}\mathbf{V})$$

where $\mathbf{r} = \mathbf{R}\mathbf{K}\mathbf{X}$ is the vector of residuals and $\mathbf{R}$ is the residual-forming matrix given by

$$\mathbf{R} = \mathbf{I} - \mathbf{G}^{*}(\mathbf{G}^{*\mathbf{T}}\mathbf{G}^{*})^{-\mathbf{1}}\mathbf{G}^{*\mathbf{T}}$$

The distribution of T above under the null hypothesis of no effect is approximately distributed as a t random variable with "effective degrees of freedom" ν. The correct formula for the effective degrees of freedom, given by Worsley and Friston, is computed to be

$$\nu = \frac{2E(\epsilon^2)^2}{Var(\epsilon^2)} = \frac{trace(\mathbf{R}\mathbf{V})^2}{trace(\mathbf{R}\mathbf{V}\mathbf{R}\mathbf{V})}$$

Having developed a theory producing SPM{T} maps for any linear compound hypothesis, they apply the theory of stationary spatial random fields to perform inferences about regions of significant activation. More details of inference procedures in this context are discussed in Chapter 7.

The purpose of the temporal smoothing introduced by the matrix $\mathbf{K}$ to obtain "a more powerful test" is analogous to what is done in analysis of variance of mixed or random effects models. In the case of multiple or nested error strata, the F-test performed in analysis of variance compares the sums of squares within the error stratum concerned, avoiding contributions to the overall sum of squares due to uncorrelated errors in the other strata. If the high frequency noise in the data is totally uncorrelated with the signal, temporally smoothing before analysis is increasing the power of a test within the low frequency error stratum.

2.3.5 Eigenimages, functional connectivity, multivariate statistics

Connectivity between brain regions is of particular interest in human brain mapping. The organization of brain function for particular mental tasks can be split between discontiguous cortical regions. A substantial amount of work has been done in the psychology literature on identifying such functional connectivity, typically using temporal coherence and cross-correlation as indications.

Friston et al [27] propose use of principal components analysis (PCA) of PET data to identify functional connectivity. They argue that each of the principal components of the correlation matrix of ANCOVA-adjusted rCBF data "represents a truly distributed brain system within which there are high intercorrelations." They also claim that since the principal components are orthogonal to each other, each system so identified is functionally unconnected.

To reduce the computational workload implied by eigenanalysis of the correlation matrix of thousands of image voxels, the authors propose a recursive PCA subroutine written in `matlab` to compute the eigenvalues and eigenvectors of the correlation matrix. Weaver [61] shows that use of the `eig` and `cov` functions in `matlab` executed twenty to forty times faster and had an error six to ten orders of magnitude smaller for data sets at usual imaging sizes.

Friston et al [27] compute the principal components based on a subset of their data consisting of the pixels for which significant differences between scans were detected by the ANCOVA F-test. Their first principal components accounted for 71% of the variance of the data and identified separated brain regions of activation corresponding to anatomically connected areas used for processing the verbal fluency task their experimental regime was testing.

In Friston et al [22], a distinction is made between what the authors call functional connectivity and effective connectivity. Functional connectivity is defined as "temporal correlations between spatially remote neurophysiological events", where effective connectivity is "the influence one neuronal system exerts over another." To quantify "the amount a pattern of activity ... contributes to the functional connectivity or variance-covariances observed", the authors propose use of the vector 2-norm of the pattern multiplied by the data matrix. They claim this metric combined with use of the singular value decomposition "has been used to good effect in demonstrating abnormal prefrontotemporal integration in schizophrenia" in Friston et al [32]. Noting that the singular value decomposition produces singular vectors which are the eigenvectors of their functional connectivity matrix, they put forth the SVD as a way of "decomposing a neuroimaging time-series into a series of orthogonal patterns that embody, in a step-down fashion, the greatest amounts of functional connectivity." The eigenimages show spatial arrangements of areas with high principal component values, accounting for large fractions of the variability in the data.

The authors further borrow from multivariate statistics and mathematics, proposing application of multidimensional scaling to map anatomy into functionally similar regions, use of mutual information as a metric for intrahemispheric integration comparisons between normal and schizophrenic patients, and undetermined linear modeling of effective connectivity as well as exploring nonlinear models.

A number of authors, such as Gonzalez-Lima [36] and McIntosh and Gonzalez-Lima [49], propose use of structural equation modeling to quantify the interactions among neural regions. This methodology, also known as path analysis, has been roundly criticized by many statisticians, an excellent example of which is the paper by Freedman [21] and its many responses and rejoinders.

Friston et al [24] propose a MANCOVA test based on their general linear model framework described above. Specifically, they compute a Wilk's Lambda statistic to test the null hypothesis that β is zero, and transform it so it has an approximate χ^2 distribution. The degrees of freedom they indicate is based on the incorrect effective degrees of freedom proposed in Friston et al [29]. Once testing via MANCOVA indicates a significant effect, they suggest use of canonical variate analysis (CVA) to characterize the spatiotemporal dynamics of the effects. In an example data set, the canonical variates identify transient brain responses that are biphasic and have differential responses based on the activation condition. The authors describe canonical images as "de-noised eigenimages that are informed by (and attempt to discount) error."

2.3.6 Summary

Statistical techniques used in the analysis of PET and fMRI data sets for human brain mapping experiments run the gamut from subtraction-based t-tests and correlation coefficients to linear and non-linear modeling approaches and use of multivariate statistical tools. The literature review in this chapter has focused primarily on techniques of estimation and modeling, and detailed consideration of issues involved in statistical inference for such data sets are deferred to Chapter 7. As is seen by the range of statistical techniques considered in the literature, the scope for involvement of professional statisticians is vast.

Characteristic features of the analysis of fMRI data sets include high dimensionality, relatively short time series at each pixel, the low signal to noise ratio of the BOLD effect, the need to reduce variability through replication and intersubject averaging, experimental design issues such as selection bias and treatment presentation orderings, and multiple comparison corrections. For the most part, these have been reasonably well considered in the literature. Many of the papers carefully validate their methodological proposals with experimental data. The enthusiasm with which papers describing the application of such a great variety of statistical techniques may be of potential concern to statisticians. Several of the papers by Friston and others have contained mathematical errors or inaccuracies due to heuristic or ad hoc reasoning [26, 29] or inefficient, inaccurate implementations proposed as innovative [27]. The same authors have proposed some of the landmark papers in the fMRI literature, for example [26, 31]. Some of the assumptions taken for granted in this literature may benefit from closer inspection by statisticians. The nature and adequacy of the convolution model proposed by Friston, Jezzard and Turner [31] and the generalization proposed by Lange and Zeger [44] will be examined in the next chapter. The statistical rigor of inference techniques, discussed in Chapter 7, is substantial. Several aspects of fMRI data analysis pre-processing could bear further examination, but are not considered here. These include issues of image registration, ensuring subsequent images are aligned and comparing the same brain regions. The pre-eminent technique used in the literature is that proposed by Woods et al [63], although Friston et al [33] propose their own technique. The detrending of the time course drift prior to analysis rather than as part of analysis may be introducing systematic bias and reducing the effectiveness of subsequent statistical analysis. Finally, the temporal smoothing proposed by Friston et al [29, 67] with a Gaussian kernel is quite a strong presumption on the form of the temporal auto-correlation of the data, and more data-driven analysis of its appropriateness would be of interest.

Chapter 3

Spline Basis Expansion Models

Identification of brain activity using functional magnetic resonance imaging (fMRI) depends on blood flow replenishing activated neuronal sites. In this chapter we describe how previous studies have sought to model the hemodynamic response function with restricted parametric models, and examine some of the inadequacies imposed by the implicit restrictions. A study of primary visual cortex activation is used to illustrate. We investigate more flexible estimation of hemodynamic response using cubic spline basis expansions, both of the time course itself and of the hemodynamic response function. In the latter, a deconvolution estimate of hemodynamic response is estimated from the observed fMRI time series at each pixel location and the designed temporal input stimulus. The estimated hemodynamic responses include both monophasic and biphasic forms, comparable with the more limited model proposed in Friston et al [28]. Bootstrapping allows us to show that for our example data, a Poisson-based convolution model performs no better than fitting a sinusoid to the data, but that for regions of activation, both the spline-based convolution model and periodic spline models do. Classification of hemodynamic response estimates using a statistical clustering algorithm such as Hartigan's k-means yields an informative brain activation map.

3.1 Time-course models at each pixel

While neural activation occurs on a millisecond time-scale, the detectable blood flow replenishing the activation sites depends on local vasculature and can occur as long as several seconds after activation. As discussed in the previous chapter, a number of models have been proposed by various authors to detect activation while allowing for that time delay and dispersion. The concept of the *hemodynamic response function* introduced by Friston, Jezzard and Turner [31] and their convolution model has been enthusiastically adopted in the fMRI literature.

To recap briefly, they propose the following model: "$x(t) = \Lambda(\epsilon(v) + z(t))$ where $x(t)$ is the observed signal at a particular time (t), $\Lambda(.)$ is a convolution operator modeling the effect of the response function, $\epsilon(v)$ is the evoked neuronal transient as a function of time from the stimulus onset (v) and $z(t)$ is an *uncorrelated* neuronal process representing fast intrinsic neuronal dynamics." In terms more familiar to statisticians, if h is the hemodynamic response function, their model is $x(t) = (h*(\epsilon+z))(t)$. They propose use of a Poisson form for the hemodynamic response function (to be convolved, for example, with the temporal treatment profile), with a single parameter λ globally describing delay, dispersion, and hemodynamic response function shape. Lange and Zeger [44] and Boynton, Engel and Heeger [8] use Gamma forms for the hemodynamic response. Lange and Zeger estimate parameters at each pixel location rather than globally, and use a frequency domain fitting procedure.

Friston et al [28, 24] observe that evoked hemodynamic responses can be biphasic or have differential 'early' or 'late' profiles. Friston et al [28] fit a model consisting of a linear combination of two monophasic parametric curves to explore this behavior, for example. They set the parameters of their model's monophasic curves to prechosen integers and do not estimate them using the data. The hemodynamic response functions admitted by their model take the form $Af_4(t) + Bf_{-1}(t)$ where

$$f_k(t) = \sin(\frac{\pi t}{n+1}) \exp(\frac{-t}{kn})$$

Figure (3.1) shows examples of the family of curves provided using $Af_{1/4}(t)+Bf_{-1}(t)$. The exponential modulation described by Friston et al [28] in the text of the paper does not seem to correspond to the curves displayed in Figure 1 of that paper, but using the reciprocal rate constants seems to correct this.

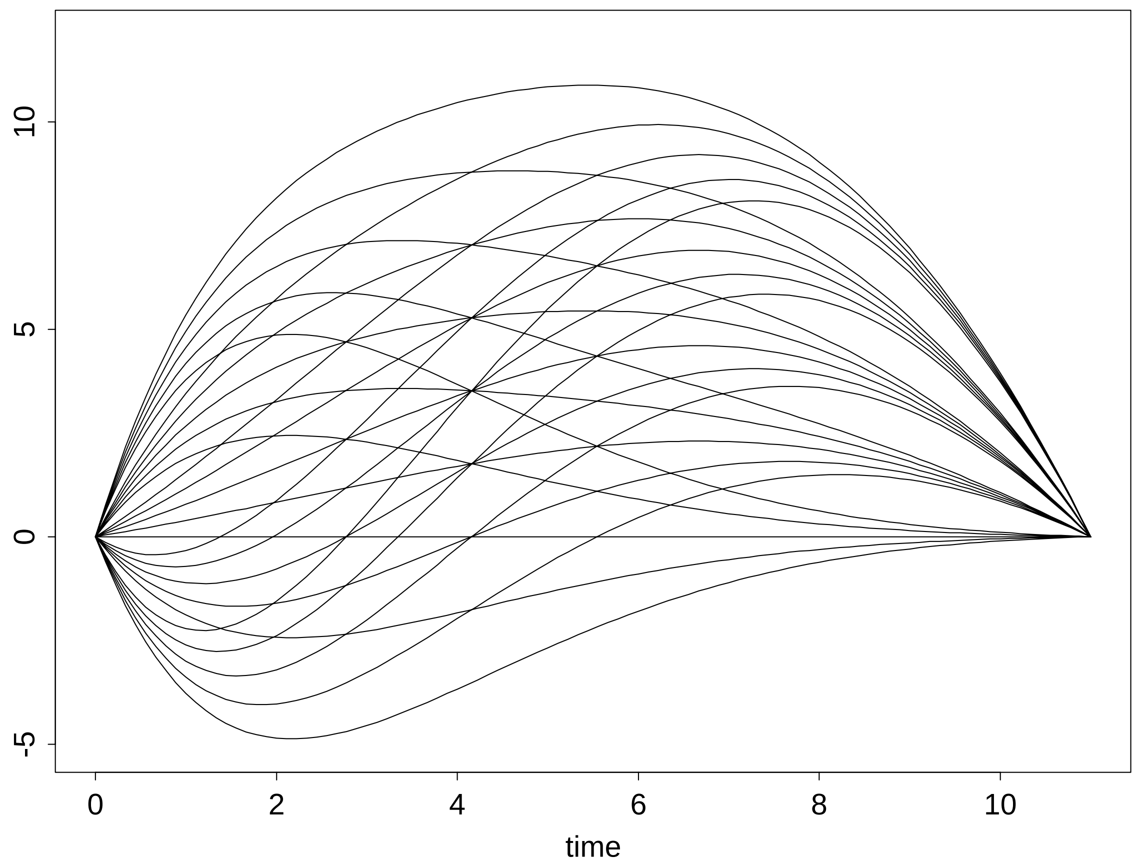

Figure 3.1: *An example family of curves admitted by the model proposed in Friston et al [28].*

An acknowledged deficiency of the Poisson model for the hemodynamic response function is that it confounds estimation of delay and dispersion into a single parameter. Given the possible inadequacy of the Poisson and Gamma models in describing the true nature of vascular delay and acknowledging the possibility of a biphasic response, we propose a more data-driven approach to the estimation of the hemodynamic response. In particular, by using a more flexible modeling family, such as cubic B-splines, we are able to discover more information regarding the local workings of the brain during activation such as sensory stimulation or cognitive function.

3.2 Data

Figure (3.2) shows an anatomical MRI scan of the human brain. This is an oblique slice taken as part of an investigation into activation of primary visual cortex, which is located along the calcarine sulcus. The area of activation for this particular experiment is expected to lie within the boxed 16 by 16 pixel subregion along the sulcus. The stimulus-rest regime for this series of functional scans consists of stimulus for 15 scans then rest for 15 scans repeated for 4 cycles, a total of 120 functional scans. An image was taken every 1.5 seconds in a continuous spiral scan.

Figure (3.3) shows an example of a time series produced at a single pixel, with the stimulus regime superimposed at the top of the figure. Note the delay between the change in treatment and the change in signal. Observe also the periodic nature of the response in an almost sinusoidal pattern. One common form of analysis of such data consists of thresholding a map of the correlation coefficient between the data at each pixel and the best fitting sinusoid at the stimulation frequency (with respect to amplitude and phase).

Figure (3.4) on page 42 shows the brain location of four selected pixels, whose mean-corrected time courses are shown in Figure (3.5) on page 43. Superimposed over

Figure 3.2: *An oblique anatomic MRI scan localized around the calcarine sulcus. The 16 by 16 subregion indicated contains the area of primary visual cortex being stimulated.*

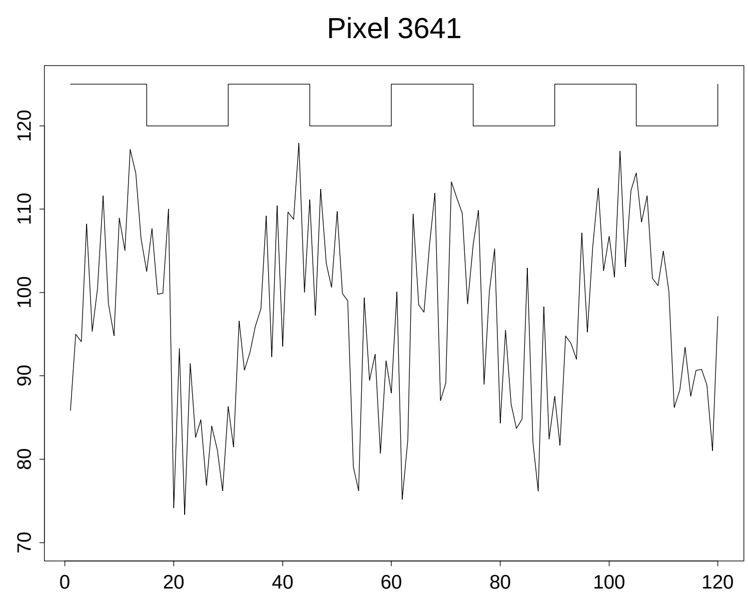

Figure 3.3: *The time course from a single pixel is displayed. Superimposed is the stimulus regime, with the higher level indicating stimulus and the lower rest.*

the data are the fits produced by the Lange-Zeger procedure using a Poisson model for the hemodynamics. We first model the observed data directly, using smooth cubic spline fits constrained to be periodic. Figure (3.6) on page 43 shows the fits produced at these four pixels for comparison, and Figure (3.7) on page 44 shows a single cycle of the periodic splines compared to the Lange-Zeger estimates at the corresponding pixels. Precise details of the periodic spline model are described in the next section.

Figure 3.4: *The location of four selected pixels used in examples below.*

A large number of the periodic spline fits of the data exhibited the characteristic double-humped pattern seen in these figures. Increasing the number of knots did not dramatically alter the overall shape of the fits, suggesting this pattern is not an artifact of the fitting procedure.

The structure apparently uncovered by these exploratory fits suggested that a

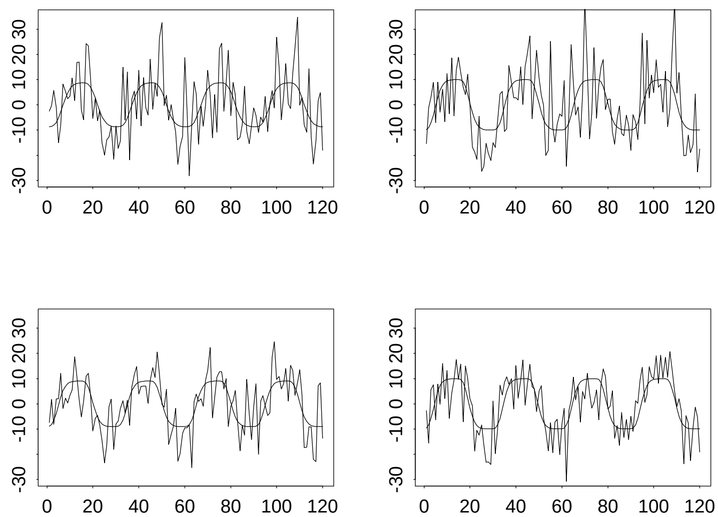

Figure 3.5: *Mean-corrected time-courses at the four pixel locations indicated above (in order from left to right). Superimposed over the time-courses are the parametric fits obtained using the Lange-Zeger procedure for those pixels.*

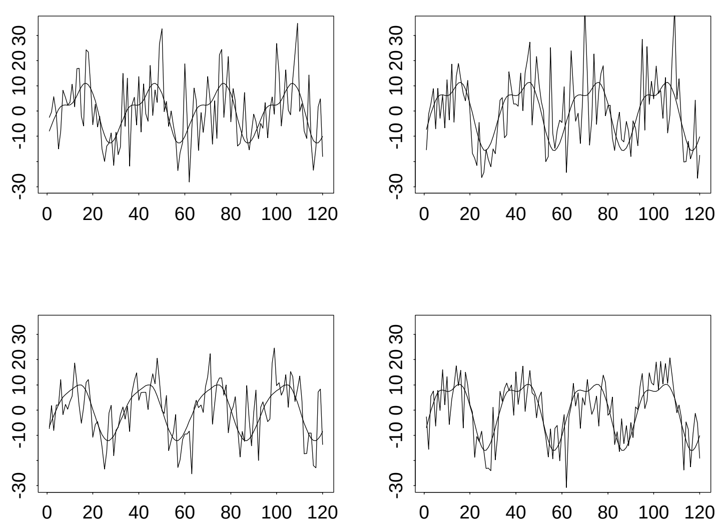

Figure 3.6: *The panels above show mean-corrected time-courses at the four pixel locations marked in black at left (in order from left to right). Superimposed over the raw time-courses are the fits obtained for those pixels using periodic cubic splines.*

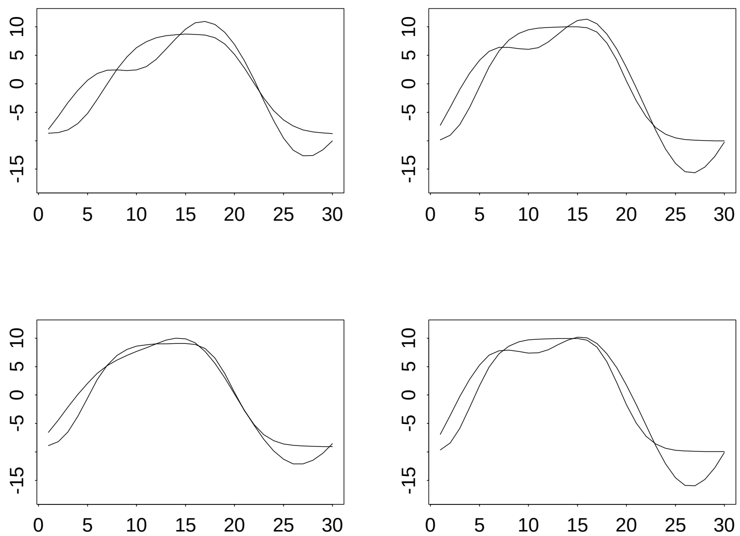

Figure 3.7: *The monophasic kernel fits of the Lange-Zeger procedure are shown compared to the periodic spline fits over a single cycle at each of the 4 pixels displayed above.*

more flexible model might provide greater insights into the nature of the brain's response to activation. The double-humped pattern seemed to be suggestive of a biphasic hemodynamic response. This led us to also consider cubic spline basis expansion models for the hemodynamic response function.

3.3 Models

The data we are considering consists of sequences or time series of image data. Let y_i denote the i^{th} image in the sequence, for $i = 1, \ldots, n$, where each image is viewed as a vector of $M = m_1 \times m_2$ pixels, where m_1 and m_2 are the image dimensions. If we consider the image data as an $n \times M$ matrix Y, each row represents the pixels of an image in the sequence, and each column represents the time-series of responses at individual pixel locations. The experimental subject is scanned while being exposed to alternating periods of rest and (in our example above, visual) stimuli. Encoding whether an image is scanned during a period of rest or stimulation, we can create a covariate $x(t)$, for which we have the discretized vector with $x_i = x(t_i)$ equal to one if the i^{th} scan occurs during the stimulus condition, or zero during rest.

3.3.1 Periodic Spline Model for pixel time series

Since the simplest commonality of the experimental regime is its periodic nature, we first explored the time-series using regression with periodic splines. Here we are modeling the observed data directly, with a model of the form

$$Y(t) = S(t) + \epsilon(t)$$

where $S(t) = \sum_{l=1}^{L} \theta_l b_l(t)$. The spline fits are based on equally spaced knots over a single cycle, constrained to periodically wrap at the cycle boundaries with continuous zero'th through second derivatives at the knots. The functions $b_l(t)$ for $l = 1, \ldots, L$ are the cubic B-spline basis functions for L equally spaced knots over a single stimulus-rest cycle, and the values of θ_l must be such that the periodicity constraints are satisfied.

The times at which fMRI scans are taken, t_i, are converted to new times $t_i^* = t_i \bmod T$. We can then construct a basis for cubic-splines periodic on $[0, T]$ as follows. We have available a function to generate a basis of cubic B-splines $b_l(t)$ with L equally spaced knots in $[0, T]$, and their derivatives [9]. Using their zero'th through third derivatives at 0 and T, we can obtain a linear constraint matrix C such that $C\theta = 0$ enforces the periodic boundary conditions on the parameters θ. We now construct a basis matrix B using the periodic time points t_i^*. Using C, we can reduce B to B^* by standard linear algebra [35], where B^* has the constraints built in. We can then regress the pixelwise series Y_{ij} at the j^{th} pixel on the columns of B^* for $j = 1, \ldots, M$.

The standard linear algebra mentioned above can be computed as follows. First we generate a basis of cubic B-splines with K uniformly spaced knots in $[0, T]$. Any knot spacing can be modified to a periodic basis; uniform spacing here merely reflects the least presumptive design. The cubic B-spline basis generates a linear function subspace of piecewise cubic functions which are smooth in the sense that their zero'th through second derivatives match at each knot. The columns of the matrix B are the values of the B-spline basis evaluated at the "wrapped" time points $t_i^* = t_i \bmod T$. In order to satisfy the additional requirement of periodicity, the functions must also satisfy the constraint that their zero'th through third derivatives at 0 and T must also match. Note the periodic interval endpoints are not necessarily knots, but the cubic polynomial in the first and last subintervals must wrap onto each other exactly. By matching the zero'th through third derivatives, we ensure the same cubic polynomial between the last interior knot and the "wrapped" knot above it, and similarly for the subinterval up to the first interior knot from the "wrapped" knot below it. Alternatively, we can force the periodic boundary to be a knot by only matching the zero'th through second derivatives. Now we have a basis B for the cubic B-splines space, but wish to further restrict ourselves to a linear subspace determined by the constraint $C\theta = 0$, which amounts to a basis for the null space of C. Note that $R(A)^\perp = N(A^T)$. By taking the QR decomposition of C^T, we obtain an orthonormal matrix Q_{const} such that the first 4 columns are an orthonormal basis for $R(C^T)$ and the remaining columns form a basis for $R(C^T)^\perp = N(C)$. Our basis B^* can thus be constructed by removing the first 4 columns of BQ_{const}.

3.3.2 Hemodynamic Spline Model for response function

Alternatively, we consider the time-series Y_{ij} at the j^{th} pixel, for $i = 1, \ldots, n$ as formed by the convolution of an unknown hemodynamic response function, $h(t)$ with the stimulus regime, $x(t)$,

$$Y(t) = \mu + (h * x)(t) + \epsilon(t).$$

Noting that we are sampling at uniformly-spaced time points, we denote by the vector h consisting of $h(k)$ for $k = 1, \ldots, K$, the discretized version of $h(t)$ evaluated at the uniform time points. Since we allow for distinct hemodynamic responses at each pixel location, we may refer to the hemodynamic response at the j^{th} pixel using $h_j(t)$ or its discretization h_j. We can write the convolution model as

$$Y_{ij} = \mu + \sum_{k=1}^{K} x_{i-k+1} h_j(k) + \epsilon_i$$

for each pixel j and time index i. Negative subscripts are to be understood as zero value quantities. (Cyclically lagging values modulo n may be an appropriate alternative in some circumstances.) This can be rewritten in vector form as

$$Y = \mu + Xh + \epsilon$$

if the $n \times K$ matrix X is constructed with first column given by the vector x described above, and lagged versions of x in subsequent columns. (If h is periodic, cyclic lags may be quite appropriate, especially in the typical case where initial scans are not considered while the magnetic resonance relaxation effects of tissue stabilize.) For physiological interpretation, we consider μ as a baseline (control) response level, and values of $h(t)$ to be non-negative, corresponding to how much blood is delayed by that amount. We can now readily estimate h using least-squares, subject to a non-negativity constraint. That is, at the j^{th} pixel, we minimize

$$Q_j = \sum_{i=1}^{n} (Y_{ij} - \mu - Xh_j)^2 \tag{3.1}$$

with respect to μ and $h_j(k), k = 1, \ldots, K$ subject to the constraints $h_j(k) \geq 0$ for $k = 1, \ldots, K$.

In this description so far, we have offered a similar formulation to the convolution models proposed by Friston et al and Lange and Zeger and others. They model h using a Poisson or Gamma model, which could be fit in this framework by iterative non-linear least squares techniques in the time domain. Friston et al [28] use a general linear model estimation procedure as described in Friston, Holmes et al [30, 29] and Worsley and Friston [67]. As described above, their model assumes an uncorrelated error term as the sole source of error which is added to the response function before convolution by the hemodynamic response. Lange and Zeger transform into Fourier space and perform what is effectively iteratively re-weighted non-linear least squares in the frequency domain. Friston, Jezzard and Turner [31] assume the same Poisson model across all pixels; Lange and Zeger allow for locally varying parameter values between pixels, and use Gamma models to allow for more general local variability.

Rather than presume a specific distributional form for h, we propose instead to estimate it semi-parametrically at each pixel location based on the time course data. Specifically, we propose a basis expansion model of the form $h(t) = \sum_{l=1}^{L} \theta_l b_l(t)$ and the values of θ are the parameters to be estimated. The studies referred to above use scaled probability distributions such as the Poisson and Gamma to estimate hemodynamic response. Given the physiological interpretation of h, we expect $h(t)$ to be "smooth" and suggest the use of cubic B-splines as basis functions to ensure smoothness without requiring a specific monophasic form. Note that the cubic B-spline basis functions $b_l(t)$ for $l = 1, \ldots, L$ are themselves probability density functions. Modeling $h(t)$ as $\sum_{l=1}^{L} \theta_l b_l(t)$, we are estimating the hemodynamic response as a non-negative linear combination of densities. The coefficients may be negative so long as the estimate of $h(t)$ is non-negative. The non-negativity constraint on h makes this a constrained but linear least-squares problem (note that the constraints are also linear). This is easily implemented using minimization software such as CFSQP [45]. We discretize the basis functions $b_l(t)$ for $l = 1, \ldots, L$ at the uniform points at which $h(t)$ was discretized into the columns of a matrix B. The minimization from equation (3.1) becomes: minimize

$$Q_j = \sum_{i=1}^{n} (Y_{ij} - \mu - XB\theta_j)^2 \tag{3.2}$$

with respect to μ and $\theta_j(l), l = 1, \ldots, L (L \leq K)$ subject to the constraints $B\theta_j(l) \geq 0$ for $l = 1, \ldots, L$.

3.4 Example

Examples of the periodic spline fits to several pixels were provided previously in Figures (3.6) and (3.7). They appear to follow the observed data quite well and detect structural patterns not observed in the comparable Lange-Zeger fits. Now observe what is found by the hemodynamic spline modeling approach.

Recall Figure (3.3) displaying the time series produced at a single pixel, with the stimulus regime superimposed at the top of the figure. In this hemodynamic spline model, we use a cubic B-spline basis with eight equally spaced knots on the interval [0,22.5] seconds. Recall that a scan is taken every 1.5 seconds and this corresponds to the time taken for 15 scans. Consider the construction of the matrix X in the hemodynamic spline model described above for this stimulus regime. Since the test and stimulus phases last 15 scans each, the columns become linearly dependent if

X has more than 15 columns. It follows that estimability of $h(t)$ beyond 22.5 seconds is not possible. Lange and Zeger cite Bandettini et al [3] and Friston, Jezzard and Turner [31] as estimating hemodynamic delays in humans to be roughly between 4 to 10 seconds so this should not be restrictive. We have chosen $K = 15$ and $L = 8$ for this example. Figure (3.8) shows examples of the estimates obtained at two example pixels. The two pixels selected are the upper two of the four displayed in Figures (3.5), (3.6), and (3.7). The estimated hemodynamic response function is shown on the left, and the fitted reconvolutions on the right. For each case the corresponding curves based on the best fitting Poisson model is displayed for comparison. There is little, if any, obvious improvement or difference in the reconvolution estimates. Also, neither convolution models' estimates appear to capture the double-humped pattern seen in the periodic spline fits.

The general shape of the estimates of hemodynamic response for these pixels are fairly close in gross features. In many cases the estimates obtained for the hemodynamic response closely follow the matching Poisson fit. Figure (3.9) below shows the corresponding estimates at two pixels with more obvious biphasic dissimilarity. Observe that the estimate of the hemodynamic response in each case is notably different, apparently uncovering structure not observed in the Lange-Zeger and Friston models. The reconvolutions, however, are still quite similar. The periodic spline fits (not shown) in these cases are similar to both convolution models and are more sinusoidal, not exhibiting the double-humped pattern noted previously.

3.5 Inadequacies of convolution models

To more clearly display the differences between the two convolution models and the periodic spline fits, Figure (3.10) below shows a single cycle of each model. The fits displayed are from the first example pixel in Figure (3.7), and the first example pixel

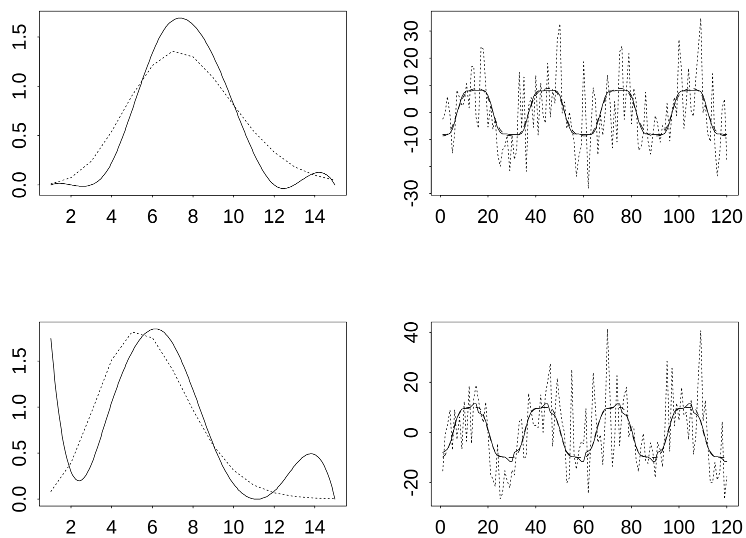

Figure 3.8: *On the left are estimates of the hemodynamic response function. On the right appear the convolutions of both our and the Lange-Zeger hemodynamic estimates superimposed over the original time series. For the hemodynamic response function estimates, the dashed line represents the Lange-Zeger Poisson estimate, and the solid line the hemodynamic spline model fit.*

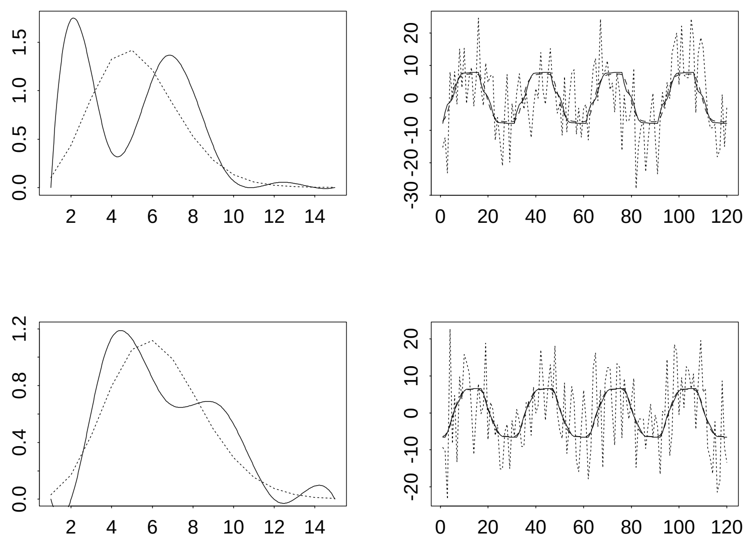

Figure 3.9: *On the left are estimates of the hemodynamic response function. On the right appear the convolutions of both our and the Lange-Zeger hemodynamic estimates superimposed over the original time series. For the hemodynamic response function estimates, the dashed line represents the Lange-Zeger Poisson estimate, and the solid line the hemodynamic spline model fit.*

in Figure (3.9). The convolution models are convolving a square wave with the hemodynamic response function. As the stimulus regime switches from rest to stimulus, the convolution accumulates a sum of the values of the hemodynamic response function. As it switches back from stimulus to rest, the accumulated terms are dropped in the order they were added. This accounts for the mirror anti-symmetry at the half cycle point of the convolution model reconstructions. This is more clearly seen in Figure (3.11), which shows one cycle of the hemodynamic spline model from the first set of panels of Figure (3.9).

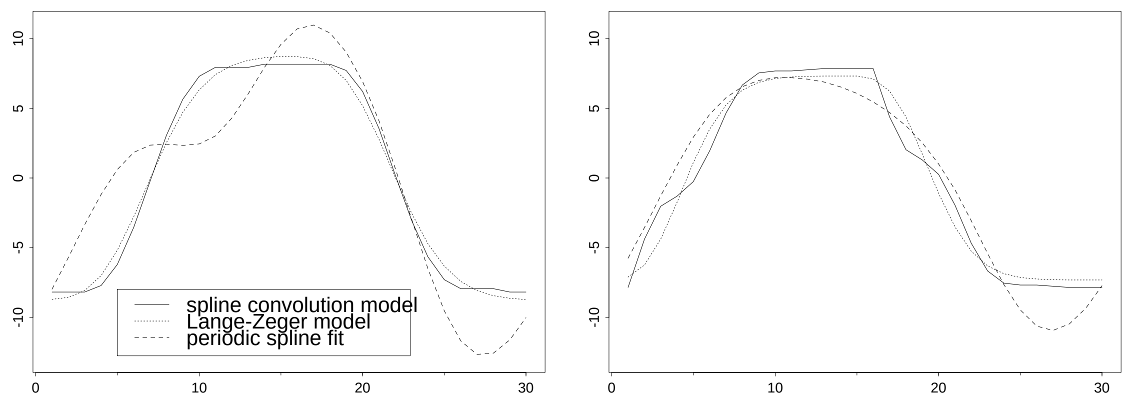

Figure 3.10: *A single cycle of each of the convolution models together with the periodic spline fit. At left are the fits from the first example pixel above. At right are the fits from the strongly biphasic pixel seen in Figure (3.9).*

The window smoothing nature of the convolution models for this stimulus regime also forces the estimates to be monotonically non-decreasing up to the half cycle and monotonically non-increasing in the second half cycle. By contrast, the periodic spline model is only constrained to be smooth and periodic and admits different rise and decay behaviors. An attempt to model the periodic spline fitted values from the first panel of Figure (3.10) with a convolution model was undertaken. Fitting a completely general hemodynamic response function produces the deconvolution and

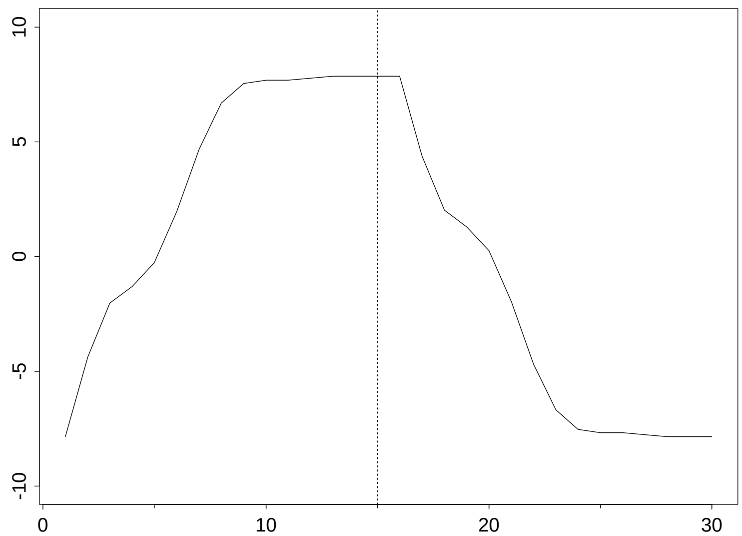

Figure 3.11: *A single cycle of the strikingly biphasic hemodynamic spline estimates reconvolution. Observe both the half-cycle monotonicity and mirror anti-symmetry.*

reconvolution estimates shown in Figure (3.12). Clearly the convolution model is unable to capture the types of patterns suggested by the spline fits to the data.

3.6 Signal or Noise?

Our investigations above suggest that the convolution models used by Friston et al and Lange and Zeger may be unfairly restrictive and fail to represent structure present in the data. It should be stressed that thus far these are empirical results, and we need to better understand the nature of variability of the the data. The observed time series can be quite noisy, and we need to be sure we are not just fitting models to the noise. Observe the closeness of the convolution model reconstructions which call into question how much signal information is contained in the data. The example power spectrum in Figure (3.13) on page 56 shows the periodogram of the time series

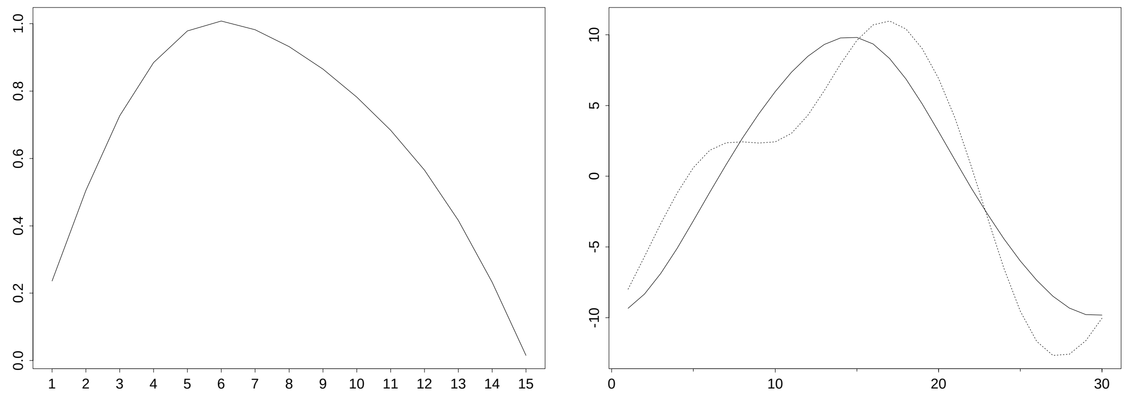

Figure 3.12: *On the left is the best deconvolution kernel estimate when attempting to fit the periodic spline fitted values using a convolution model. On the right are the curve being fitted and the convolution reconstruction. Note the convolution model is unable to adequately describe the shape of the spline fit.*

displayed in Figure (3.3) on page 41. Most of the power lies at the fundamental frequency. Implicit in the convolution model is the assumption that signal power at the off-harmonic frequencies comes from noise. The spectral power at the harmonic frequencies is of comparable magnitude to that seen at the off-harmonic frequencies. This poses the following questions: if all of the spectral power of the time series lies at the fundamental frequency, do any of these models perform significantly better than fitting a single sinusoid at that frequency, and can we statistically test if the pattern observed in the periodic spline fits is real?

3.6.1 Poisson convolution model versus sinusoid

Observing the power of the periodogram above to lie almost solely at the fundamental frequency, we consider first the question of whether fits produced by the Poisson-based convolution models are significantly different from simply fitting a sinusoid at the fundamental frequency to the time-course data. We do this via the bootstrap, using the residuals from the best sinusoidal fits at each pixel to form our resampling

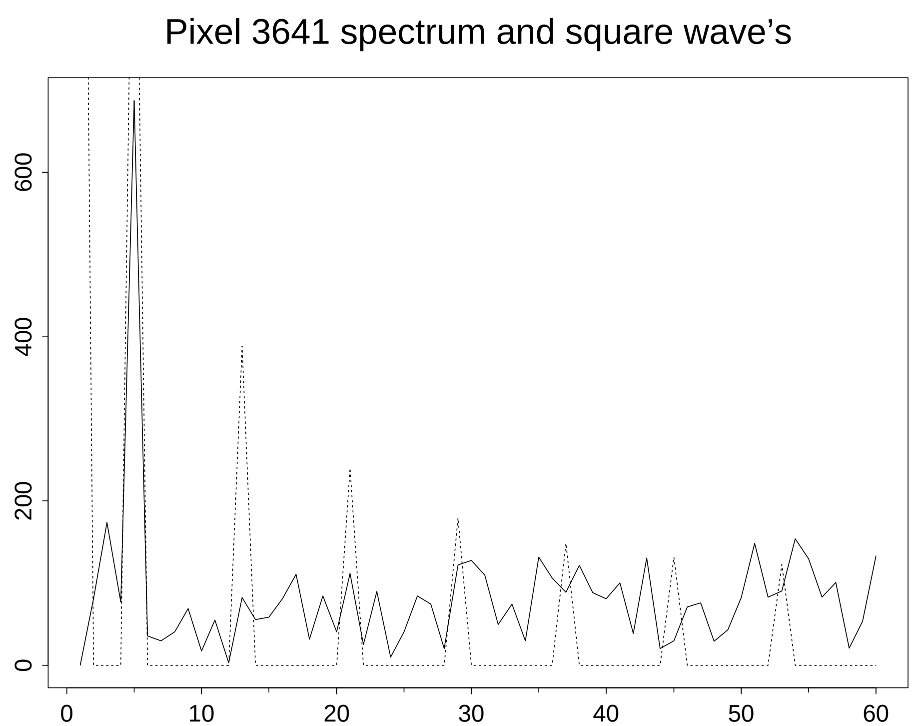

Figure 3.13: *The power spectrum of the first example time series above. Superimposed is a scaled periodogram of the stimulus regime square wave, whose power lies at odd harmonics of the fundamental frequency.*

distribution. Specifically, we compute the best fitting sinusoid at each pixel and then take each cycle of the residuals as our sampling units. In the example above, each pixel time-course thus yields four cycles of residuals. We resample in temporal blocks of cycles in an attempt to retain the temporal autocorrelation structure of the underlying data. Under the null model that the data are ideally modeled by a sinusoid at the fundamental frequency, these can reasonably be assumed to arise from the same distribution.

At each pixel, $B-1$ bootstrap resamples are formed by taking the best fitting sinusoid for that pixel, and adding a separate resampled residual cycle to each of its four cycles to produce a bootstrap simulated time-course. For each of these $B-1$ simulated time series, the best fitting Poisson-based convolution model for that data was computed. The Poisson-based convolution model fit of that pixels original data was included in this ensemble for comparison, and ranked based on a score quantifying the difference between the convolution model fits and the sinusoidal fit. If the pixels convolution fit lay within the top or bottom $\alpha B/2$ values, it can be considered significantly different from the sinusoid model.

To produce a single number summary of how different a convolution model fit was from the best sinusoidal fit, we took the best Poisson-based convolution model and the best sinusoidal model at each pixel, and took their difference across a single cycle. Both models are periodic, so a single cycle suffices for comparison. These difference curves or vectors were put in a matrix, and the singular value decomposition of that matrix was taken to determine linear combinations explaining the majority of the variability of that data which might be used to discriminate between the two models. Views of the singular vectors and the principal component loadings were examined seeking a good discrimination of these differences for pixels likely to be activated[1].

[1]Differences between the two models at pixels where no activation was detected could show greater difference of a distinctly different pattern than that accounted for in the first principal component.

These showed that the second principal component of the difference data is a good discriminator of time courses over pixels where activation appears to be present. Figure (3.14) below shows a single cycle of both the Poisson-based convolution model and best sinusoidal model at each pixel. Figure (3.15) below shows the projections of the differences between the fits onto the space of the first two principal components. Observe here that the second principal component discriminates between the fits well for the high amplitude pixels, while the first principal component is discriminating between pixels with low amplitude sinusoidal fits and convolution models that are effectively flat. Note that the monotonicity constraint implicit in the convolution models means that sinusoids with a wide range of phase delays may fit the data well but cannot be modeled by the convolution model. This appears to explain many of these cases.

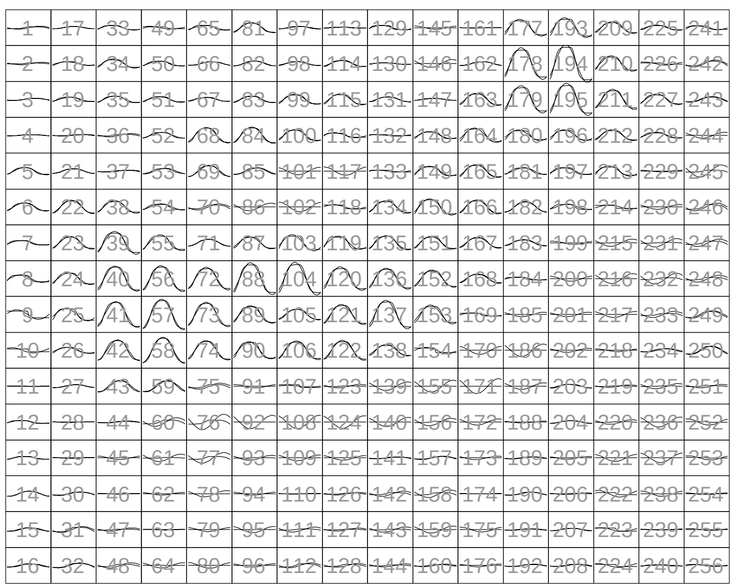

Figure 3.14: *At each pixel one cycle of the best Poisson-based convolution model and best sinusoidal fit are displayed. Pixels are labeled for comparison with later figures.*

These pixels typically had very noisy time courses and low amplitude sinusoid fits, which are not of interest here.

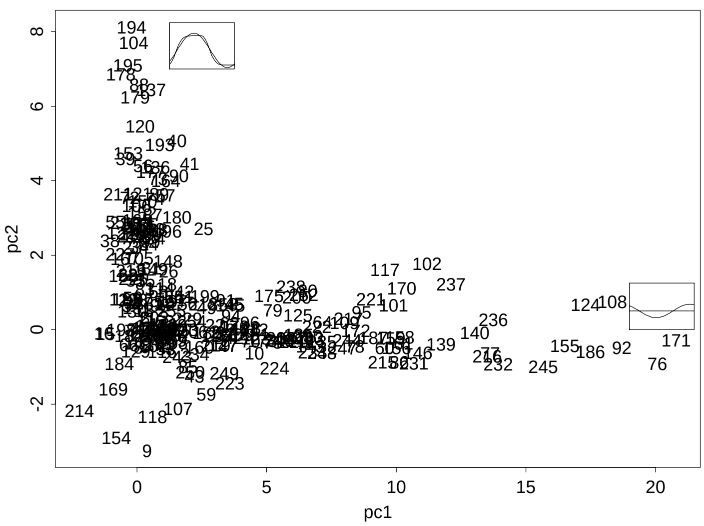

Figure 3.15: *First two principal components loadings of the difference between Poisson-based convolution and sinusoid models. Example shapes of both the sinusoid and Lange-Zeger model fits are shown for high values of each principal component. The first principal component is seen to discriminate the case where an out-of-phase sinusoid may fit the data while the convolution model estimate is flat, as it is unable to describe such data.*

Simulations as described above were performed with $B = 500$, computing the second principal component loading of the convolution model fits of each simulated time-course. The rank ordering of the score for the true data at each pixel within the B simulated values was recorded. Ranks close to 1 or B would indicate that the difference between the convolution model fit was so substantially different from the sinusoidal fit that it was unlikely to have arisen purely by chance. The 'one-sided' ranks out of 500 were converted to 'two-sided' ranks out of 250 indicating smallest rank from both extremes. Figure (3.16) displays these two-sided ranks at each pixel. Figure (3.17) displays the corresponding p-values. The smallest rank across the entire image is 62 out of 250, corresponding to a p-value of $\alpha = 24.8\%$, which is far from being statistically significant. Among the pixels with high-amplitude signals where activation is likely to lie, ranks in excess of 100 or 200 were common. We can conclude confidently that for this fMRI data set, fitting a Poisson-based convolution model offers no significant improvement over simply fitting a sinusoid at the fundamental frequency.

3.6.2 Periodic splines versus sinusoid

Having determined that the Poisson-based convolution model is no better than fitting a sinusoid at the fundamental frequency for this data, the question remains as to whether the double-humped pattern of the periodic spline fits are modeling signal or noise. We approach this question in an identical fashion, assessing whether periodic spline fits to the data are significantly different from fitting a sinusoid at the fundamental frequency.

At each pixel, $B-1$ bootstrap resamples were formed by taking the best fitting sinusoid for that pixel, and adding a separate resampled residual cycle to each of its four cycles to produce a bootstrap simulated time-course. For each of these $B-1$

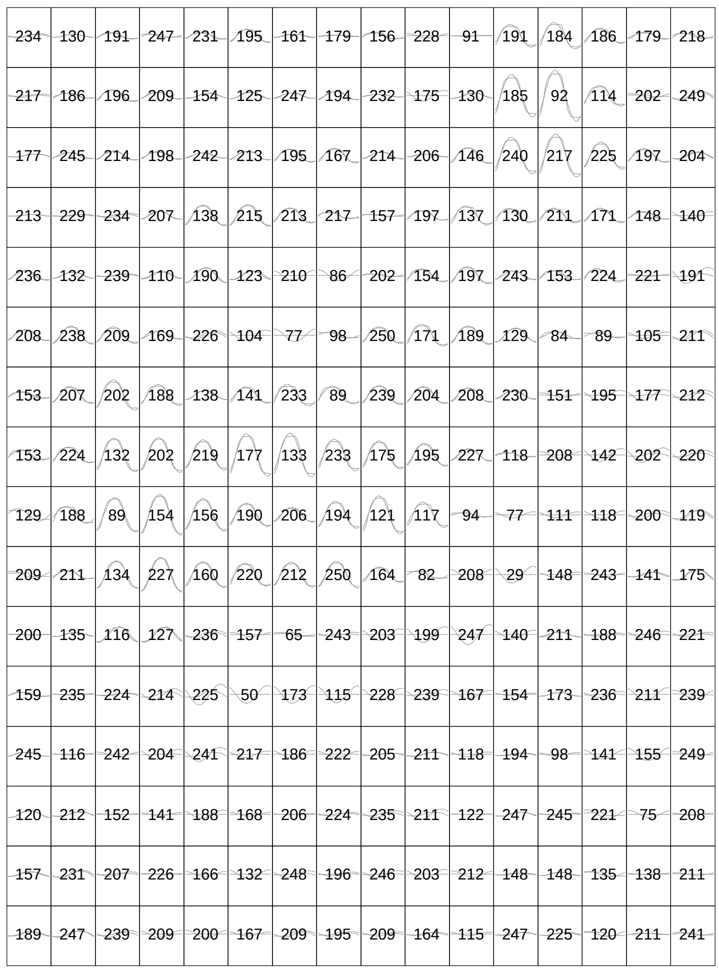

Figure 3.16: *Rank out of 250 obtained in bootstrap simulation comparing Poisson convolution model to sinusoid. Low values would indicate significance.*

0.936	0.52	0.764	0.988	0.924	0.78	0.644	0.716	0.624	0.912	0.364	0.764	0.736	0.744	0.716	0.872
0.868	0.744	0.784	0.836	0.616	0.5	0.988	0.776	0.928	0.7	0.52	0.74	0.368	0.456	0.808	0.996
0.708	0.98	0.856	0.792	0.968	0.852	0.78	0.668	0.856	0.824	0.584	0.96	0.868	0.9	0.788	0.816
0.852	0.916	0.936	0.828	0.552	0.86	0.852	0.868	0.628	0.788	0.548	0.52	0.844	0.684	0.592	0.56
0.944	0.528	0.956	0.44	0.76	0.492	0.84	0.344	0.808	0.616	0.788	0.972	0.612	0.896	0.884	0.764
0.832	0.952	0.836	0.676	0.904	0.416	0.308	0.392	1	0.684	0.756	0.516	0.336	0.356	0.42	0.844
0.612	0.828	0.808	0.752	0.552	0.564	0.932	0.356	0.956	0.816	0.832	0.92	0.604	0.78	0.708	0.848
0.612	0.896	0.528	0.808	0.876	0.708	0.532	0.932	0.7	0.78	0.908	0.472	0.832	0.568	0.808	0.88
0.516	0.752	0.356	0.616	0.624	0.76	0.824	0.776	0.484	0.468	0.376	0.308	0.444	0.472	0.8	0.476
0.836	0.844	0.536	0.908	0.64	0.88	0.848	1	0.656	0.328	0.832	0.116	0.592	0.972	0.564	0.7
0.8	0.54	0.464	0.508	0.944	0.628	0.26	0.972	0.812	0.796	0.988	0.56	0.844	0.752	0.984	0.884
0.636	0.94	0.896	0.856	0.9	0.2	0.692	0.46	0.912	0.956	0.668	0.616	0.692	0.944	0.844	0.956
0.98	0.464	0.968	0.816	0.964	0.868	0.744	0.888	0.82	0.844	0.472	0.776	0.392	0.564	0.62	0.996
0.48	0.848	0.608	0.564	0.752	0.672	0.824	0.896	0.94	0.844	0.488	0.988	0.98	0.884	0.3	0.832
0.628	0.924	0.828	0.904	0.664	0.528	0.992	0.784	0.984	0.812	0.848	0.592	0.592	0.54	0.552	0.844
0.756	0.988	0.956	0.836	0.8	0.668	0.836	0.78	0.836	0.656	0.46	0.988	0.9	0.48	0.844	0.964

Figure 3.17: *p-values from bootstrap simulation comparing Poisson convolution model to sinusoid. Low values would indicate significance.*

simulated time series, the periodic spline fit for that data was computed. The periodic spline fit of that pixels original data was included in this ensemble for comparison, and ranked based on a score quantifying the difference between the periodic spline fits and the sinusoidal fit. Similarly to the comparison of Poisson-based convolution models and sinusoid, the score quantifying difference was based on the first two principal components of the difference vectors of one cycle of the periodic spline fits and best fitting sinusoids. The first and second principal components together account for over 85% of the variability in this difference data set. The norm of the projection into the subspace spanned by the first two singular vectors was a good discriminator of the difference between the periodic spline fits, especially for the high-amplitude signal pixels. Figure (3.18) shows a plot of the norm of the projection into the subspace spanned by the first two singular vectors against amplitude.

Simulations as described above were performed with $B = 500$, computing the norm of the projection onto the subspace spanned by the first two singular vectors[2] of the periodic spline fits of each simulated time-course. The rank ordering of the score for the true data at each pixel within the B simulated values was recorded. Ranks close to 1 or B would indicate that the difference between the convolution model fit was so substantially different from the sinusoidal fit that it was unlikely to have arisen purely by chance. The 'one-sided' ranks out of 500 were converted to 'two-sided' ranks out of 250 indicating smallest rank from both extremes. Figure (3.19) displays these two-sided ranks at each pixel. Figure (3.20) displays the corresponding p-values. Several pixels among areas thought to be active show ranks lower than 20. In one case, the pixel is the most extreme of the bootstrap resamples. For those pixels, this is strong evidence of a significant difference between the periodic spline fits

[2]The singular vectors of the data set formed by taking the difference between the periodic spline fits and the best fitting sinusoid at each pixel. The singular vectors need not be recomputed with each bootstrap resampling

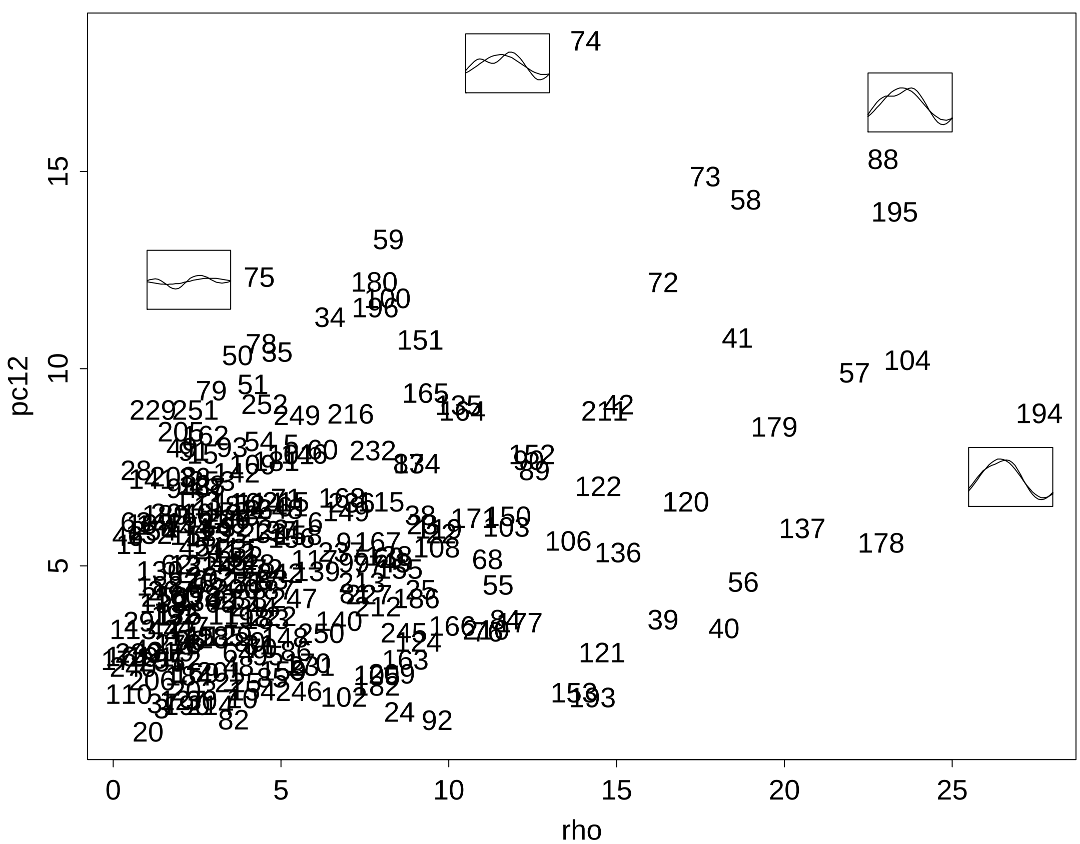

Figure 3.18: *Norm of the projection onto the subspace spanned by the first two singular vectors of the difference between periodic spline and sinusoid fits against signal amplitude. Examples of the shapes of both the sinusoidal and periodic spline fit are displayed near several of the extreme points.*

and fitting a sinusoid at the fundamental frequency. Moreover, significantly different pixels appear in spatially clustered regions.

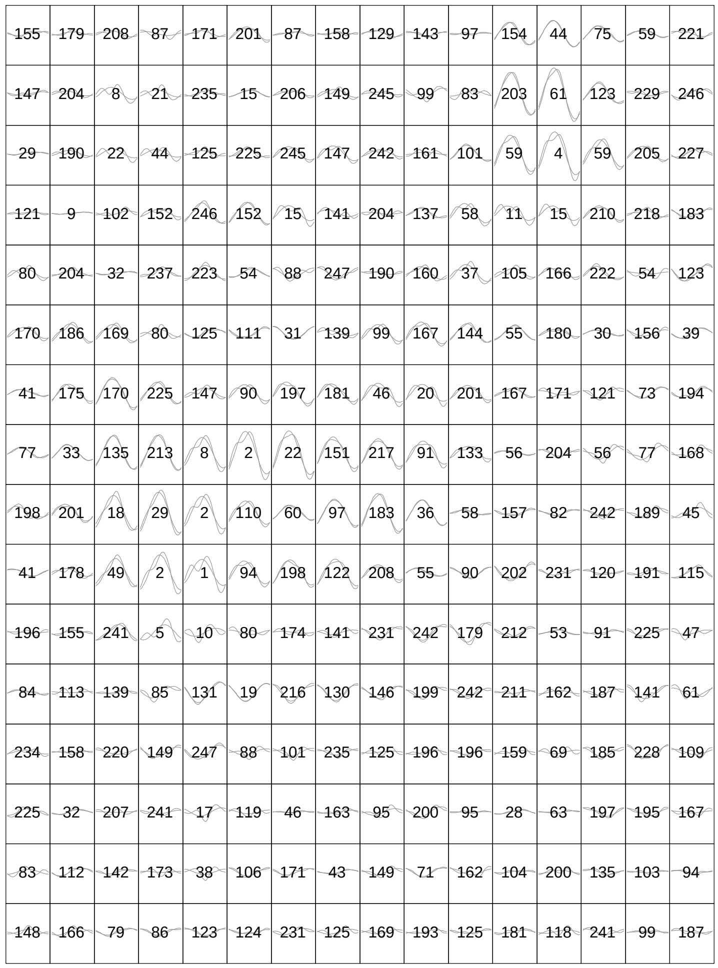

Figure 3.19: *Rank out of 250 obtained in bootstrap simulation comparing periodic spline fits to sinusoid. Low values would indicate significance.*

Validation with Friston, Jezzard and Turner data

To confirm that these results are not specific to the example data used, similar analyses were performed on a reference data set. Peter Jezzard provided us with access to a copy of the data analyzed in Friston, Jezzard and Turner [31], which is also the

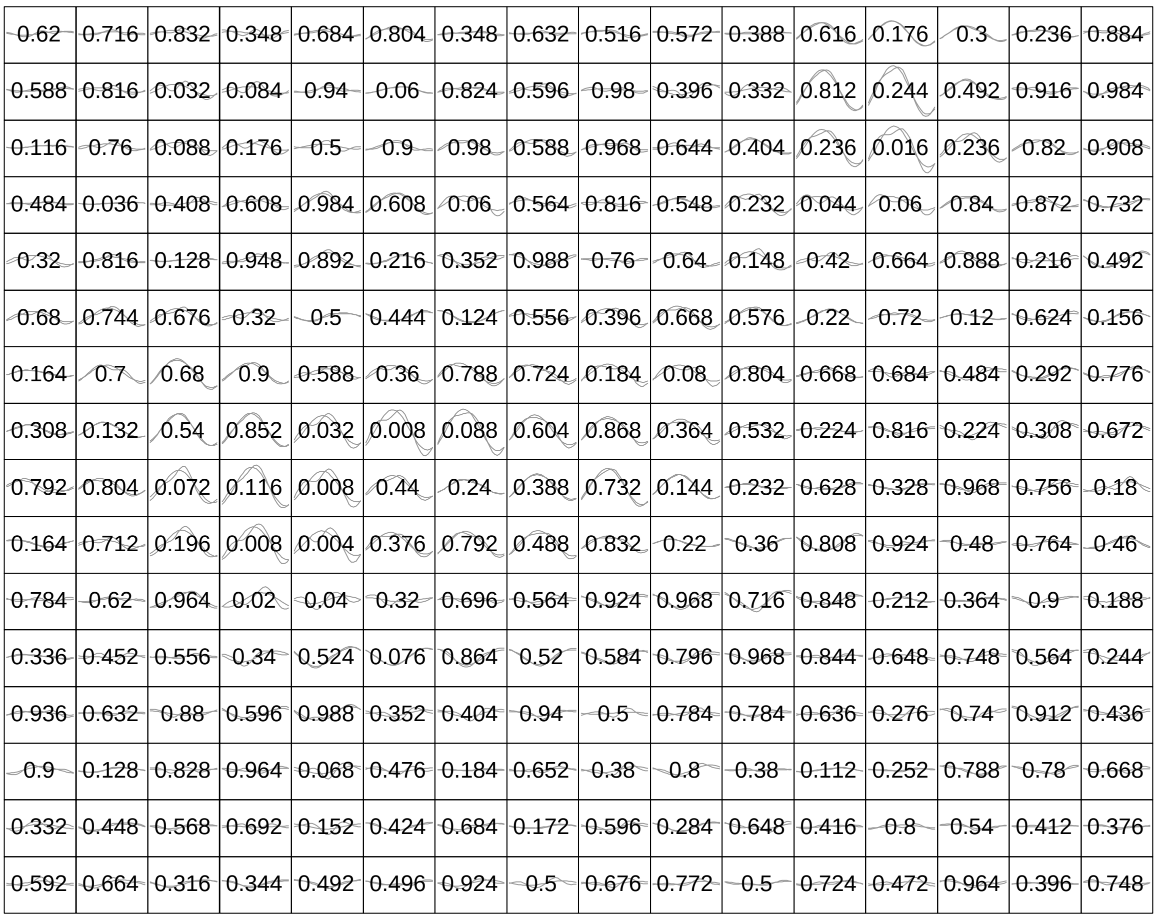

Figure 3.20: *p-values from bootstrap simulation comparing periodic spline fits to sinusoid. Low values would indicate significance.*

example data set used and described in Lange and Zeger [44]. The data from Friston, Jezzard and Turner [31] were taken from "a time series of gradient-echo EPI single coronal slices through the calcarine sulcus and extrastriate areas". Only the subset of the images containing the brain were considered, constituting an 18×30 voxel region[3]. The experimental subject was subjected to 30 seconds of photic stimulation followed by 30 seconds of rest, for three complete cycles, with a scan taken every 3 seconds. The image sequence thus comprises 60 scans, with three cycles of 10 stimulus images followed by 10 rest images. The average image taken over all experimental scans is displayed in Figure (3.21). The analysis in Friston, Jezzard and Turner identifies regions of activation for this data in "V1 bilaterally and (on the right) V2 dorsal to the calcarine fissure" as well as in "extrastriate regions in the fusiform, lingual, and dorsolateral cortices." Figure (3.22) below shows the difference image between an average across stimulus images and an average across rest images.

The same basic procedure as used in the previous sections was applied to the Friston, Jezzard and Turner data. At each pixel, $B-1$ bootstrap resamples were formed by taking the best fitting sinusoid to the pixel time course and adding a resampled residual cycle to each of its three cycles to produce a bootstrap simulated time course. For each of the $B-1$ simulate time series, the periodic spline fit to that data was computed, and the periodic spline fit to the pixels original data included in the ensemble. The original fit was ranked within the ensemble in accordance to a score which discriminating difference between periodic spline fits and the sinusoidal fit. For this data, the score used was the norm of the projection into the subspace spanned by the first three singular vectors, which showed good discrimination between the periodic spline fits and the sinusoids. The first three principal components

[3]Both the Friston, Jezzard and Turner paper and the Lange and Zeger paper consider interpolated voxels and consequently consider the same region of data at a resolution of 36×60 voxels. We chose to use the raw data resolution rather than the interpolated resolution.

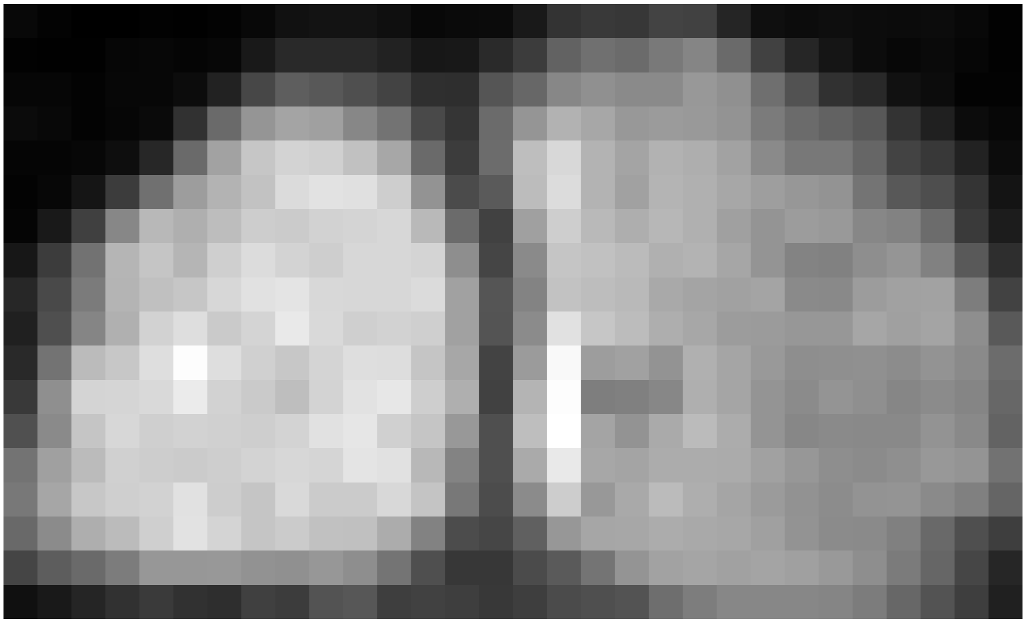

Figure 3.21: *Friston, Jezzard and Turner's fMRI data set, averaged across all experimental scans.*

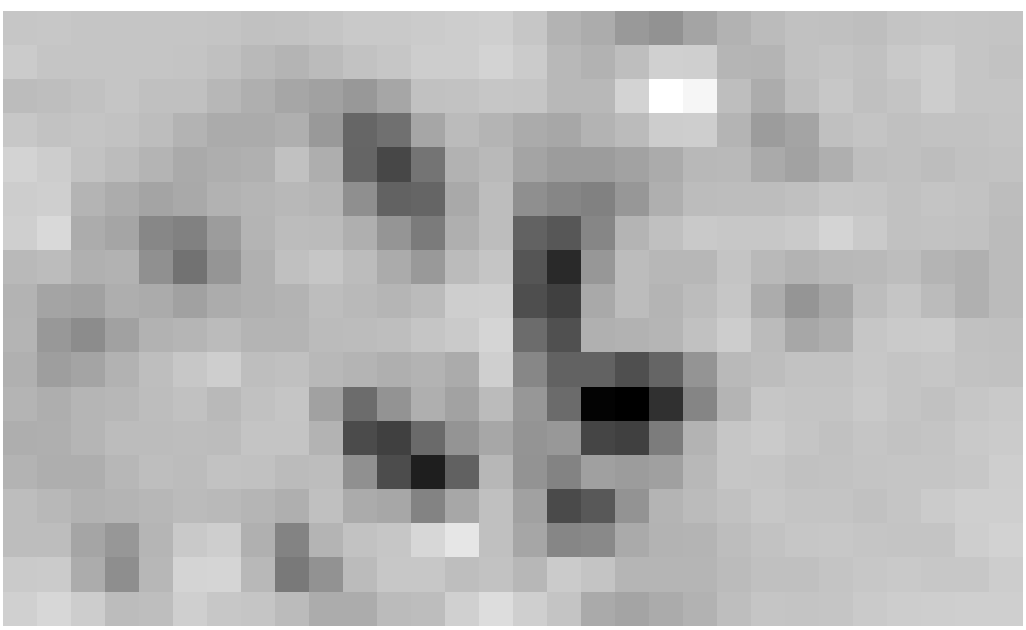

Figure 3.22: *Friston, Jezzard and Turner's fMRI data set, difference image between average across stimulation scans and average across rest scans. Darker regions correspond to likely areas of activation.*

together accounted for approximately 93% of the variability of this difference data set. Bootstrap simulations with $B = 500$ were performed, and the rank order of the periodic spline fit recorded. As above, 'two-sided' ranks and p-values were computed and are displayed for this data in the following figures. Figure (3.23) on page 70 shows the two-sided rank for each pixel of the Friston, Jezzard and Turner data, and Figure (3.24) on page 71 displays the corresponding p-values. Periodic spline fits to the several of the regions in which activation was identified are seen to have a similar double-humped shape to those observed previously, and the bootstrap ranks and p-values at these locations are seen to be highly significant.

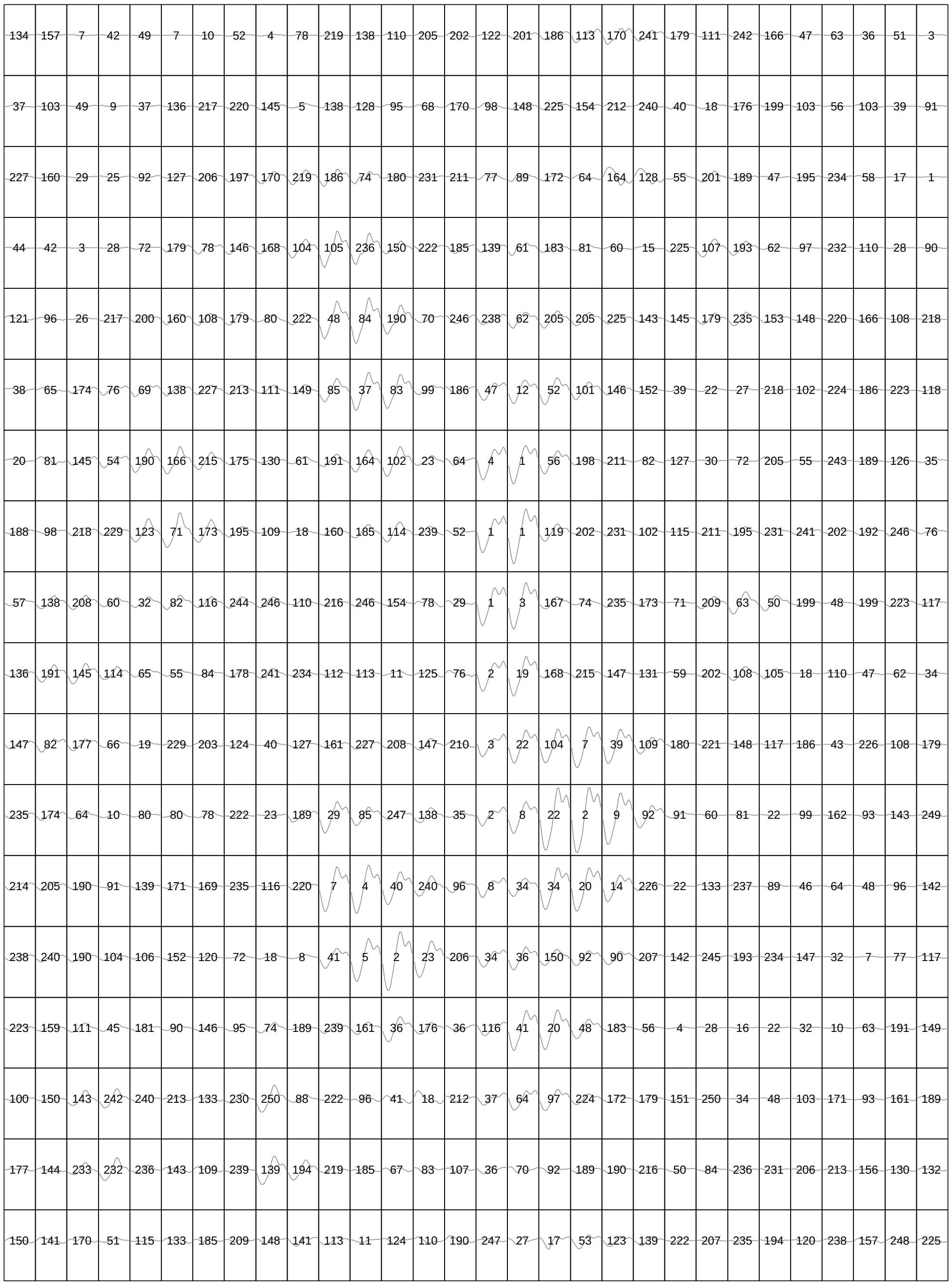

Figure 3.23: *Rank out of 250 obtained in bootstrap simulation comparing periodic spline fits to sinusoid for the Friston, Jezzard and Turner data. Low values would indicate significance.*

Figure 3.24: *p-values from bootstrap simulation comparing periodic spline fits to sinusoid for the Friston, Jezzard and Turner data. Low values would indicate significance.*

3.6.3 Hemodynamic spline models versus sinusoid

For completeness, it seems only reasonable to consider whether the models afforded by the hemodynamic spline modeling approach perform better than the sinusoidal model at any pixel. An identical procedure to the two above was followed with the example model fits performed above (where $L = 15$ and $K = 8$). The score used to separate the spline-based convolution model fits from the best fitting sinusoids was again the second principal component. As for the Poisson-based convolution model case, the first principal component for the spline-based convolution models was seen to discriminate the cases of better fit by a sinusoid for low amplitude cases where the convolution constraint is restricted to an almost flat estimate.

As before, B bootstrap resamples were computed at each pixel, one being the spline-based convolution model fit to the true pixel data, and $B - 1$ by adding a resampled cycle of residuals to each cycle of the best fitting sinusoid at that pixel. The rank of the second principal component score among the B resamples was recorded, and the ranks and corresponding p-values are displayed in Figures (3.25) and (3.26) below.

Note that several of the pixels have ranks and p-values suggesting they are significantly different from simply fitting a sinusoid to the data at the fundamental frequency. Although the convolution model has the inherent restrictions described above, flexibly modeling the hemodynamic response function using cubic splines appears to afford a non-trivial improvement in capturing the underlying pattern of the data. In addition, it has the advantage of being directly comparable to the convolution models proposed and used by Friston and others, which the periodic spline model lacks. The question remains as to how much to believe in these model fits. The nature of the model is unable to capture what appears to be the true structure as estimated by the periodic spline fits.We are however fitting K parameters per time

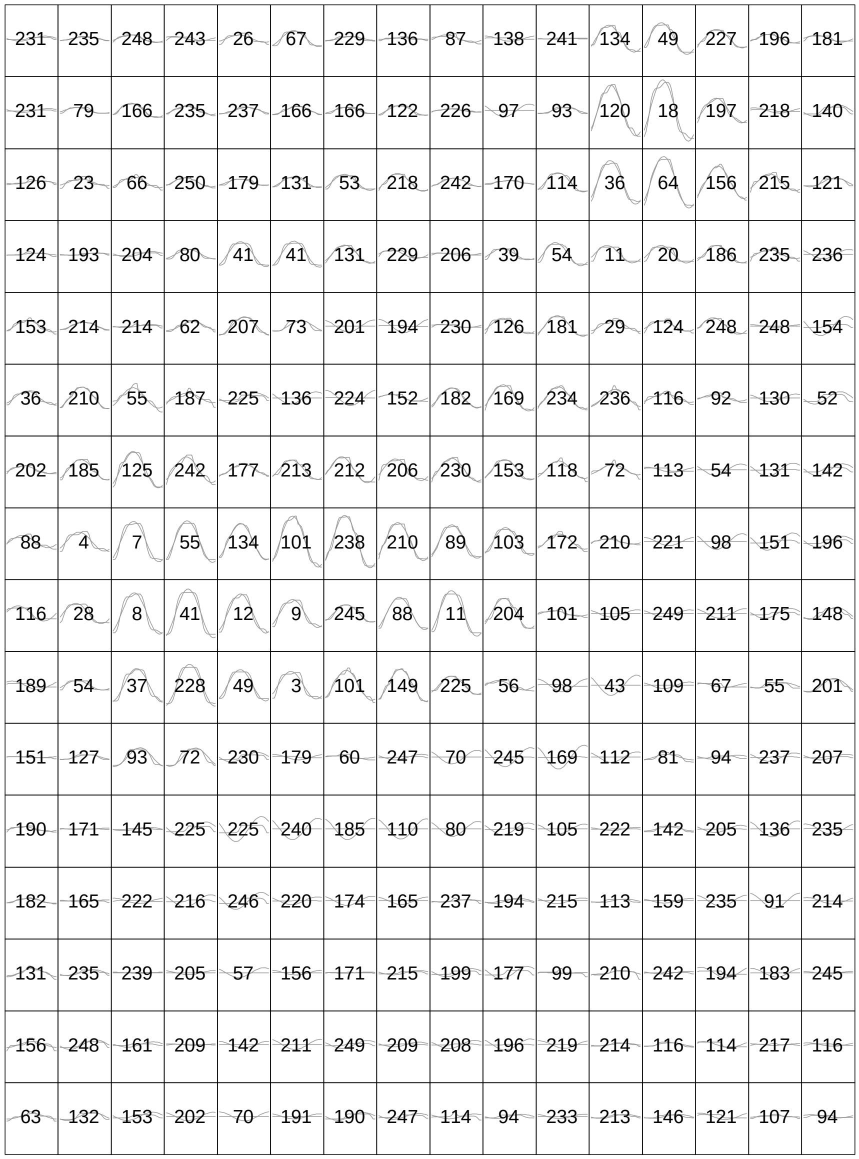

Figure 3.25: *Rank out of 250 obtained in bootstrap simulation comparing spline-based convolution model fits to sinusoid. Low values would indicate significance.*

0.924	0.94	0.992	0.972	0.104	0.268	0.916	0.544	0.348	0.552	0.964	0.536	0.196	0.908	0.784	0.724
0.924	0.316	0.664	0.94	0.948	0.664	0.664	0.488	0.904	0.388	0.372	0.48	0.072	0.788	0.872	0.56
0.504	0.092	0.264	1	0.716	0.524	0.212	0.872	0.968	0.68	0.456	0.144	0.256	0.624	0.86	0.484
0.496	0.772	0.816	0.32	0.164	0.164	0.524	0.916	0.824	0.156	0.216	0.044	0.08	0.744	0.94	0.944
0.612	0.856	0.856	0.248	0.828	0.292	0.804	0.776	0.92	0.504	0.724	0.116	0.496	0.992	0.992	0.616
0.144	0.84	0.22	0.748	0.9	0.544	0.896	0.608	0.728	0.676	0.936	0.944	0.464	0.368	0.52	0.208
0.808	0.74	0.5	0.968	0.708	0.852	0.848	0.824	0.92	0.612	0.472	0.288	0.452	0.216	0.524	0.568
0.352	0.016	0.028	0.22	0.536	0.404	0.952	0.84	0.356	0.412	0.688	0.84	0.884	0.392	0.604	0.784
0.464	0.112	0.032	0.164	0.048	0.036	0.98	0.352	0.044	0.816	0.404	0.42	0.996	0.844	0.7	0.592
0.756	0.216	0.148	0.912	0.196	0.012	0.404	0.596	0.9	0.224	0.392	0.172	0.436	0.268	0.22	0.804
0.604	0.508	0.372	0.288	0.92	0.716	0.24	0.988	0.28	0.98	0.676	0.448	0.324	0.376	0.948	0.828
0.76	0.684	0.58	0.9	0.9	0.96	0.74	0.44	0.32	0.876	0.42	0.888	0.568	0.82	0.544	0.94
0.728	0.66	0.888	0.864	0.984	0.88	0.696	0.66	0.948	0.776	0.86	0.452	0.636	0.94	0.364	0.856
0.524	0.94	0.956	0.82	0.228	0.624	0.684	0.86	0.796	0.708	0.396	0.84	0.968	0.776	0.732	0.98
0.624	0.992	0.644	0.836	0.568	0.844	0.996	0.836	0.832	0.784	0.876	0.856	0.464	0.456	0.868	0.464
0.252	0.528	0.612	0.808	0.28	0.764	0.76	0.988	0.456	0.376	0.932	0.852	0.584	0.484	0.428	0.376

Figure 3.26: *p-values from bootstrap simulation comparing spline-based convolution model fits to sinusoid. Low values would indicate significance.*

course, which may be over-parameterization. Nonetheless, the estimates so obtained appear to include the biphasic, 'early', and 'late' profiles described by Friston et al [28]. In the next section, we will also see that using these estimates in a clustering procedure identifies interesting spatial arrangements of brain response.

3.7 Clustering

Figure (3.27) shows the estimated hemodynamic response function for each pixel in the 16 by 16 subregion identified in Figure (3.2). Observe that regions of activity tend to lie along the upper bank of the sulcus, and that patterns of 'early', 'late' and biphasic response occur, as observed in Friston, Frith, Turner and Frackowiack [28]. Note also these regions appear to be spatially arranged, with neighboring pixels likely to be similar, and biphasic regions between 'early' and 'late' regions. If we use the hemodynamic response function estimates as vector inputs into a multivariate clustering procedure, the resulting segmentation of the brain may be of physiological interest to neuroscientists.

We used Hartigan's [39, 38] k-means clustering procedure to produce the classification displayed in Figures (3.28) and (3.29). Figure (3.28) shows the class centroids and Figure (3.29) shows the classification generated. The inputs to the clustering algorithm are the estimates of the hemodynamic response. The number of clusters chosen was selected in accordance with Hartigan's suggestion of a heuristic test for the value of k. Specifically, if SS_j is the within-group sum of squares for the k-means clustering into j groups and n is the number of input vectors, we use the k-means clustering into $j+1$ groups if $(n-j-1)(SS_j/SS_{j+1}-1)$ is greater than 10. Note the classification into regions of no activation, monophasic activation with 'early' and 'late' profiles, and biphasic profiles. Note also the spatial arrangement of regions with

Figure 3.27: *The estimated hemodynamic response functions in the 16 by 16 subregion of visual cortex identified in Figure (3.2), fit using a cubic B-spline basis expansion.*

like profiles clustered together and biphasic regions spatially between 'early' and 'late' regions.

While the hemodynamic spline model offers direct comparability with the conjectures of Friston et al [28], the constraints imposed by the convolution model and the similarities of the reconstructions based on it call into question whether the shapes of the spline-based hemodynamic response estimates can be reliably and robustly estimated. In short, the shapes used as inputs to this clustering procedure may or may not be real. Given that we are much more confident that the pattern in the periodic spline model fits are real, what sort of a classification is obtained using them as inputs?

Figure (3.30) shows a single cycle of the periodic spline model fits to each pixel of the same 16 by 16 subregion. Using the same clustering procedure as described above on these inputs, we obtain the class centroids displayed in Figures (3.31) and (3.32),

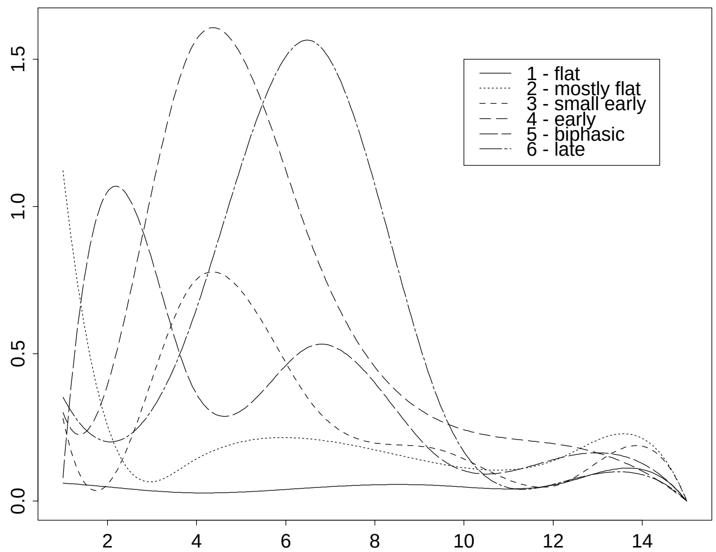

Figure 3.28: *The class centroids of a k-means procedure for $k = 6$ with inputs being the cubic B-spline estimate of the hemodynamic response deconvolution.*

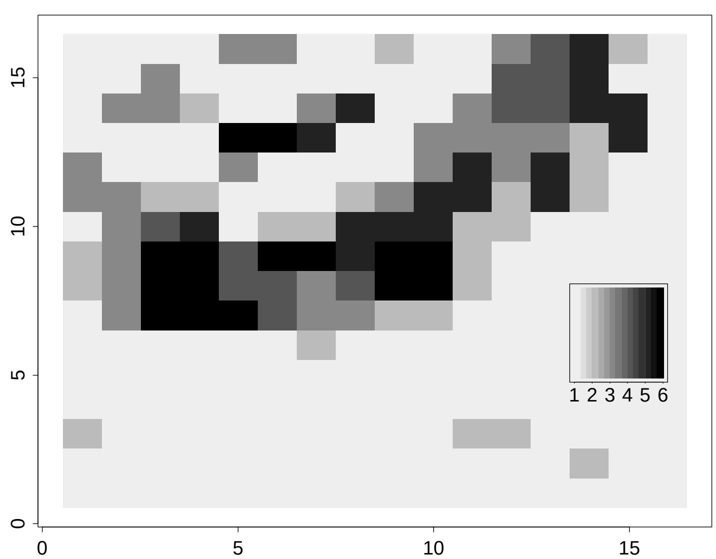

Figure 3.29: *The classification produced by the k-means procedure.*

and the classification generated is shown in Figure (3.33). Again we see areas of similar response spatially clustered together. The last three classes typify the pattern previously described in the data and are clustered as spatial neighbors of each other. The first three classes appear to be mostly noise patterns, and the fourth has a pattern quite out of phase with the activation profile.

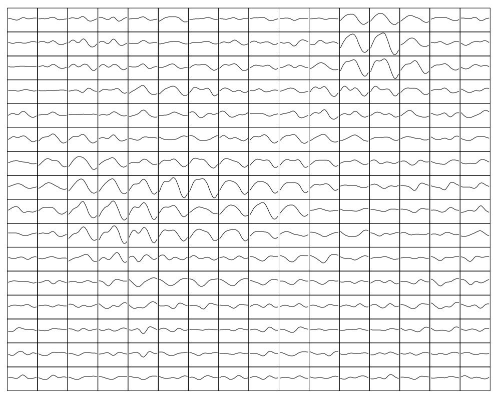

Figure 3.30: *One cycle of the periodic spline model fit at each pixel in the 16 by 16 subregion of visual cortex identified in Figure (3.2), fit using a cubic B-spline basis expansion.*

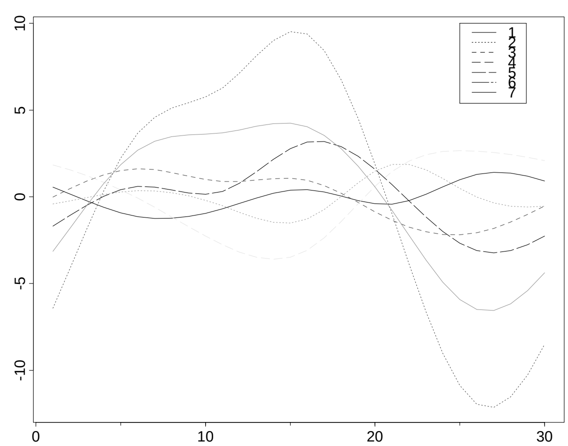

Figure 3.31: *The class centroids of a k-means procedure for $k = 7$ with inputs being one cycle of the periodic spline fit to the pixel time course.*

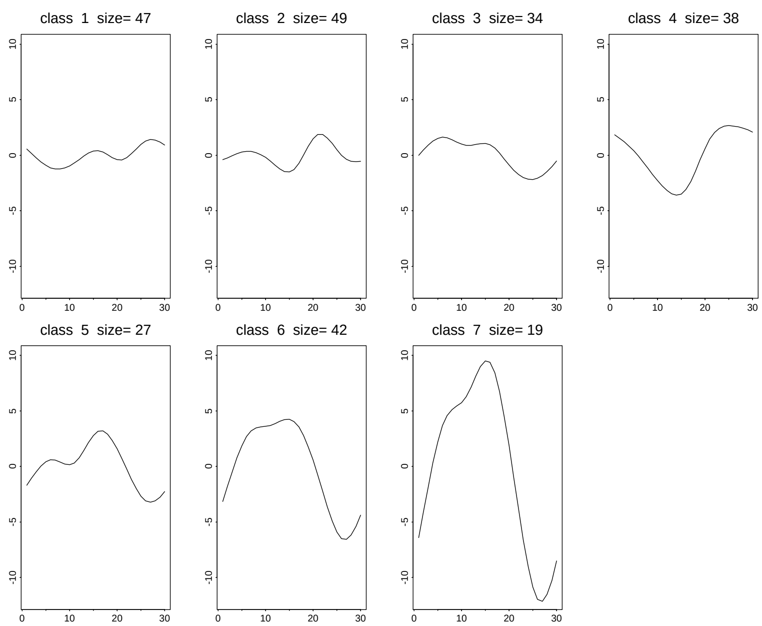

Figure 3.32: *The class centroids of a k-means procedure for $k = 7$ with inputs being one cycle of the periodic spline fit to the pixel time course.*

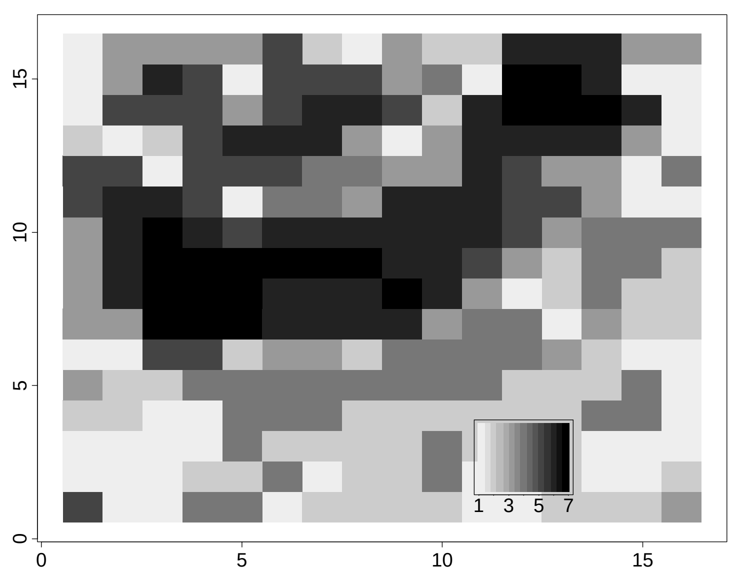

Figure 3.33: *The classification produced by the k-means procedure on the periodic spline fits inputs.*

Lee, Glover and Meyer [46] and others use Fourier domain phase information to discriminate likely venous vessel responses from cortical activity. Classification using model-based estimates may provide further useful insights into the nature and spatial arrangement of brain activity.

3.8 Inference

In order to determine regions of the brain that are exhibiting significant activation, Friston, Jezzard and Turner [31] compute a normalized cross-covariance and use the theory of stationary random Gaussian fields for multiple-comparison correction. Lange and Zeger propose a model in which each voxel has a response magnitude coefficient which is multiplied by the Poisson model. For inference about activation, they consider a vector of the response magnitude parameter estimates and use approximate multivariate Gaussian theory with a patterned variance-covariance matrix based on spatio-temporal auto-covariance. The test for significant activation is essentially a test that the response magnitude parameter is significantly different from zero, and Worsley [64] has developed the distribution theory of such a t-field on the image lattice for such inference. By comparison with a single magnitude parameter multiplied by a probability density, we now have several parameters to estimate and a non-negative linear combination of densities being estimated. A reasonably straightforward approximate inference procedure would be to assess significant activation based on a hypothesis test of the null hypothesis that all the parameter estimates are zero. Based on the same mathematical theory of stationary random Gaussian fields, an approximate F-field can be computed to determine where significant brain activation has occurred. This is not pursued further here. A more detailed summary of inference procedures is contained in Chapter 7.

3.9 Summary

This chapter shows several important results about fMRI time course data. The first is that the convolution model contains intrinsic structural constraints on the nature of the reconstructions or model fits which cannot capture the patterns seen in the periodic spline fits to the data. One possible modification which could improve matters would be to allow the hemodynamic response function to take negative values, at the cost of more difficult physiological interpretation[4]. Models for the hemodynamic response function are almost uniformly positive, usually Poisson, Gamma or Gaussian density functions however, and as such the models will not capture the patterns seen in the periodic spline fits.

Another important result is that fitting the Poisson-based convolution model is shown conclusively to perform no better than fitting a sinusoid with arbitrary phase and amplitude at the activation frequency. However, fitting the spline-based convolution model does outperform the sinusoidal fits, despite the inherent constraints of the convolution model. The periodic spline fits to the time course data are quite substantially different from the sinusoid fits, and the patterns seen in them can be argued to be quite real. This is further confirmed by observing the patterns remain when the number of knots is increased.

We observe additionally that the responses are spatially clustered. The patterns seen in the spline fits to the data at the activation areas are similar between spatial neighbors. The locations seen to be different from the sinusoidal fits are spatially clumped. This leads us to investigate clustering procedures applied to the model fits, which provided interesting and informative brain maps.

[4]David Heeger, in personal communication, proposes the following possible mechanism: immediately following neural activation but before vascular replenishment, there is a higher level of deoxyhemoglobin at activation sites. This could be modeled by negative hemodynamic response values. This usually happens faster than images are rescanned however.

A flexible family of models, cubic B-splines, has provided more detailed insights into the nature of evoked hemodynamic response and its spatial arrangement. Despite the constraint asserted by Friston, Frith, Turner and Frackowiack [28] that the number of temporal basis functions be kept to a minimum, both meaningful inference and interpretability are seen to be usefully preserved.

An interesting question to ask is whether the spatio-temporal correlation of the data can be exploited to improve estimation and inference. For example, the hemodynamic responses in a spatial neighborhood's pixels could be fit in some joint fashion, encouraging regional similarity or smooth spatial variation. Alternatively, by exploiting the multivariate nature of the responses, we may be able to "borrow strength" from neighboring pixels in a James-Stein shrinkage sense to improve estimation. The Hartigan k-means procedure can be easily adapted such that the distance used in determining cluster membership be generalized to include a metric based on spatial proximity of the estimates within the image. These generalizations are touched on briefly in Chapter 5.

Chapter 4

Restricted Singular Value Decompositions

In this chapter, singular value decompositions with basis expansion restrictions on the left and right singular vectors are considered. The expected periodicity and temporal smoothness of the time courses and spatial smoothness of the the brain response can be incorporated into a singular value decomposition of the image sequence data offering new views of the data.

4.1 Principal components and SVD

As described in Chapter 2, Friston et al [27, 31, 22] and others have used principal component analysis and the SVD of PET and fMRI image sequences to produce informative views of the high dimensional data. We consider the matrix Y, as before, to be the $n \times M$ matrix of images, here centered with both row and column averages subtracted throughout. There exists a unique singular value decomposition of the matrix Y, given by $Y = UDV^T$, such that U and V are orthogonal and D is diagonal with non-increasing diagonal entries. These diagonal entries are the singular values, with the corresponding rows of U and V being the singular vectors.

The leading principal component of Y (with the average image removed) is the linear combination of pixel values which explains the maximal amount of variability among the rows of Y, the images in the image sequence. It is thus a contrast image which shows the most spatial variation, and is informative when displayed as an image. The corresponding time plot shows each image in the series projected onto this component or axis (an inner product) and shows how these maximal variations occurred in time. This leading principal component is also the first right singular vector of the matrix Y. It may be sensible to confine attention by masking the image to exclude areas of "non brain" material and white matter, to have only grey matter included. (Lange [43] describes a simple method for doing this, using histogram thresholding.)

Similarly one can treat the pixelwise time series as vectors, and produce principal components for them. Assuming again that the pixelwise time series are centered, the first principal component is the linear combination of pixel values which explains the maximal amount of variability among the columns of Y (equivalently, the rows of Y^T), which are the time courses at each pixel. This leading principal component is consequently also the first left singular vector of Y. The corresponding spatial plot showing how these maximal variations occur within the anatomical images, projects each pixels time course onto this component and can be color coded over anatomical images of the brain. Using the singular value decomposition, one can obtain both of these simultaneously at no extra cost. In identical fashion, subsequent orthogonal principal components correspond to the appropriate singular vectors.

4.2 Restricted SVD: motivation and theory

Here we would like to consider more structured approaches. For example, we might constrain the right singular vectors to be spatially smooth, or limited in other ways, and the left singular vectors likewise to be smooth and periodic functions of time.

In general, the unrestricted SVD of a centered matrix Y solves the optimization problem

$$\inf_{U,V} \left\| Y - UDV^T \right\|^2$$

subject to orthonormality constraints $U^TU = I$, $V^TV = I$ and a diagonal form for D with non-increasing diagonal entries.

If we reparameterize $U = H_U U^*$ and $V = H_V V^*$ where H_U is a restricted orthonormal basis for the time components and H_V is a restricted orthonormal basis for the spatial components then we can solve for the reduced parameter matrices U^* and V^* instead. The restricted orthonormal basis matrix H_U is formed by taking an orthogonal set of $n_U (\leq n)$ basis vectors for a linear subspace of $\mathbb{R}^n$ as columns. Similarly, H_V is formed taking $M_V (\leq M)$ orthonormal basis vectors for a linear subspace of $\mathbb{R}^M$ as columns. In this way, the singular vectors are restricted to lie in the linear subspaces so chosen. Note the cases of $n_U = n$ and $M_V = M$ impose no restriction on their corresponding singular vectors.

It is not hard to show that these restricted singular vectors are obtained from the reduced or filtered SVD problem with matrix $Y^* = H_U^T Y H_V$, and solve the optimization problem:

$$\inf_{\tilde{U},\tilde{V}} \left\| Y^* - \tilde{U}\tilde{D}\tilde{V}^T \right\|^2$$

where $Y^* = H_U^T Y H_V$, again subject to $\tilde{U}$ and $\tilde{V}$ being orthonormal and $\tilde{D}$ diagonal with non-increasing diagonal entries.

The proof of the above is as follows: If $U = H_U U^*$, $V = H_V V^*$, let $Q_U =$

$\begin{bmatrix} H_U & H_{U\perp} \end{bmatrix}$ and $Q_V = \begin{bmatrix} H_V & H_{V\perp} \end{bmatrix}$ where $H_{U_\perp}$ is an orthonormal basis completion of H_U over $I\!R^n$ and $H_{V_\perp}$ is an orthonormal basis completion of H_V over $I\!R^M$, then

$$\begin{aligned}
\left\|Y - UDV^T\right\|^2 &= \left\|Q_U^T\right\|^2 \left\|Y - UDV^T\right\|^2 \|Q_V\|^2 \quad (Q_U \text{ and } Q_V \text{ are orthonormal}) \\
&= \left\|Q_U^T(Y - UDV^T)Q_V\right\|^2 \\
&= \left\|Q_U^T Y Q_V - Q_U^T UDV^T Q_V\right\|^2 \\
&= \left\| \begin{bmatrix} H_U^T \\ {H_{U_\perp}}^T \end{bmatrix} Y \begin{bmatrix} H_V & H_{V_\perp} \end{bmatrix} - \begin{bmatrix} H_U^T \\ {H_{U_\perp}}^T \end{bmatrix} H_U U^* D V^{*T} H_V^T \begin{bmatrix} H_V & H_{V_\perp} \end{bmatrix} \right\|^2 \\
&= \left\| \begin{bmatrix} H_U^T Y H_V & H_U^T Y H_{V_\perp} \\ {H_{U_\perp}}^T Y H_V & {H_{U_\perp}}^T Y H_{V_\perp} \end{bmatrix} \right. \\
&\qquad \left. - \begin{bmatrix} (H_U^T H_U) U^* D V^{*T} (H_V^T H_V) & (H_U^T H_U) U^* D V^{*T} (H_V^T H_{V_\perp}) \\ ({H_{U_\perp}}^T H_U) U^* D V^{*T} (H_V^T H_V) & ({H_{U_\perp}}^T H_U) U^* D V^{*T} (H_V^T H_{V_\perp}) \end{bmatrix} \right\|^2 \\
&= \left\| \begin{bmatrix} H_U^T Y H_V & H_U^T Y H_{V_\perp} \\ {H_{U_\perp}}^T Y H_V & {H_{U_\perp}}^T Y H_{V_\perp} \end{bmatrix} - \begin{bmatrix} U^* D V^{*T} & 0 \\ 0 & 0 \end{bmatrix} \right\|^2 \\
&= \left\| \begin{bmatrix} H_U^T Y H_V - U^* D V^{*T} & H_U^T Y H_{V_\perp} \\ {H_{U_\perp}}^T Y H_V & {H_{U_\perp}}^T Y H_{V_\perp} \end{bmatrix} \right\|^2 \\
&= \left\|H_U^T Y H_V - U^* D V^{*T}\right\|^2 + \text{ terms not depending on } U \text{ or } V
\end{aligned}$$

It follows that minimizing $\left\|Y - UDV^T\right\|^2$ for orthonormal U and V of the form $U = H_U U^*$ and $V = H_V V^*$ is equivalent to minimizing $\left\|Y^* - \tilde{U}\tilde{D}\tilde{V}^T\right\|^2$ where $Y^* = H_U^T Y H_V$, as stated above. Observe that the dimensionality of the singular value problem is reduced to finding the singular vectors of the $n_U \times M_V$ matrix Y^*, which can produce a substantial computational advantage.

4.3 Application and Examples

As an example of how this works for a particular fMRI data set, we take the experimental data described in Chapter 2 presented in Figure (2.1). For purposes of comparison, we first compute the first four (unrestricted) singular vectors for that data, obtaining the curves and images displayed in Figures (4.1) and (4.2). The curves displayed are the scaled singular vectors, the singular vectors multiplied by the corresponding singular values.

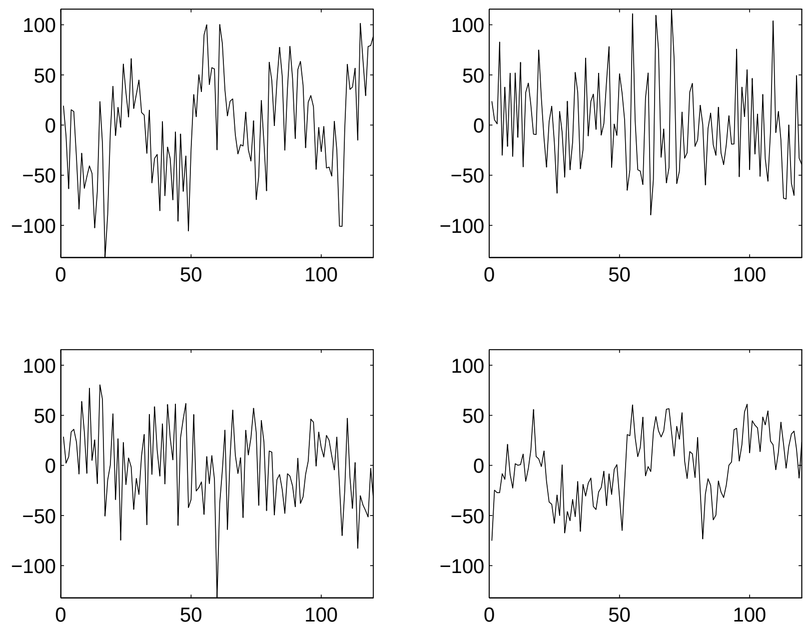

Figure 4.1: *The first four U loadings of the unrestricted SVD, scaled by the corresponding singular values.*

Note there appears to be a certain periodicity to several of the singular vector loadings corresponding to the activation-rest cycles. There is some indication of an overall linear drift. Also note that the spatial contrasts appear aligned with the location of the calcarine sulcus. In addition, observe that the left singular vectors, the U loadings, like the original data, appear to have a great deal of high frequency noise. The principal components are trying to explain as much variability as they

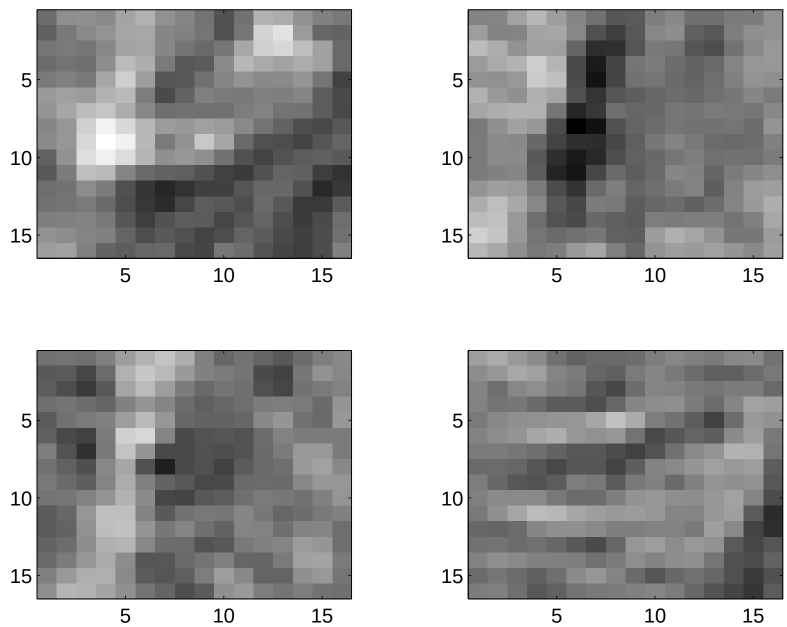

Figure 4.2: *The first four V loadings of the unrestricted SVD, scaled by the corresponding singular values.*

can, and high frequency variability is clearly part of this. Recall Friston et al [29] and Worsley and Friston [67] remove this high frequency noise by temporally smoothing the data, premultiplying the data by a matrix $\mathbf{K}$ based on a fixed width Gaussian kernel. Restricting the left singular vectors to a temporally smooth linear subspace of $I\!R^n$ provides an alternative method of discounting this high frequency noise. The nature of the temporal smoothing imposed by the $\mathbf{K}$ matrix is extremely prescriptive, determined by the bandwidth of the Gaussian kernel. This approach allows the singular vectors to be as rough as the chosen subspace restriction allows.

4.3.1 Restriction on the left

For example, suppose we were to restrict the left singular vectors to be both periodic and smooth, as members of a basis expansion of periodic cubic splines with 5 equally spaced knots in each treatment-rest cycle. If we do this, we obtain a restricted singular value decomposition whose first four singular vectors are displayed in Figures (4.3) and (4.4) below.

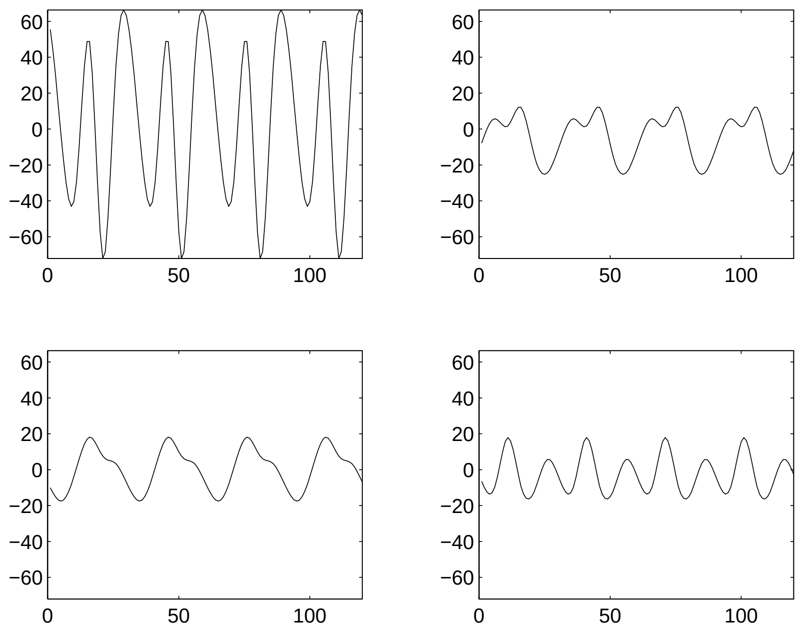

Figure 4.3: *The first four U loadings of the SVD restricted to have left singular vectors from a basis of periodic cubic splines with 5 equally spaced knots across a cycle, scaled by the corresponding singular values.*

As another example, consider a basis matrix consisting of sines and cosines at multiples of the activation frequency. The scaled singular vectors displayed in Figures (4.5) and (4.6) below were produced by taking an orthonormalized basis of sines and cosines at frequencies of once, twice and three times the activation frequency. These are by nature periodic and smooth. In effect such a basis imposes an extreme low-pass filter on the signal.

Note the patterns of the leading left singular vectors in the above two cases display similar temporal profiles to those discerned in Chapter 3 with the periodic spline model fits. These are the principal components describing the pattern of maximal variation. Interestingly, restricting to a trigonometric basis consisting of just the first

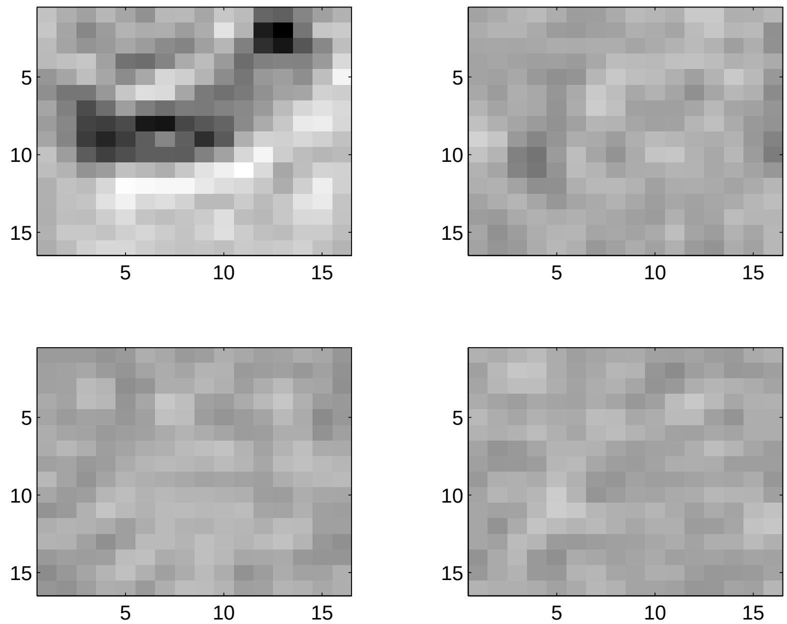

Figure 4.4: *The first four V loadings of the SVD restricted to have left singular vectors from a basis of periodic cubic splines with 5 equally spaced knots across a cycle, scaled by the corresponding singular values.*

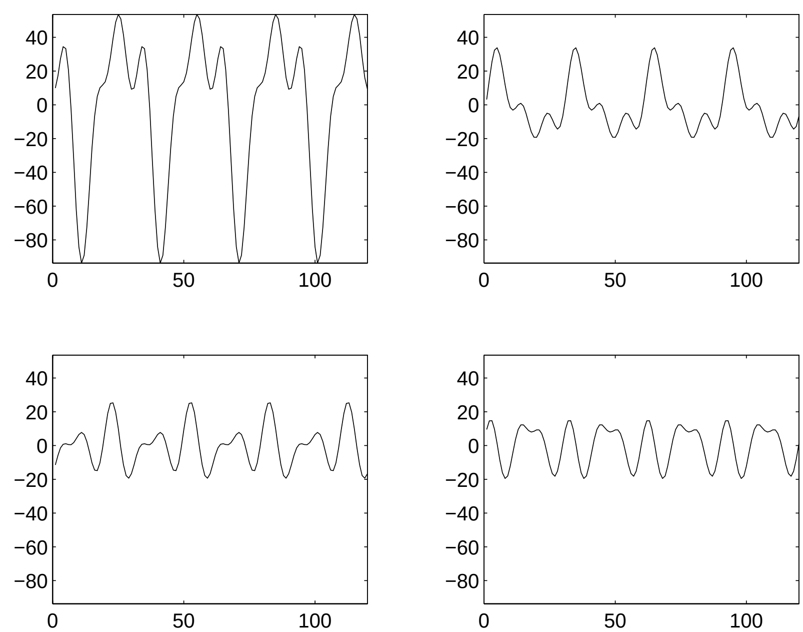

Figure 4.5: *The first four U loadings of the SVD restricted to have left singular vectors from a low frequency trigonometric basis, scaled by the corresponding singular values.*

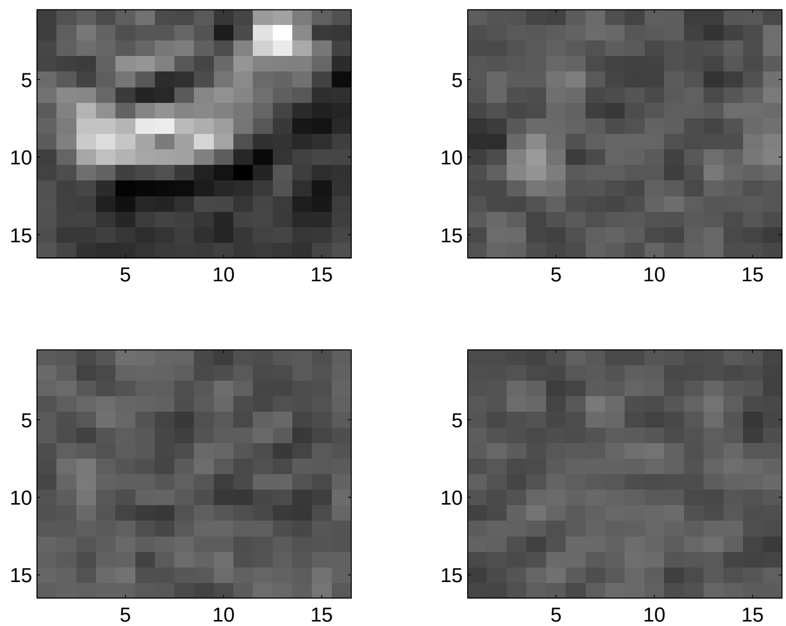

Figure 4.6: *The first four V loadings of the SVD restricted to have left singular vectors from a low frequency trigonometric basis, scaled by the corresponding singular values.*

and third harmonics (not shown), we do not observe this characteristic pattern, as that subspace incorporates the mirror antisymmetry property showcased in Figure (3.11).

4.3.2 Restriction on the right

To restrict the right singular vectors in a meaningfully interpretable way in this fashion, we require a sensible set of spatial basis functions. Note that spatial masking as used by Lange [43] is itself a form of spatial restriction. Masking can also be incorporated into the choice of spatial basis functions, but we do not explore this explicitly here. An obvious candidate for a spatial basis consists of tensor products of univariate basis functions in either spatial direction. This is the method underlying much of two-dimensional wavelet analysis, for example. An easy to understand example of such a basis is the Haar basis, displayed in Figures (4.7) and (4.8) below.

While the Haar basis is easy to explain and understand, it is somewhat deficient

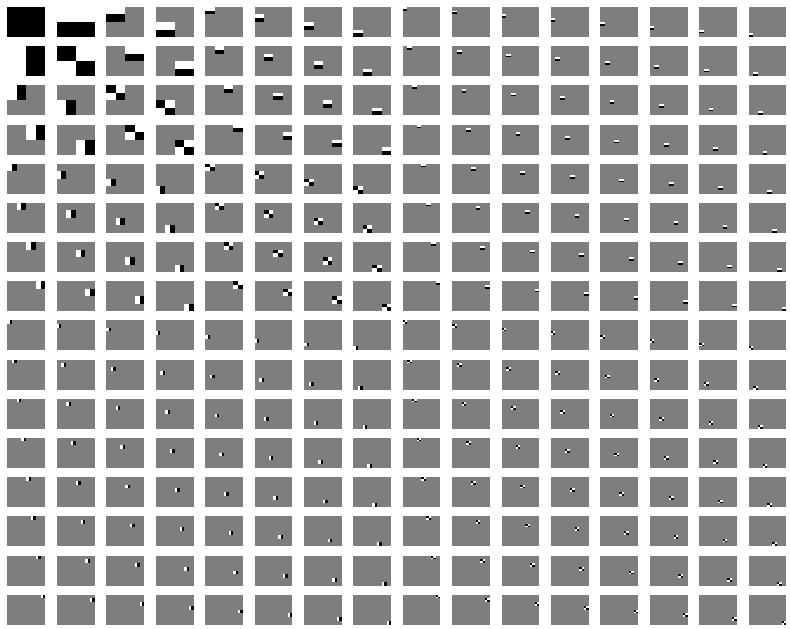

Figure 4.7: *A spatial basis formed of tensor products of Haar basis wavelets.*

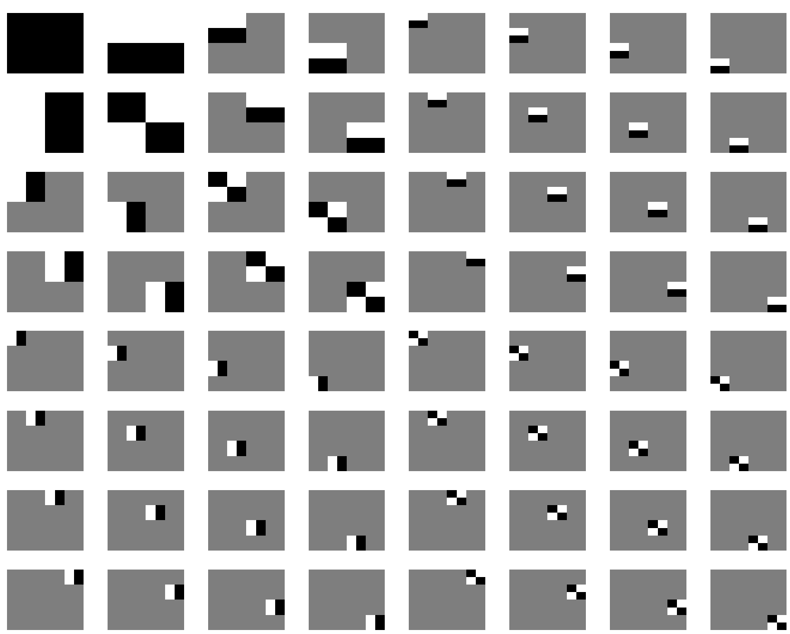

Figure 4.8: *The upper 8-by-8 corner of Figure (4.7)*

for our purposes here. The first coefficient fits a constant to the entire image, the second to the difference between the left and right halves of the image and so on down to coefficients for the differences between neighboring pixels or squares of pixels. This basis is thus a basis of piecewise constant functions defined over square subregions down to the pixel level. Taking the full set of basis functions in Figure (4.7) is then no restriction on the image space, and so we need to take some reasonable subset of that basis. One possible restriction is represented by the basis subset of Figure (4.8), which restricts the image space to piecewise constant functions over subtiles of the image space of size two pixels by two pixels. Our assumption is that image data from fMRI experiments is likely to be spatially smooth in some fashion. The discontinuous derivatives at subtile boundaries of the Haar basis make it somewhat unappealing in this context. Fortunately, the dictionaries of orthonormal two-dimensional tensor product wavelet bases we have access to is vast, and many have such attractive properties as smooth derivatives up to arbitrary orders. For purposes of example, the basis shown below in Figures (4.9) and (4.10) are of Symmlets with parameter 8. These are wavelets that are smooth and symmetric. The restricted spatial basis used will be the subset represented in Figure (4.10).

When the right singular vectors are restricted to elements from the upper-left 8×8 corner, the coarser elements, of this Symmlet(8) basis, with no restriction on the left singular vectors, the first four restricted singular vectors appear as in Figures (4.11) and (4.12).

Note the left singular vectors appear very similar to the unrestricted case, while the right singular vector images are spatially smoother, as is to be expected. The constrained right singular vectors show similar patterns of contrast to the unrestricted case, with variation apparently corresponding to contrasts between spatially contiguous regions around the calcarine sulcus.

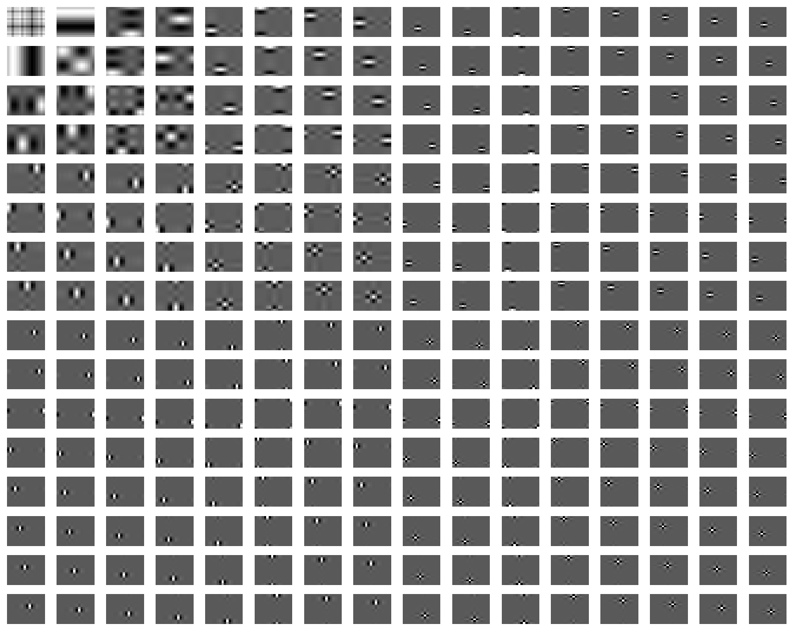

Figure 4.9: *A spatial basis formed of tensor products of Symmlet(8) basis wavelets.*

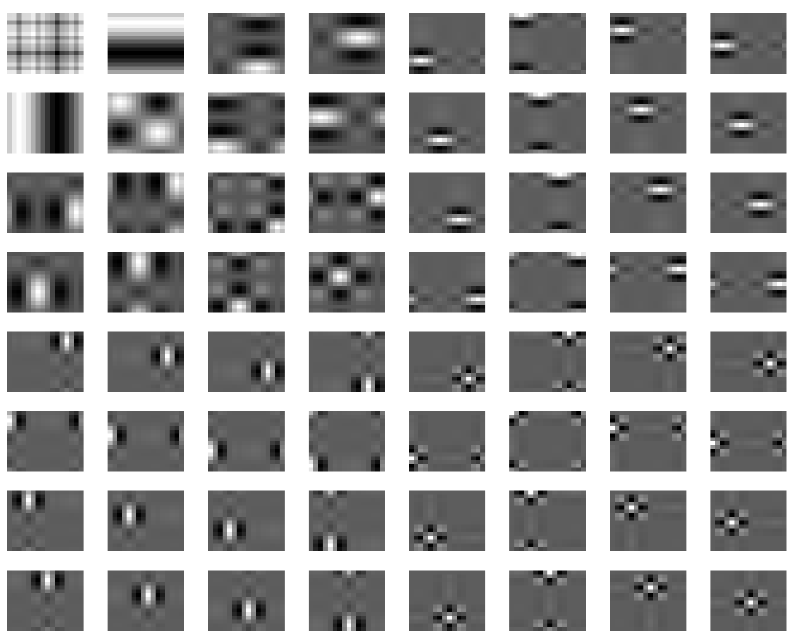

Figure 4.10: *The upper 8-by-8 corner of Figure (4.9)*

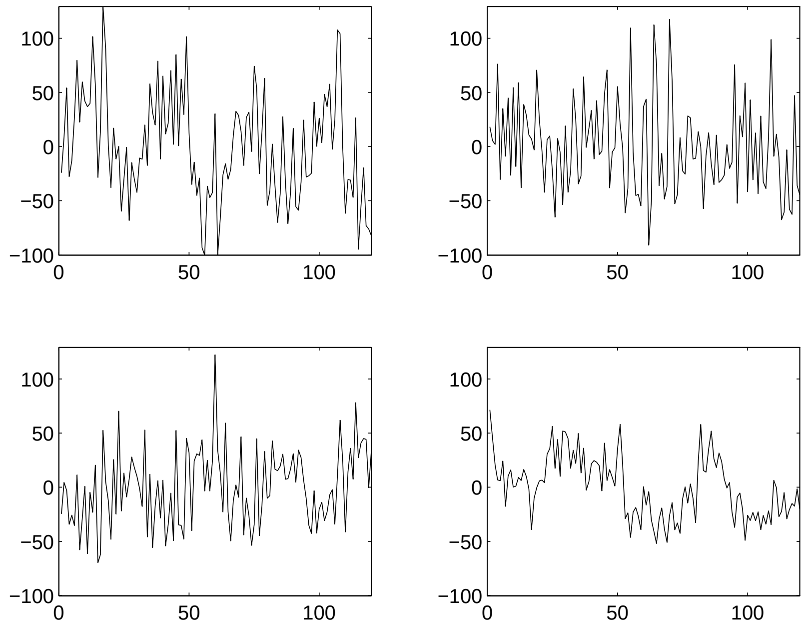

Figure 4.11: *The first four U loadings of the SVD restricted to have right singular vectors from the Symmlet(8) basis.*

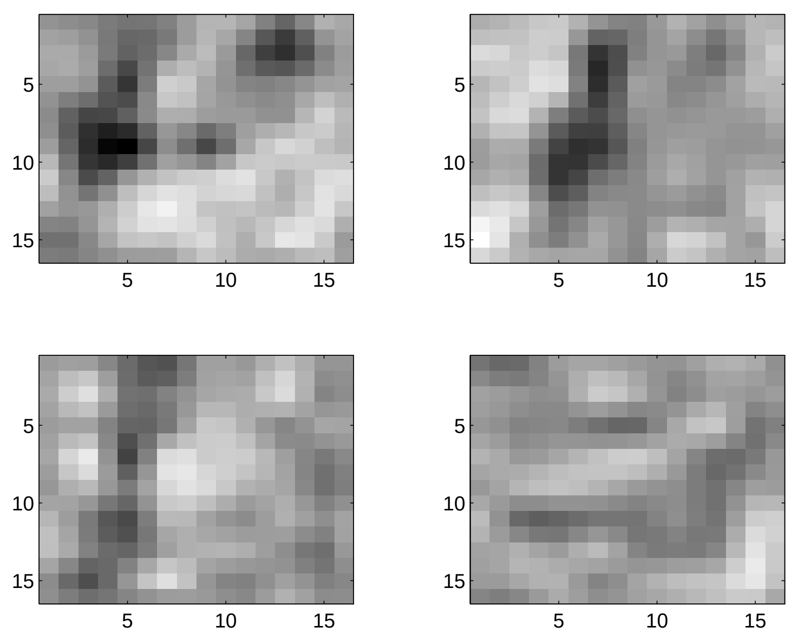

Figure 4.12: *The first four V loadings of the SVD restricted to have right singular vectors from the Symmlet(8) basis.*

4.3.3 Restriction on both left and right

If we enforce both sets of restrictions simultaneously, we obtain the first four singular vectors represented in Figures (4.13) and (4.14). In this example we combine both the periodic spline basis restriction and the Symmlet(8) spatial basis restriction used above.

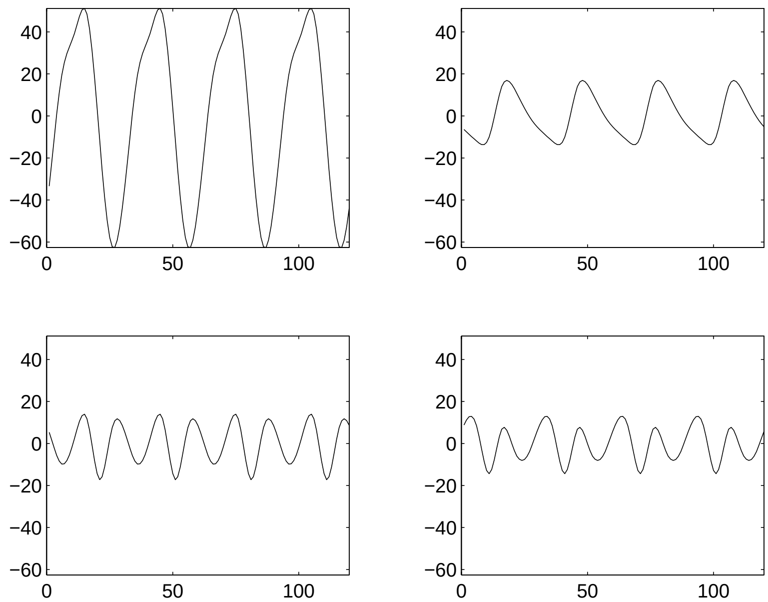

Figure 4.13: *The first four U loadings of the SVD restricted to have left singular vectors from a basis of periodic cubic splines with 5 equally spaced knots across a cycle and right singular vectors from the Symmlet(8) basis.*

As another example, Figures (4.15) and (4.16) show the singular vectors produced when the left singular vectors are restricted to the orthonormalized trigonometric basis of the first three harmonics of the activation frequency described above, and the right singular vectors are restricted to the upper 4 by 4 subset of the Symmlet(8) basis. This latter restriction impose a much greater amount of spatial smoothness than the previous example which used the upper 8 by 8 subset of the Symmlet(8) basis. For the left singular vector subspace restrictions, there are four degrees of freedom in this trigonometric basis example versus five in the case of the periodic spline restriction.

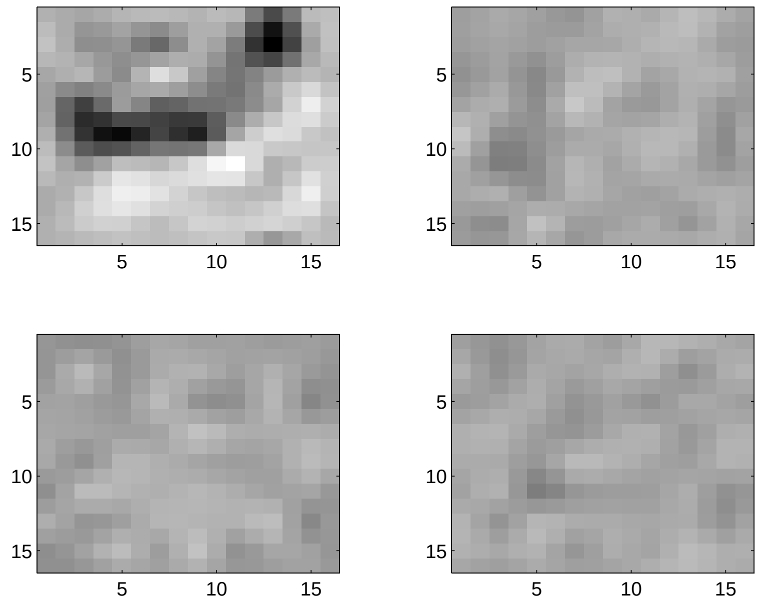

Figure 4.14: *The first four V loadings of the SVD restricted to have left singular vectors from a basis of periodic cubic splines with 5 equally spaced knots across a cycle and right singular vectors from the Symmlet(8) basis.*

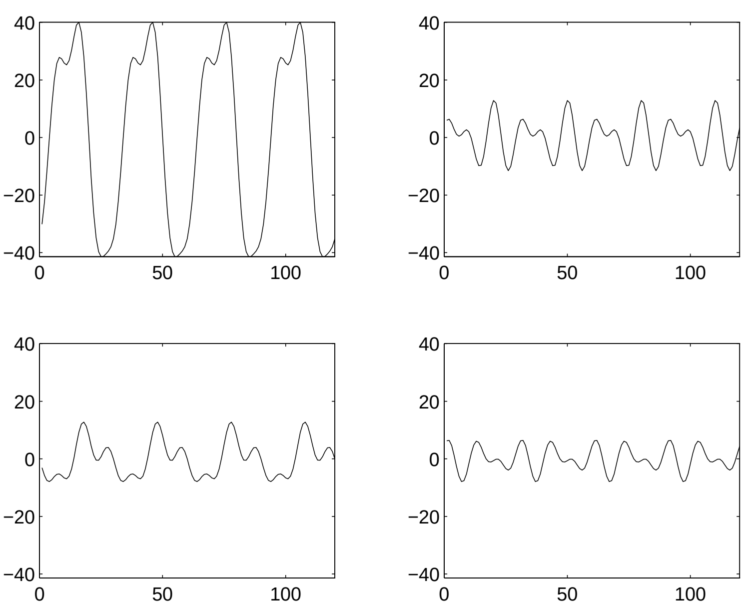

Figure 4.15: *The first four U loadings of the SVD restricted to have left singular vectors from a low frequency trigonometric basis and right singular vectors from the upper 4-by-4 corner of the Symmlet(8) basis.*

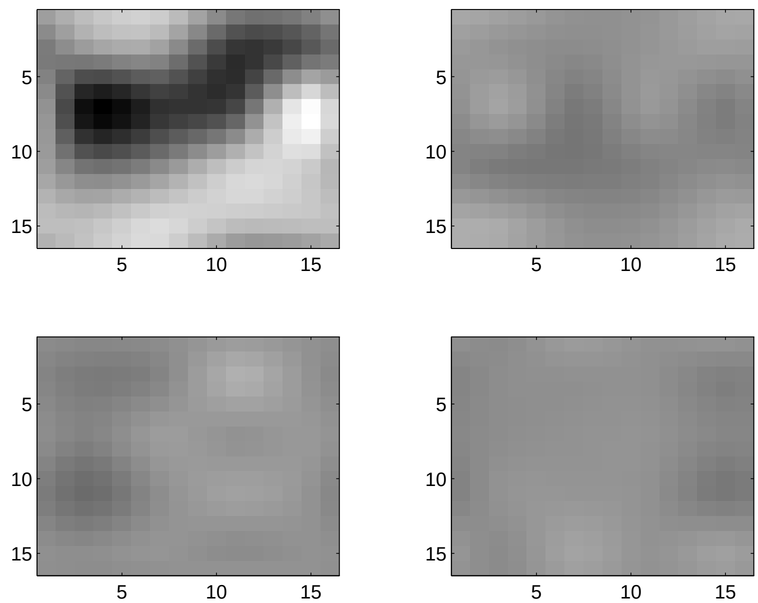

Figure 4.16: *The first four U loadings of the SVD restricted to have left singular vectors from a from a low frequency trigonometric basis and right singular vectors from the upper 4-by-4 corner of the Symmlet(8) basis.*

Observe in both cases that the first left singular vector shows the temporal pattern of maximal variability within this restricted framework has the characteristic pattern of the data observed in Chapter 3. The fMRI magnitude response to activation appears to rise to show a small double hump, and then falls to a level lower than it starts, showing what is sometimes referred to as undershoot. The effect of both restrictions is to smooth the data, removing high frequency variability from consideration, both temporally and spatially. The leading right singular vector in both cases is indicating that the maximal variation of interest lies at the expected locations along the calcarine sulcus. The remaining singular vectors appear uninformative, suggesting that the variability found in subsequent singular vectors in the unrestricted case may be largely due to the high frequency temporal or spatial variability respectively.

4.4 Summary

When applying principal components analysis or the singular value decomposition to raw fMRI data, high frequency noise in the data may make the singular vectors difficult to interpret. Use of the restricted singular value decomposition described in this chapter allows one to restrict the class of possible singular vectors to members of a basis expansion subspace with properties as desired. Thus the singular vectors represent orthonormal contrasts from within their respective restricted subspaces which display maximal variability.

Applying the restricted SVD to impose periodic, temporally smooth left singular vectors for Y and spatially smooth right singular vectors, presents contrasts which show a leading left singular vector with the characteristic double-humped temporal pattern and undershoot seen in Chapter 3 and a leading right singular vector showing maximal variation around the expected activation regions of the calcarine sulcus. The remaining singular vectors (and values) appear non-informative suggesting that remaining variability in the data is due to high frequency temporal or spatial variability.

Chapter 5

Regularized Image Sequence Regression

In this chapter we briefly consider regression performed on image sequences. Basis expansion estimation of regression parameters is considered, including a specific penalized variant which gives rise to a linear transformation and shrinkage procedure. An analogous non-linear shrinkage procedure based on the VisuShrink wavelet shrinkage procedure of Donoho and Johnstone [12] is then considered.

5.1 Image Regression

We consider first some possible variations of the least squares regression problem

$$\hat{\beta}_{LS} = \arg\min_{\beta} \| Y - X\beta \|^2$$

Note this regression can be performed as separate pixel-wise regressions and as such takes no account of any spatial correlation of the responses.

Suppose we wish to impose some structure on the rows of β *a priori*. One way to do this is to take a set of basis functions h_j for $j = 1, \ldots, J$ and to model β from the linear space spanned by those basis functions. Specifically, let H be the $M \times J$ matrix whose columns are values of the basis functions h_j evaluated at the spatial locations of the image pixels. Model β as $\beta = \theta H^T$. Without loss of generality assume the columns of H are orthonormal and denote by Γ the matrix $(H\ H_\perp)$ where $H_\perp$ is any orthocomplement of H $(HH^T + H_\perp {H_\perp}^T = I)$. Note then that

$$\begin{aligned}
\| Y - X\beta \|^2 &= \| (Y - X\beta)\Gamma \|^2 \\
& \quad \text{since the Frobenius and 2- norms are invariant} \\
& \quad \text{under orthonormal transformations} \\
&= \| (YH - X\theta H^T H) \|^2 + \| YH_\perp \|^2 \\
&= \| Y^* - X\theta \|^2 + \| YH_\perp \|^2 \quad \text{where} \quad Y^* = YH
\end{aligned}$$

It follows that

$$\begin{aligned}
\hat{\theta} &= (X^T X)^{-1} X^T Y H \\
&= \hat{\beta}_{LS} H
\end{aligned}$$

and $\hat{\beta}_H = \hat{\beta}_{LS} H H^T$. In other words, $\hat{\beta}_H$ is a smoothed version of the least squares coefficients, where $S = HH^T$ is a spatial smoothing projection operator, projecting into the column space of H.

Suppose further we wish to spatially regularize our regression coefficients using a quadratic smoothness penalty of the form

$$\hat{\beta} = \arg\min_{\beta}(\| Y - X\beta \|^2 + \lambda \| \beta \|_\Omega^2)$$

where $\| \beta \|_\Omega^2 = tr(\beta \Omega \beta^T)$ which amounts to a sum of quadratic penalties for each coefficient image, each parameterized by an identical “smoothing” parameter. This may potentially be generalized to separate smoothing parameters for each coefficient but this has not as yet been explored here.

Suppose $\Omega = HDH^T$ where H is $M \times J$ with $J \leq M$ orthonormal columns and D is a diagonal penalty sequence (with “rougher” columns receiving higher penalties). Now H is an orthonormal set of “basis function” columns with $HH^T + H_{\perp}{H_{\perp}}^T = I$. Suppose as before that $\beta = \theta H^T$ where the β coefficients are expressed as linear combinations of the “basis function” columns of H. Note then that

$$\begin{aligned} \beta\Omega\beta^T &= \theta H^T HDH^T H\theta^T \\ &= \theta D\theta^T \end{aligned}$$

leading to the result that

$$\| Y - X\beta \|^2 + \lambda tr(\beta\Omega\beta^T) = \| Y^* - X\theta \|^2 + \lambda tr(\theta D\theta^T) + \| YH_{\perp} \|^2 \qquad (5.1)$$

and hence we can minimize with respect to θ and transform into results for β. Differentiating (5.1) with respect to θ yields the normal equations

$$X^T X\hat{\theta} + \lambda\hat{\theta}D = X^T Y^*$$

which when considered pixel-wise give the solutions

$$\hat{\theta}_j = (X^T X + \lambda d_j I)^{-1} X^T y_j^*$$

where y_j^* is the j^{th} column of $Y^* = YH$. If in addition we have an orthonormal experimental design ($X^T X = I$), then it follows that

$$\hat{\theta} = X^T Y^* (I + \lambda D)^{-1} \tag{5.2}$$

Transforming to the original covariate basis, this gives us an estimate for β which is

$$\begin{aligned}
\hat{\beta}_\Omega &= \hat{\theta} H^T = X^T Y^* (I + \lambda D)^{-1} H^T \\
&= X^T Y H (I + \lambda D)^{-1} H^T \\
&= \hat{\beta}_{LS} (H (I + \lambda D)^{-1} H^T) \\
&= \hat{\beta}_{LS} S_\Omega
\end{aligned}$$

where S_Ω is a linear "smoother" operating on the rows of Y or $\hat{\beta}_{LS}$ enforcing a spatial smoothness constraint on the coefficient images β.

When we rewrite equation (5.2) as

$$\hat{\theta} = (X^T Y H)(I + \lambda D)^{-1}$$

the procedure can be described as follows: The image sequence is transformed on the right into spatial basis coefficients, on the left by orthogonal regression contrasts, and then operated on by a linear shrinkage operator on the right. Note that since each of these operations are linear mappings, the order in which they are performed is irrelevant.

The linear procedure above presents the possibility of an analogous nonlinear procedure, wherein the image sequence is transformed (linearly) on the right into some spatial basis coordinate system, into (linear) orthogonal contrasts on the right, and a nonlinear shrinkage operator is applied. For example, if the spatial basis used is wavelets, the shrinkage operator used could be SureShrink as proposed in Donoho and Johnstone [11] or VisuShrink as proposed in Donoho and Johnstone [12] or some similar wavelet basis nonlinear shrinkage operator. The transformations on left and right into spatial basis coefficients and orthogonal contrasts respectively remain linear and may be applied in either order prior to shrinkage. Using such wavelet shrinkage techniques will annihilate wavelet basis coefficients that are too small, isolating spatial regions of the coefficient images that best describe the effect of the corresponding covariate. Transforming back to the image domain, the regression coefficient images for each covariate will have been spatially "smoothed" by the procedure, with large scale signal hopefully well described, spatial structure left intact, and high frequency noise removed.

Figure (5.1) below shows an example of such a procedure applied to a the results of an fMRI experiment on the region of primary visual cortex first described in Chapter 2. In Figure (5.1), upper left, are displayed the results the unpenalized (pixel-wise) least squares regression fits of the time courses against a sinusoidal covariate at the stimulation frequency, allowing for optimal phase matching. Figure (5.1), upper right, shows the result of a linear shrinkage heavily penalizing neighboring pixel differences. Figure (5.1), lower left, shows the same regression subjected as described above to nonlinear wavelet shrinkage using VisuShrink soft thresholding, and Figure (5.1), lower right, is the result of the same procedure but using VisuShrink soft thresholding.

Note the spatial smoothness of the latter two figures when compared against the pixel-wise regressions of Figure (5.1), upper left. All four images are presented using the same color map for comparability. The VisuShrink procedure with hard thresholding clearly a more visually appealing display in terms of spatial smoothness, but shrinkage has noticeably reduced the parameter values across the image. Nonetheless, indication of the higher parameter estimates around the expected areas of activation calcarine sulcus in a spatially smoothed estimate map is clearly observed.

Figure (5.2) below shows another example of the above procedure on the same data set. In this example, the norm of projection of the pixel time course into the periodic spline regression basis used in Chapter 3 is used. The matrix X used in this example comprises an orthonormalized basis for the periodic splines with period equal to a complete stimulus-rest cycle. As before, the upper left panel shows the result for the unpenalized regression, the upper right panel the result of linear shrinkage, the lower left shows the result of VisuShrink soft thresholding, and the lower right shows the result of VisuShrink hard thresholding. The shrinkage procedures again produce a more spatially smooth image and reduce the parameter values. VisuShrink hard thresholding again yields the most visually appealing display.

5.2 Summary

Regularized image regression using the shrinkage techniques above can be accomplished quite efficiently computationally. However, the spatial smoothness obtained by shrinking the parameter estimates in this is difficult to interpret and does seem to afford tremendous gains in analysis or understanding.

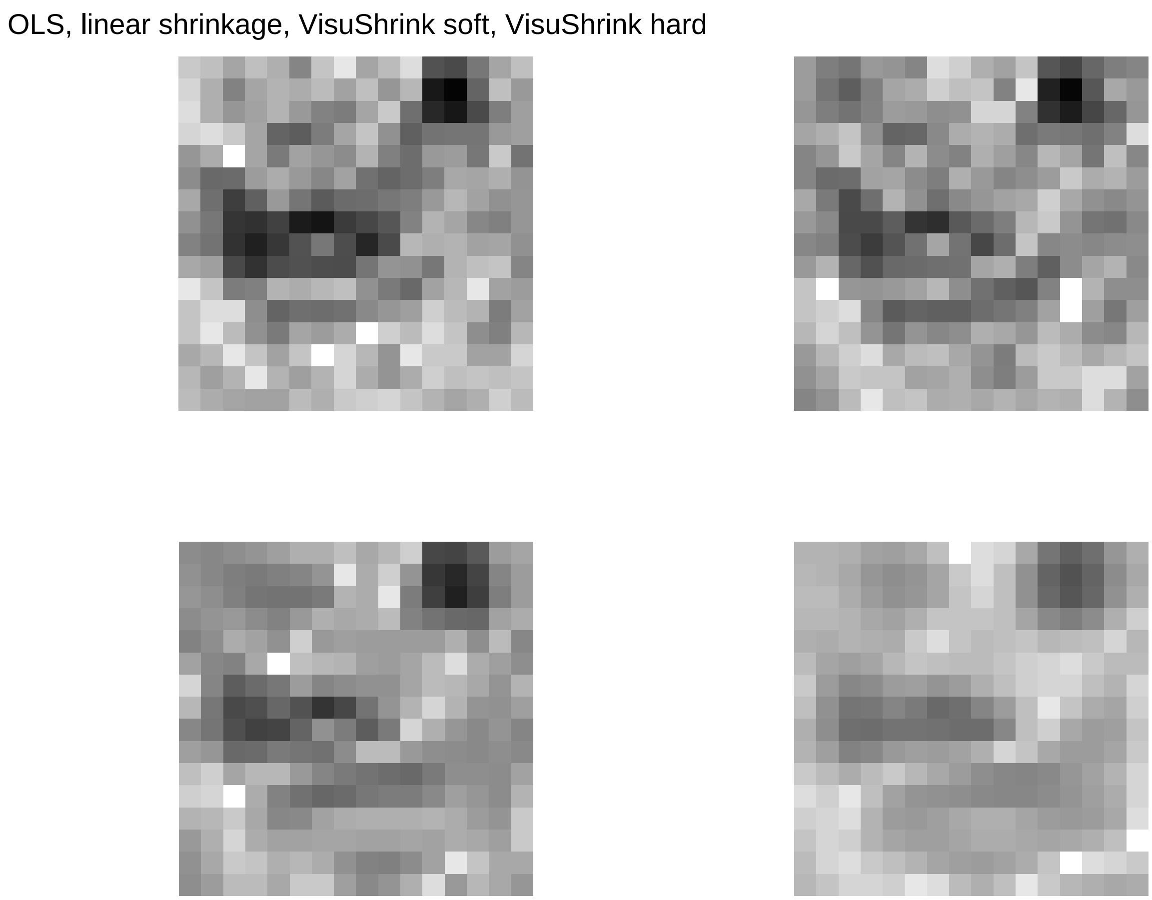

Figure 5.1: *The upper left figure shows the ordinary least squares regression estimate for the coefficient image fitting an arbitrary phase shift sinusoid at the stimulus frequency to each pixel time course. The upper right is an estimate based on linear shrinkage, with heavier penalization of neighboring pixel differences. The lower left figure is the result of VisuShrink shrinkage with soft thresholding, and the lower right displays the result of VisuShrink with hard thresholding*

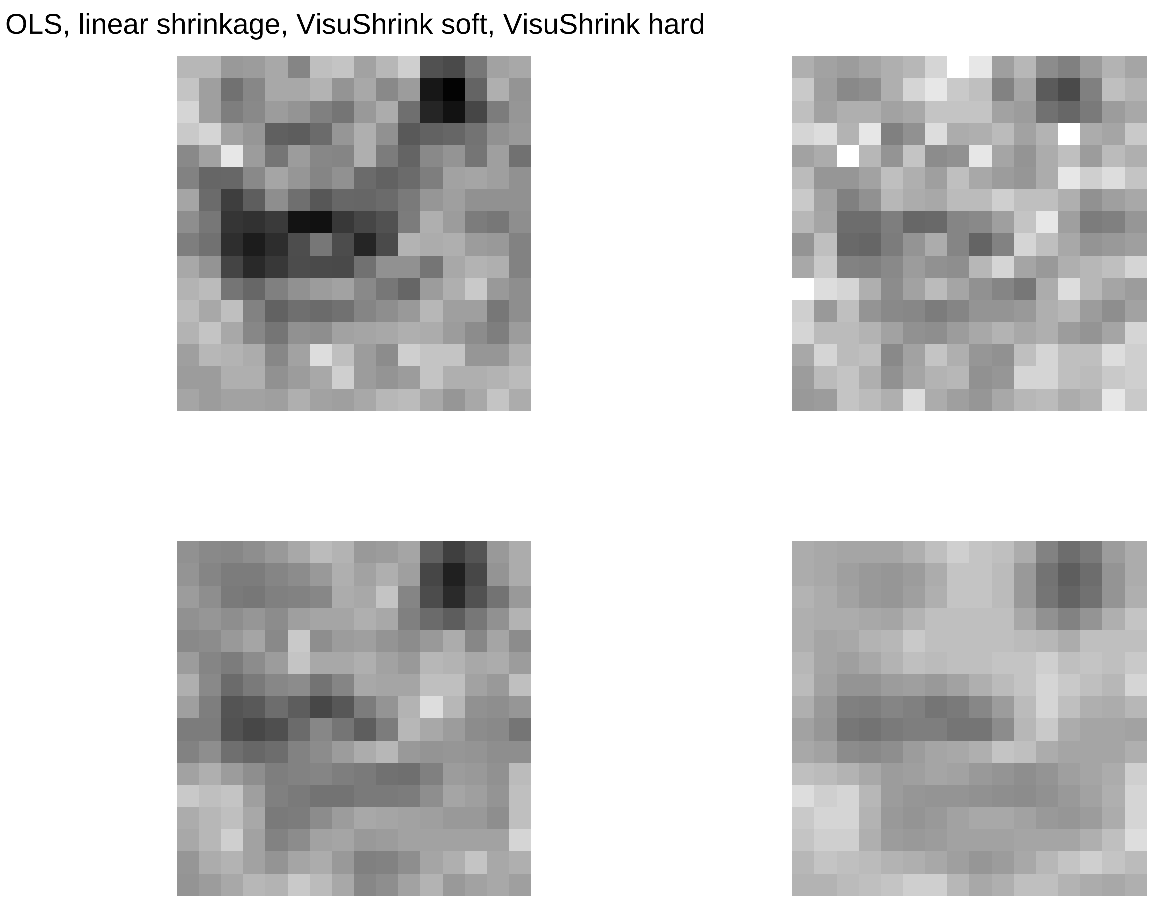

Figure 5.2: *The upper left figure shows the ordinary least squares regression estimate for the coefficient image fitting an arbitrary phase shift sinusoid at the stimulus frequency to each pixel time course. The upper right is an estimate based on linear shrinkage, with heavier penalization of neighboring pixel differences. The lower left figure is the result of VisuShrink shrinkage with soft thresholding, and the lower right displays the result of VisuShrink with hard thresholding*

Chapter 6

Visualization

6.1 Visualization of Image Sequence Data

By their very nature, image sequence data are a time sequence of images collected in a particular order, and as such can be presented as a series of image frames (a contact sheet of images, if you like) or an image movie. Considering the spatial structure first provides an alternative view. At each pixel location, image sequence data records a time series of pixel values over the time course of data collection. As such, we have $M = m_1 \times m_2$ potential time series plots with underlying spatial interrelation to consider.

The following list provides some examples of the sort of visualization structures that are appropriate:

- ▷ Ordered Image Frames as Contact Sheets or Animations
- ▷ Region of interest visualization
    - ▷ Selecting along an area of interest, not necessarily rectangular
    - ▷ Visualizing the time course or summary/descriptive views for selected pixel regions

- ▷ Viewing the time courses of pixels or pixel regions, including spatial structure.
  - ▷ Subsampling pixel space and producing grids of time series
    - ▷ Multiple overlaying plots of pixel time series from that region
    - ▷ Averaging/Smoothing time courses over rectangular region, possibly with variability bands shown

6.2 Ordered Image Frames

For the typical dimensionality of fMRI data sets, laying out images as a contact sheet of sequential images is rarely worthwhile. Care must be taken to display each image with directly comparable color maps, much as one must plot multiple time series on identical scales. Rescaling each image to utilize the entire color range can lead to quite misleading visual comparisons. The increase in signal strength produced by the BOLD contrast is on the order of a few percent of the signal magnitude, and as such can be hard to perceive between successive image views. Indeed, this is the reason there are statistical methods used to detect it.

Animation of the image sequence is a superior technique, subject to many of the same caveats. Each image must be rendered on comparable color scales, and the magnitude of the change may be difficult to perceive visually. The human visual system is however much better at perceiving change when it is presented in this form. Animation is considerably less effective if successive frames are not double-buffered. Where double buffering is not done, each successive movie frame is drawn onto the same area of screen memory, sometimes erasing the previous frame before beginning the next. This visual interruption strongly detracts from the animation. In double-buffering, the next frame is rendered off-screen into a portion of memory that copied into screen memory all at once replacing the previous frame, giving a smooth continuity to the animation. The software package `matlab` provides for animation of image sequences through its `immovie` function. A sequence of images can be made into a movie matrix by first constructing what `matlab` calls an image deck, a matrix with the image frames stored as rows and calling `immovie` with the image deck as argument. Movie matrices are displayed in `matlab` using the `movie` function.

As well as the issue of comparable color scales, the issue of drift may need to be considered. In Figure (2.4) we saw a common artifact of fMRI datasets wherein the mean level at a single pixel appeared to have a linear drift. Taken across the entire image sequence, it may be desirable to detrend each pixel time-course or the mean intensity level of each image prior to analysis or visualization.

6.3 Region of interest visualization

Considering the image sequence data as a spatially organized collection of time courses at each pixel, we may wish to examine how these time courses or summaries derived from them such as model fits, appear at selected pixel locations, or over regions of selected pixels.

Several software visualization tools were written in S-plus for interactive exploratory data analysis as part of this thesis research. The first of these `explore.brain()` allows the user to view a general waveform or set of waveforms associated with a region of interest in the brain by selecting the pixels of the region on the split screen display at left showing brain response, with the resulting waveforms displayed at right. Figure (6.1) and (6.2) show examples of the sorts of displays produced.

Multiple waveforms can be superimposed or representative aggregates computed to be displayed. This is accomplished by specifying a function argument which produces the waveform(s) of interest. In the example in Figure (6.2) that function has produced as output the pair of waveforms consisting of the pixels time course and the

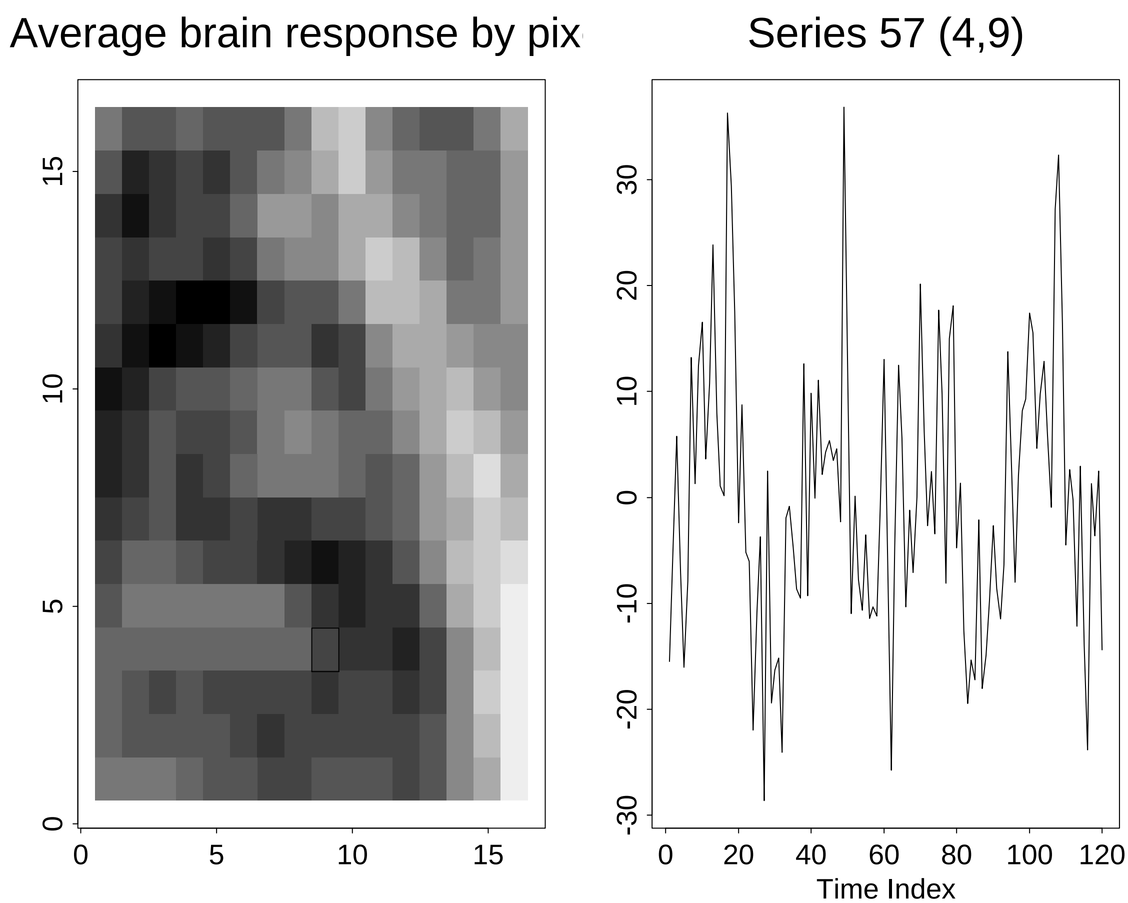

Figure 6.1: *Use of* `explore.brain` *to simply show the time course at a selected pixel.*

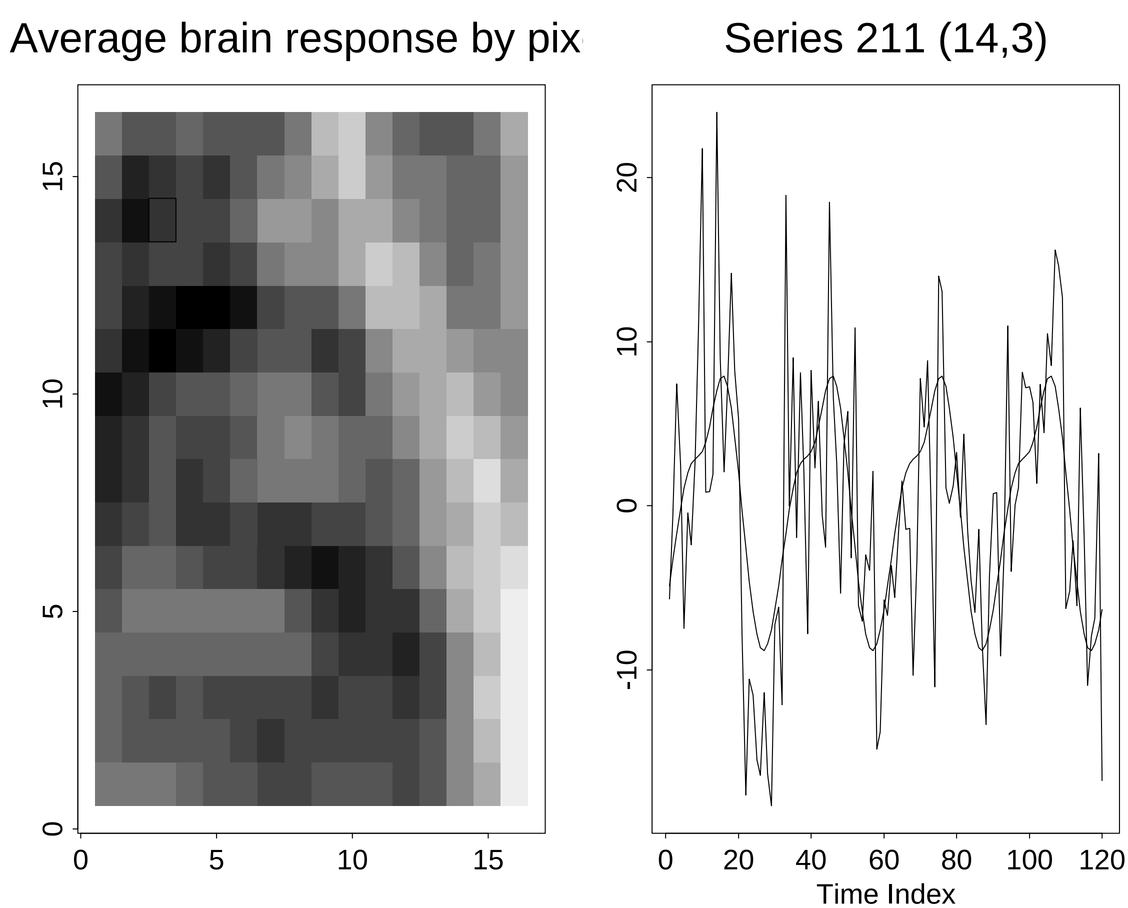

Figure 6.2: *Use of* `explore.brain` *to show the time course and a periodic spline fit at a selected pixel.*

periodic spline fit to that data. The function argument `FUN` is passed the pixel indices which the user has selected as of interest and may compute as many waveforms per pixel as the user desires.

Pixels are selected as of interest by selecting them with the left-most mouse button on the brain representation at left. The optional argument `add=T`, true by default allows successive addition of pixels to the region of interest. The function `FUN` can select from among the pixels in the region of interest, create a regional summary such as averaging across the regions pixels, or can compute single or multiple waveforms for each pixel selected. Figures (6.3) and (6.4) give examples of region of interest summaries computed this way.

Clicking outside the brain response region resets the region of interest to an empty set. The selected pixels are identified on the brain response display. For exploration on a pixel by pixel basis without having to repeatedly rest, the `add` argument can be set to false. Clicking the middle mouse button ends the interactive exploration.

6.4 Spatial structure summaries

Exploratory data interaction which automatically shows the structure of spatially neighboring pixels is also of interest. As seen in previous chapters, neighboring pixels may have similar time course behaviors or model fits. One could use `explore.brain` to select from pixels of interest and their spatial neighbors by computing the relevant pixel indices and plot the relevant summary curves, but the display of `explore.brain` does not accommodate a spatially arranged presentation of the results. For this purpose, an alternative software visualization tool was developed.

The S-plus function `boxes.brain()` tessellates the brain area into rectangular tiles and facilitates exploration of regional behavior with an equal number of panes across and down specified by argument `npanes`. Analogously with `explore.brain()` the

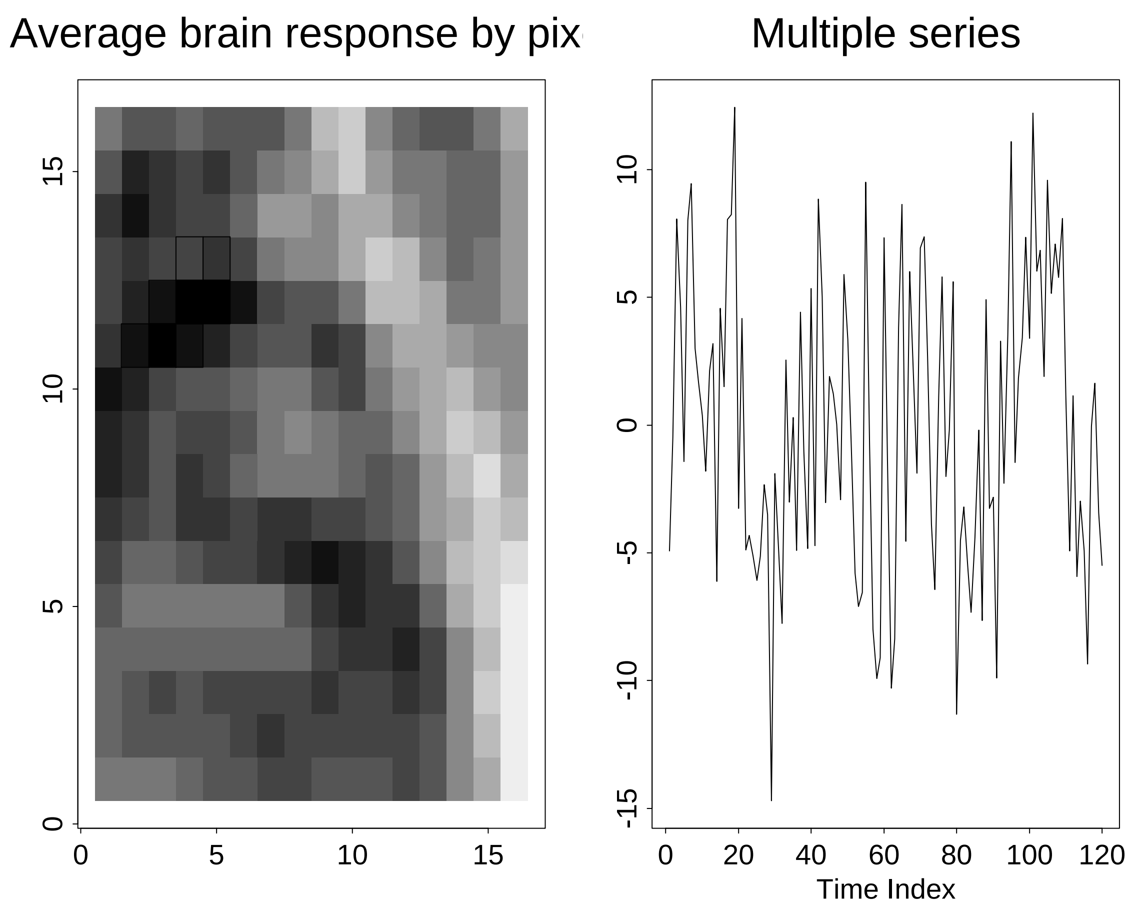

Figure 6.3: *Use of* `explore.brain` *to show the average time course averaged over the pixels in the selected region of interest.*

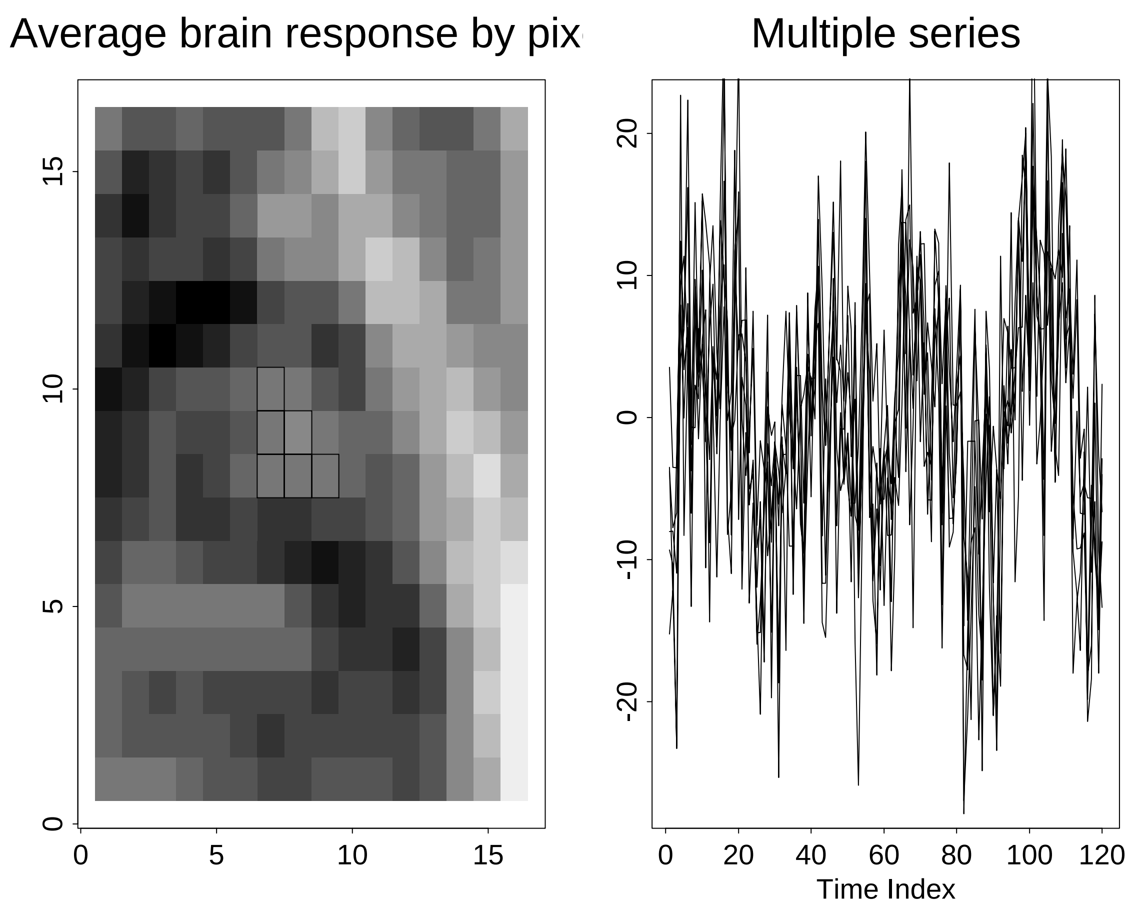

Figure 6.4: *Use of* `explore.brain` *to show all the pixels time courses for the selected region of interest.*

function argument `FUN` provides the waveform(s) output for all of the pixels contained in each pane to be represented in the corresponding pane on the right side of the display. Again, this might consist of the entire set of raw data or fitted responses of that region of the brain, or some representative computed waveform for that region. Clicking on a pane on the left zooms the tessellation to subdivide that pane, and the corresponding displays on the right are updated accordingly. The procedure may be iterated down to the resolution of individual pixels. Clicking outside the current set of panes but within the brain area zooms back out to the next lower scale. Clicking the middle mouse button ends the interactive exploration. Figure (6.5) below shows an example of the function in use.

I have found the combination of these two functions to be quite general for exploratory visualization of the pixelwise time-series in the fMRI brain mapping image sequence data examined as part of this thesis.

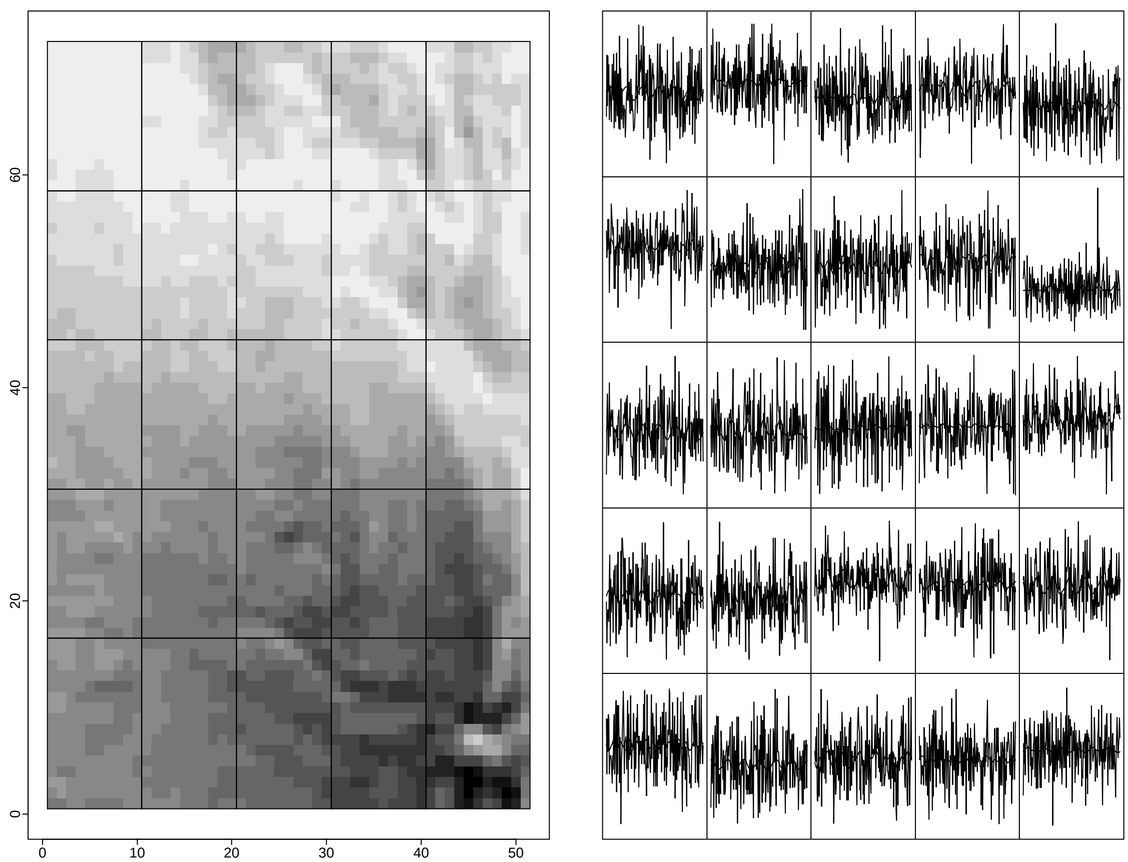

Figure 6.5: *Use of* `boxes.brain` *showing representative time course summaries of the image subtiling indicated.*

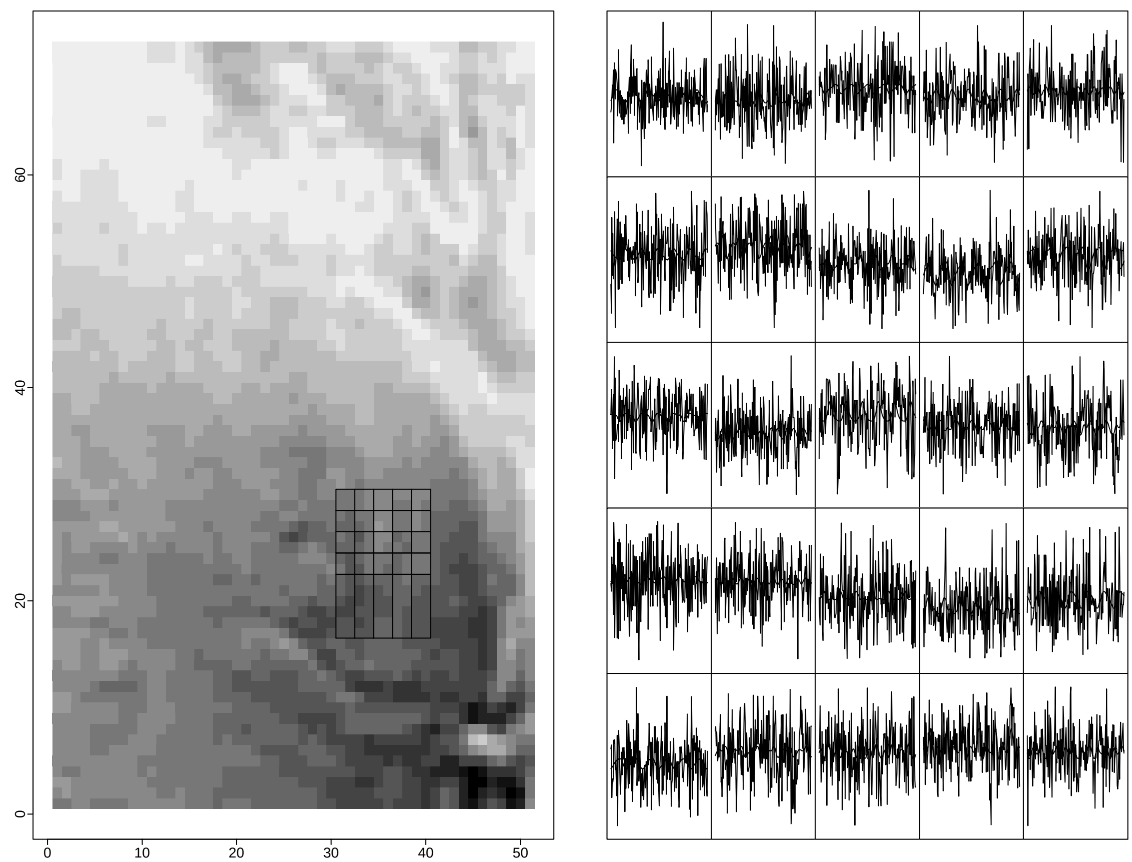

Figure 6.6: *Use of* `boxes.brain` *showing representative time course summaries of the image zoomed into to pixel resolution. The lower subtiles include several pixels each.*

Chapter 7

Inference for fMRI image sequence models

Consider any statistical parametric map (SPM) constructed in any of the previously discussed chapters. These may be formed from normalized cross-covariances, correlations with best-fitting sinusoids with respect to amplitude and phase, or coefficient estimate maps from hemodynamic, basis expansion, or regularized regression models. How does one select a critical value threshold for assessing significance of the value of the map, noting that (i) there are a large number (M) of simultaneous inferences to be made, and (ii) there is spatial and temporal autocorrelation in the data. In this chapter we review the literature of statistical inference procedures and how they account for these considerations.

7.1 Literature

7.1.1 Early Proposals

Bandettini et al [3] identify pixels with significant activation by thresholding the correlation maps they compute. They assume that the sampling distribution of the correlation coefficient is approximately normal, and relate threshold u to statistical significance p by the formula

$$p = 1 - \frac{2}{\sqrt{\pi}} \int_0^{u\sqrt{n/2}} e^{-t^2} dt$$

where n is the number of images in the series. No adjustment for multiple comparisons is proposed in this paper. Later papers describing their technique, such as De Yoe et al [10] and Binder and Rao [7] make reference to an unspecified but conservative correction and a Bonferroni correction respectively.

Much of the statistical methodology in fMRI followed out of earlier work done with PET. Several papers from the PET literature are therefore included in this review.

Fox et al [18], as described earlier, propose an omnibus test of activation based on the kurtosis measure gamma-2. ($\gamma_2 = \frac{\sum(x-\bar{x})^4}{n\hat{\sigma}^4} - 3$). They assume that γ_2 follows a t distribution for n sufficiently large (citing Zar [69] that $n > 1000$ suffices) or using significance tables such as Snedecor and Cochran [58]. Their technique included local maxima sampling as a preprocessing step to select the region of interest. Pixels were selected for consideration only if their mean change in rCBF exceeded all 26 adjacent voxels and the local maximum location was estimated using a center of mass calculation over a 14mm spherical region of interest. Consequently, the sampling distributions of γ_2 for the data from these local maxima areas exhibited bimodal symmetry about zero. They decoupled the distributions of positive and negative γ_2, computing what they call one-sided γ_2 statistics. They conclude that the one-sided γ_2 statistic for leptokurtosis (positive γ_2) was "both sensitive to and specific for stimulus-induced focal brain responses." This procedure is however an omnibus test for activation. As Fox et al note: "Omnibus testing, by g_2 statistic, determines whether any regional changes within a subtracted image achieved significance as distribution outliers. It does not, however, identify *which* focal changes are significant, physiological responses rather than image noise." They propose post-hoc analyses based on z-scores of standardized response magnitude assuming normality to identify regions of significant activation. Multiple comparison adjustments are not considered, except insofar as the sample size of their one-sided γ_2 distributions are based on the number of local maxima found, and this is accounted for.

7.1.2 Random field theory results

Friston et al [26] promote the use of statistical parametric maps and stationary Gaussian field theory for assessing significant change due to neural activity in PET scans. They observe that "the assessment of significance is confounded by the large number of pixels and by image smoothness" meaning "a large number of comparisons are made that are not independent." They note that due to the image smoothness imposed by the PET modality, Bonferroni correction to ensure that the probability of at least one pixel above the threshold is likely to overcorrect. They propose a "smoothness adjustment" to be used in conjunction with a Bonferroni-style adjustment for the number of pixels. Their smoothness adjustment requires that a smoothness parameter s based on the variance of the first partial derivatives of the image process be estimated, where

$$s = \sqrt{1/(2 \cdot S_{\partial}^2)} \tag{7.1}$$

They estimate S_{∂}^2 numerically, approximating the partial derivative in a direction by the difference between a pixel and its neighbor and "calculating the variance of the pooled interpixel differences along rows and columns of the SPM."

They assume that under the null hypothesis of no activation, the SPM is a stationary, isotropic, Gaussian random field, produced by the convolution of "a completely random uncorrelated process with a Gaussian filter of width s." In addition, they assume that "(a) pixel size is small relative to s and (b) s is small relative to the size of the SPM." The probability of a false positive (Ω) using a particular threshold (u) is heuristically argued to be approximately

$$\Omega \approx \left[32\pi \cdot s^2 \cdot e^{u^2} \cdot \overline{\Phi}(u)\right]^{-1} \tag{7.2}$$

where $\overline{\Phi}(u) = 1 - \Phi(u)$ with $\Phi(u) = \Pr(Z \leq u)$ for $Z \sim N(0,1)$. To make the appropriate multiple comparison correction for assessing significant activation within an SPM, they recommend that for achieving a significance level α, the threshold u should be chosen so that Ω in equation (7.2) equal α/N where N is the number of pixels. The results of this paper only apply to two-dimensional data, or single slices, and does not provided overall control of Type I error for multi-slice data.

Worsley et al [66] use the theory of random Gaussian fields and compute a more mathematically rigorous adjustment. For the three-dimensional t statistic image calculated from normalized differences in CBF from PET data, they assume the null distribution of the image field "is well approximated by a standard Gaussian distribution." As an aside, they cite results from Adler and Hasofer [2] and Adler [1, p.111] for the maxima of a two-dimensional slice which shows that the heuristic arguments in Friston et al [26] are incorrect by a factor of approximately $\pi/4$. The main result of this paper however, which the authors attribute to a result by Adler and Hasofer [2], is that

$$\Pr(T_{max} > t) \approx V|\Lambda|^{\frac{1}{2}}(2\pi)^{-2}(t^2-1)e^{-\frac{1}{2}t^2} \tag{7.3}$$

where V is the search volume "is large relative to the smoothness of the image", and the matrix Λ is the "3×3 variance matrix of the partial derivatives of the random field in each of the three variables x,y, and z and measures the roughness of the image." Assuming that the image under no activation can be generated by convolving a white noise Gaussian random field with a Gaussian kernel, where the principal axes of the kernels covariance matrix align with the x,y, and z direction, the determinant of matrix Λ simplifies, giving the formula

$$|\Lambda|^{\frac{1}{2}} = (FWHM_x FWHM_y FWHM_z)^{-1}(4\ln 2)^{\frac{3}{2}} \tag{7.4}$$

where $FWHM_.$ is the full width at half maximum of the kernel in the corresponding direction. The full width at half maximum is the width of the kernel at half its maximum value. Dividing the search volume V by the product of the full widths at half maximum, the authors define as the number of resolution elements or "resels" in the search volume, and can use $R = V/(FWHM_x FWHM_y FWHM_z)$ to simplify formulae used for threshold selection. (In fact, an extra step of linear interpolation in the z direction was employed to exploit interplanar smoothness, and the effective $FWHM_z$ of the interpolated, resampled voxels used, but the general principle remains identical.)

The result in equation (7.3) is derived based on the Euler characteristic of the excursion set above the threshold. Several definitions of the Euler characteristic exist. Hasofer [40] defines the Euler characteristic as an additive, integer-valued functional $\phi(A)$ over compact subsets $A \subset I\!R^n$, such that $\phi(A \cup B) = \phi(A) + \phi(B) - \phi(A \cap B)$, where additionally if $A \subset I\!R^n$ is homeomorphic to an n-dimensional sphere, $\phi(A) = 1$. A precise definition which can be used to evaluate the Euler characteristic involves counting contributions based on the curvature at critical points on the boundary of the set. The Euler characteristic is an inherently topological concept, effectively counting the number of isolated connected components of a set, minus the number of 'holes'.

For threshold values t near the global maximum of the random field, the Euler characteristic χ_t will equal zero if $T_{max} < t$ and one if $T_{max} > t$. Hasofer [40] shows that as $\Pr(\chi_t \geq 1) \to 0$ as $t \to \infty$, $\Pr(T_{max} > t) = \Pr(\chi_t \geq 1) \approx E(\chi_t)$ and this bound is quite close for high levels of the threshold t. The result of Adler and Hasofer used by Worsley et al is

$$\begin{aligned} E(\chi_t) &= V|\Lambda|^{\frac{1}{2}}(2\pi)^{-2}(t^2-1)e^{-\frac{1}{2}t^2} \\ &= R(4\ln 2)^{\frac{3}{2}}(2\pi)^{-2}(t^2-1)e^{-\frac{1}{2}t^2} \end{aligned}$$

The results of Worsley et al [66] provide exact results for a Gaussian random field computed using a pooled standard deviation across all voxels. The paper notes that the t-statistic field computed using each voxels standard deviation in the denominator has a different distribution and they suggest an approximation using the inverse probability transform to convert the field to approximate normality with a multiplicative correction factor to adjust equation (7.3). This approximate technique remains widely used largely due to its implementation in the freely distributed SPM software package [59] for matlab [47]. However, exact results for the expected Euler characteristic of the excursion set for χ^2, F, and t fields have been derived by Worsley [64].

Siegmund and Worsley [57] consider the similar but not quite identical problem of testing for a fixed signal of unknown location and amplitude within a stationary Gaussian random noise field in $\mathbb{R}^n$ where the signal width is also unknown. They are able to use results based on the Hadwiger characteristic of differential topology as well as a differential geometry approach based on the volume of tubes to characterize the null distribution of a likelihood ratio test and compute power against specific alternatives. The Hadwiger characteristic is defined on different classes of sets than the Euler characteristic, but is equal to it for all sets in the domain of definition of both. It has an iterative definition more amenable to statistical analysis, which facilitates Siegmund and Worsley's theory. They apply their results to a PET data set as an example in their paper.

Worsley [65] refines the theoretical results available for estimating the number of peaks in a random field even further. He notes that the results previously published in Worsley et al [66] are actually based on the IG (integral geometry) characteristic. He also notes that the results proved by Adler [1] are for the DT (differential topology) and IG characteristics. Both are equal to the Euler characteristic only provided that the set does not touch the boundary of the volume. Neither of these characteristics are invariant under reflections or rotations of the co-ordinate axes, which would present problems in a non-stationary or anisotropic situation. In Worsley et al [66], the authors attempt to handle this by averaging the IG characteristic over all reflections of the co-ordinate axes, which introduces a problem of interpretability of fractional values. Use of the Hadwiger characteristic is championed because "it avoids all these problems." "It takes integer values, it counts regions near the boundaries, it is defined on any set C, and it is invariant under any rotations or reflections." In an example where the method is applied to PET study data containing activation, use of the Hadwiger characteristic is seen to identify regions missed by the averaged IG criterion because of their proximity to the volume boundary.

7.1.3 Activation by spatial extent or cluster size

Several authors have observed that the threshold selection and multiple comparison correction formulae are based on the Type I error for detecting activation at a single pixel, whereas typically activation is only of interest if it occurs in clusters of multiple neighboring pixels. A variety of techniques accounting for the spatial extent of clustered pixels activated have been proposed.

Poline and Mazoyer [55] observe that functional images typically have low signal to noise ratios and that filtering (smoothing) is often used to improve this, at the expense of spatial resolution. They propose a method based on detecting high signal-to-noise pixel clusters (HSC) and comparing their sizes to a Monte-Carlo derived distribution of cluster sizes in pure noise images. Detection of the HSC pixels within a slice of data is achieved by first performing low-pass spatial filtering (either the average of a 3 pixel radius disk or a Gaussian filter with a 6.1 pixel FWHM was used), then computing a signal-to-noise ratio image based on the average signal images and a standard deviation image formed from control state scans, followed finally by thresholding at critical points of the standard normal distribution and identifying four-connected clusters above the threshold.

Their Monte-Carlo simulation of the HSC size distribution in noise-only images models temporal autocorrelation to account for the between-scan covariance. The temporal variance-covariance matrix they use is proportional to $(v-c)I+cJ$ where v is the pixel variance, c is the between-scan covariance (both estimated from the experimental data), I is the identity matrix and J is the matrix of all ones. Lange [43] notes that this is the same as Fishers intraclass correlation matrix which has an analytically exact inverse formula, where Poline and Mazoyer use numerical Jacobi diagonalization. Assuming HSC occurrences within noise images are well modeled by a Poisson process, Poline and Mazoyer simulate such images, estimate the process intensities for particular levels of cluster size and threshold in the simulated noise images. They then use a sequential procedure to control overall Type I error, accepting the largest size HSC first if the probability of such a size cluster under the assumed Poisson distribution with corresponding simulated intensity is less than the desired α. They then reduce the desired level α by a Bonferroni adjustment and repeat for the next largest HSC, stopping when a non-significant HSC is reached. They compare their procedure with the γ_2 method of Fox, and find it compares quite favorably, with the additional benefit of localizing regions of significant activation.

Friston et al [34] revisit the theory of Gaussian random fields to answer the question of how to select a threshold indicating significant activation such that the probability of the activation focus containing k voxels or more is approximately α. In order to summarize their results, it is necessary to introduce some of their notation. For a given threshold value u, the SPM contains a total of N voxels exceeding the threshold, m activation regions (connected subsets of the excursion set), and n is the number of voxels within any activated cluster. They seek to approximate the probability that the largest region has k voxels or more: $\Pr(n_{max} \geq k)$. Their previous results in Friston et al [26] and Worsley et al [66] are noted to solve this problem for $k = 1$, with $\Pr(n_{max} \geq 1) = \Pr(T_{max} \geq u) = \Pr(m \geq 1) \approx E(m)$ which can be approximated using the expected value of the Euler characteristic for high values of u.

First, the authors note that the expectations of N, m, and n satisfy

$$E(N) = E(m)E(n)$$

and that the following formulae give the expectation of each for a D-dimensional process of volume S and threshold u:

$$\begin{aligned} E(N) &= S \cdot \Phi(-u) \\ E(m) &= S(2\pi)^{-(D+1)/2} W^{-D} u^{D-1} e^{-u^2/2} \\ E(n) &= E(N)/E(m) \end{aligned}$$

where W is the smoothness factor corresponding to the determinant of the covariance matrix of the partial derivatives of the stationary Gaussian process. If the process is formed by a Gaussian point spread function with full width at half maximum in the i^{th} co-ordinate direction of $FWHM_i$, then

$$W = (4 \ln 2)^{-1/2} \prod_{i=1}^{D} FWHM_i^{1/D}$$

The expectation of N follows from the Gaussian field assumption, and the expectation of m follows from the results of Hasofer [40] as previously discussed.

The authors then make use of two approximation assumptions for the probability distributions of m and n. The distribution of m is assumed to be Poisson with mean given by $E(m)$ computed as above. The authors cite Adler [1, Theorem 6.9.3, p.161] who shows the Poisson distribution of m holds in the limit for high thresholds. For the distribution of n, they cite asymptotic results due to Nosko [51, 52, 53] reported in Adler [1, p. 158] that for large thresholds, $n^{2/D}$ has an exponential distribution with

$$E(n^{2/D}) \approx \frac{2\pi W^2}{u^2 \Gamma(D/2+1)^{2/D}}$$

They then assume that

$$\Pr(n = x) \approx \frac{2\beta}{D} x^{2/D-1} e^{-\beta x^{2/D}}$$

or

$$\Pr(n \geq x) \approx e^{-\beta x^{2/D}}$$

so that $n^{2/D}$ is exponentially distributed with expectation $1/\beta$, and

$$E(n) = \Gamma(D/2 + 1) \cdot \beta^{-D/2}$$

where β is given by

$$\beta = [\Gamma(D/2 + 1) \cdot E(m)/E(N)]^{2/D}$$

This gives the necessary machinery to compute the desired probability $\Pr(n_{max} \geq k)$ by partitioning and conditioning on all values of m and computing the probability of the complementary event "that all m regions have less than k voxels." This leads to the formula:

$$\begin{aligned}
\Pr(n_{max} \geq k) &= \sum_{i=1}^{\infty} \Pr(m = i) \cdot [1 - \Pr(n < k)^i] \\
&\approx 1 - e^{-E(m) \cdot \Pr(n \geq k)} \\
&\approx 1 - \exp(-E(m) \cdot e^{-\beta k^{2/D}})
\end{aligned}$$

This equation is then inverted to compute the critical region size k_α such that $\Pr(n_{max} \geq k) = \alpha$ giving the expression

$$k_\alpha \approx [\log(-E(m)/\log(1 - \alpha))/\beta]^{D/2}$$

Using the results above facilitates an assessment of significant activation based on cluster size or volume while still controlling Type I error. The authors note: "A spatially extensive region of activation may not necessarily reach a high threshold (e.g., 3.9), it may be significant (at $P = 0.05$) is assessed at a lower threshold (e.g., 3)."

Validation of the formulae and approximations was facilitated by simulation of Gaussian random fields as well as using data from a PET phantom data set, and showed results in good agreement with theory. A theoretical power analysis showed that power either increased or decreased with threshold depending on the width of the signal to be detected. For signals narrower than 70 percent of the resolution of the underlying SPM, power increased with threshold, and for signals broader than 70 percent of the resolution, power increased as the threshold was lowered. This suggests that "narrow focal activations are most powerfully detected by high thresholds, ... whereas broader more diffuse activations are best detected by low thresholds." The use of lower thresholds "is predicated on the assumption that real signals are broader than the resolution." For PET data, this is quite reasonable, however this may give pause when applied to fMRI data. The authors reference the paper by Siegmund and Worsley [57] as an attempt to resolve this "by searching over tuning parameters (like smoothing) as well as voxels."

Forman et al [14] also laud the use of cluster size thresholds as a means of improving power and sensitivity in detecting focal activation. They acknowledge the work of Friston et al [34] with the criticism that "the approximations used appear largely inapplicable to fMRI data sets." They model the spatially correlated SPM data "as if a Gaussian filter of width, s, had been applied to form the observed SPM." They note the formula proposed in Friston et al [26] for estimating s (Equation (7.1) above) "assumes that pixel size is small relative to s", an assumption not typically held in fMRI data. They derive an alternative formula with which to estimate s, based on careful integration assuming a bivariate Gaussian filter of width s yielding

$$s = \sqrt{-1 \Big/ \left(4 \ln \left(1 - \frac{S_\partial^2}{2S^2}\right)\right)} \tag{7.5}$$

where S^2 is the individual pixel variance across the SPM. Comparisons between estimates of s produced using equations (7.1) and (7.5) for simulated SPM data sets smoothed with Gaussian kernels of known width showed equation (7.5) to be quite superior, particularly for filter widths smaller than one pixel wide, which the authors observed in their fMRI SPMs.

Assessment of significant activation under their method is accomplished by selecting both a cluster size threshold and an acceptable level of α, the per pixel permissible error rate. For a given level of α and estimated smoothness, s, they use Monte Carlo simulation to characterize the frequency versus cluster size distribution, which are then converted into false positive probability per pixel. A selection of results for various cluster sizes and levels of α and smoothnesses considered typical for fMRI data are tabulated. Power analyses are described based on similar simulation studies. An example fMRI study considered identifies nearly five times as many pixels as active than would the equivalent Bonferroni procedure, and a set of PET data is examined to show that the identified pixel regions correspond to areas of increased blood flow.

Worsley et al [68] propose a test for detection of distributed, non-focal activation based on the average sum of squares of a stationary smooth Gaussian SPM. Their omnibus test is more sensitive than the γ_2 statistic proposed by Fox et al [18] for spatially distributed signals, and has the benefit of a theoretical rather than empirical underpinning for its claims of specificity. The test statistic, the average sum of squares of voxels in the search volume, is appropriately scaled and compared against a χ^2 distribution. The effective degrees of freedom varies with the smoothness of the SPM. A previously proposed test based on the proportion of the search volume above a threshold, proposed in Friston et al [25], is also corrected in this paper. The authors observe that their test for diffuse, non-focal activations "could be used as a prelude to variance partitioning procedures that do not allow statistical inference, such as singular value decompositions and eigenimage analysis."

7.1.4 Non-parametric tests

Holmes et al [41] describe non-parametric randomization tests for assessing the statistics images commonly used in analyzing functional mapping experiments. They provide compelling reasons to consider non-parametric analysis methods, listing several shortcomings of the commonly used approaches based on random field theory. These include the following: Low degrees of freedom resulting in highly variably variance estimates dominating calculation of the t-field SPM. The estimated smoothness of the statistic image may lead to poor approximation by rough random fields whose features have smaller spatial extent than the voxel dimensions. Statistic images are sometimes computed using a global variance estimate as in Worsley et al [66], an approximation that has been criticized in the literature for physical and physiological reasons. Smoothing by Gaussian filters, as sometimes used in preprocessing SPMs to improve the smooth random field approximation, may be inappropriate, smoothing out spatially localized activations of smaller extent. As the authors write: "When the assumptions of the parametric methods are true, these new non-parametric methods, at worst, provide for their validation. When the assumptions of the parametric methods are dubious, the non-parametric methods provide the only analysis that can be guaranteed valid and exact."

They consider t-statistic images, and a generalization they call "pseudo" t-statistics based on a smoothed variance estimate. The smoothed variance estimate computed by smoothing the voxel-wise variance estimate image by a Gaussian kernel of known full-widths at half maximum in each co-ordinate direction. As an example, they first describe the simplest case of an alternating stimulus regime of two conditions. They point out that a non-parametric randomization distribution applied to the labeling of conditions to scans can be used to test hypotheses of differences between conditions, in either per-pixel, omnibus, or regional activation tests. Under the null hypothesis of no difference between stimulus conditions, statistic images based on any relabeling of conditions to scans are equally likely. To control Type I error on the omnibus test, "the probability of observing a statistic image with maximum intracerebral value as or more extreme that the observed value T_{max} is simply the proportion of randomization values greater than or equal to it. This gives a p-value for the omnibus null hypothesis." The possible values of the maximum statistic under each possible labeling is computed, and an approximate level α test is constructed by taking the $100(1-\alpha)^{th}$ percentile of those values as threshold. To identify activated voxels using the randomization distribution, an adjusted p-value image is computed, replacing each voxels value with the proportion of the randomization distribution of the maximal statistic greater than its t-statistic value. Any pixel with a value greater than α has a pixel value greater than the level α threshold. The combinatorial complexity of enumerating the maximal statistics randomization values is reduced in this case by noting that reversing labels yields the negative value of the maximal statistic, halving the number of cases to be considered.

The presence of more than one activation, they further note, may skew the distribution of randomization values of the maximal statistic, increasing the critical threshold value of the test. This would lead to the above test being conservative for detecting secondary areas of activation. To counteract this effect, they propose a sequentially rejective test adapted to randomization testing. To achieve this, they iteratively compute a step-down sequence of adjusted p-value images. In effect, once a voxel is considered active, it is removed from consideration and an adjusted sampling distribution of the maximal statistic is computed from the remaining voxels. At each step, pixels with adjusted p-values less than or equal to α are added to the set of activated pixels. Overall Type I error is strongly controlled, and an efficient implementation due to Westfall and Young [62] makes the computations more feasible.

They apply the technique to data from a PET study, and find that it detects regions of activation slightly larger than those found by the technique of Friston et al [26]. Analyzing the t-field using the expected Euler characteristic results of Worsley [64] yielded only a small number of pixels in a primary activation region, where the other techniques located secondary regions. In their discussion, they describe how to apply randomization methods to other experimental designs and test statistics used in functional mapping experiments, including many of the techniques referred to above. They note that the enumeration of all possible labelings may be computationally impractical, but point out that a sufficiently large random sampling from all possible labelings will approximate the randomization distribution quite well for very little power loss if around 1000 samples are taken. In a final comment, they remark that randomization methods can be applied to small and specific regions of interest, where random field assumptions and approximations would not apply.

Bibliography

[1] R.J. Adler. *The Geometry of Random Fields.* Wiley, New York, 1981.

[2] R.J. Adler and A.M. Hasofer. Level crossings for random fields. *Annals of Probability*, 4:1–12, 1976.

[3] P.A. Bandettini, A. Jesmanowicz, E.C. Wong, and J.S. Hyde. Processing strategies for time-course data sets in functional MRI of the human brain. *Magnetic Resonance in Medicine*, 30(2):390–397, 1993.

[4] P.A. Bandettini, E.C. Wong, R.S. Hinks, R.S. Tikofsky, and J.S. Hyde. Time course EPI of human brain function during task activation. *Magnetic Resonance in Medicine*, 25(2):390–397, 1992.

[5] J.W. Belliveau, D.N. Kennedy, R.C. McKinstry, B.R. Buchbinder, R.M. Weisskoff, M.S. Cohen, J.M. Vevea, T.J. Brady, and B.R. Rosen. Functional mapping of the human visual cortex by magnetic resonance imaging. *Science*, 254:716–719, November 1991.

[6] J.W. Belliveau, K.K. Kwong, D.N. Kennedy, J.R. Baker, C.E. Stern, R. Benson, D.A. Chesler, R.M. Weisskoff, M.S. Cohen, R.B.H. Tootell, P.T. Fox, T.J. Brady, and B.R Rosen. Magnetic resonance imaging mapping of brain function: Human visual cortex. *Investigative Radiology*, 27(suppl. 2):SS59–S65, 1992.

[7] J.R. Binder and S.M. Rao. *Human Brain Mapping with Functional Magnetic Resonance Imaging*, chapter 7. Academic Press, 1994.

[8] G.M. Boynton, S.A. Engel, and D.J. Heeger. Linear systems analysis of fMRI. *Submitted*, 1996.

[9] C. de Boor. *A Practical Guide to Splines.* Applied Mathematical Sciences. Springer-Verlag, New York, 1978.

[10] E.A. DeYoe, P. Bandettini, J. Neitz, D. Miller, and P. Winans. Functional magnetic resonance imaging (FMRI) of the human brain. *Journal of Neuroscience Methods*, 54:171–187, 1994.

[11] D.L. Donoho and I.M. Johnstone. Adapting to unknown smoothness via wavelet shrinkage. Technical Report 425, Statistics Department, Stanford University, June 1993.

[12] D.L. Donoho and I.M. Johnstone. Ideal spatial adaptaion by wavelet shrinkage. *Biometrika*, 81:425–455, 1994.

[13] S.A. Engel, D.E. Rumelhart, B.A. Wandell, A.T. Lee, G.H. Glover, E.-J. Chichilnisky, and M.N. Shadlen. fMRI of human visual cortex. *Nature*, 369:525, June 1994.

[14] Steven D. Forman, Jonathan D. Cohen, Mark Fitzgerald, William F. Eddy, Mark A. Mintun, and Douglas C. Noll. Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): Use of a cluster-size threshold. *Magnetic Resonance in Medicine*, 33(5):636–647, 1995.

[15] Peter T. Fox, Harold Burton, and Marcus E. Raichle. Mapping human somatosensory cortex with positron emission tomography. *Journal of Neurosurgery*, 67:34–43, 1987.

[16] Peter T. Fox, Francis M. Miezin, John M. Allman, David C. Van Essen, and Marcus E. Raichle. Retinotopic organization of human visual cortex mapped with positron-emission tomography. *The Journal of Neuroscience*, 7(3):913–922, March 1987.

[17] Peter T. Fox, Mark A. Mintun, Marcus E. Raichle, Francis M. Miezin, John M. Allman, and David C. Van Essen. Mapping human visual cortex with positron emission tomography. *Nature*, 323:806–809, October 1986.

[18] P.T. Fox, M.A. Mintun, E.M. Reiman, and M.E. Raichle. Enhanced detection of focal brain responses using intersubject averaging and change-distribution analysis of subtracted PET images. *Journal of Cerebral Blood Flow and Metabolism*, 8(5):642–653, 1988.

[19] P.T. Fox, J.S. Perlmutter, and M.E. Raichle. A stereotactic method of anatomical localization for positron emission tomography. *Journal of Computer Assisted Tomography*, 9:141–153, 1985.

[20] J. Frahm, K.-D. Merboldt, and W. Hänicke. Functional MRI of human brain activation at high spatial resolution. *Magnetic Resonance in Medicine*, 29(1):139–144, 1993.

[21] D.A. Freedman. As others see us: A case study in path analysis. *Journal of Educational Statistics*, 12(2):101–128, Summer 1987. Responses and rejoinders continue to page 223.

[22] K.J. Friston. Functional and effective connectivity in neuroimaging: A synthesis. *Human Brain Mapping*, 2(1–2):56–78, 1994.

[23] K.J. Friston. Statistical parametric mapping: Ontology and current issues. *Journal of Cerebral Blood Flow and Metabolism*, 15(3):361–370, 1995.

[24] K.J. Friston, C.D. Frith, R.S.J. Frackowiak, and R. Turner. Characterizing dynamic brain responses with fMRI: A multivariate approach. *NeuroImage*, 2(2):166–172, June 1995.

[25] K.J. Friston, C.D. Frith, P.F. Liddle, R.J. Dolan, A.A. Lammertsma, and R.S.J Frackowiack. The relationship between global and local changes in pet scans. *Journal of Cerebral Blood Flow and Metabolism*, 10(4):458–466, 1990.

[26] K.J. Friston, C.D. Frith, P.F. Liddle, and R.S.J. Frackowiak. Comparing functional (PET) images: The assessment of significant change. *Journal of Cerebral Blood Flow and Metabolism*, 11(4):690–699, 1991.

[27] K.J. Friston, C.D. Frith, P.F. Liddle, and R.S.J. Frackowiak. Functional connectivity: The principal component analysis of large (PET) data sets. *Journal of Cerebral Blood Flow and Metabolism*, 13(1):5–14, 1993.

[28] K.J. Friston, C.D. Frith, R. Turner, and R.S.J. Frackowiak. Characterizing evoked hemodynamics with fMRI. *NeuroImage*, 2(2):157–165, June 1995.

[29] K.J. Friston, A.P. Holmes, J.-B. Poline, P.J. Grasby, S.C.R. Williams, R.S.J. Frackowiak, and R. Turner. Analysis of fMRI time-series revisited. *NeuroImage*, 2(1):45–53, March 1995.

[30] K.J. Friston, A.P. Holmes, K.J. Worsley, J-P. Poline, C.D. Frith, and R.S.J. Frackowiak. Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, 2(4):189–210, 1994.

[31] K.J. Friston, P. Jezzard, and R. Turner. Analysis of functional MRI time-series. *Human Brain Mapping*, 1(2):153–171, 1994.

[32] K.J. Friston, P.F. Liddle, and C.D. Frith. Dysfunctional frontotemporal integration in schizophrenia. *Neuropsychopharmacology*, 10:538S, 1994.

[33] K.J. Friston, S. Williams, R. Howard., R.S.J. Frackowiak, and R. Turner. Movement-related effects in fMRI time-series. *Magnetic Resonance in Medicine*, 35(3):346–355, 1996.

[34] K.J. Friston, K.J. Worsley, R.S.J. Frackowiak, J.C. Mazziotta, and A.C. Evans. Assessing significance of focal activations using their spatial extent. *Human Brain Mapping*, 1(3):210–220, 1994.

[35] G. Golub and C. Van Loan. *Matrix computations*. Johns Hopkins University Press, 1983.

[36] F. Gonzalez-Lima. Neural network interactions related to auditory learning analyzed with structural equation modeling. *Human Brain Mapping*, 2(1–2):23–44, 1994.

[37] F.E. Grubbs. Procedures for detecting outlying observations in samples. *Technometrics*, 11:1–21, 1969.

[38] J. A. Hartigan and M. A. Wong. A k-means clustering algorithm. *Applied Statistics*, 28:100–108, 1979.

[39] J.A. Hartigan. *Clustering Algorithms*. Wiley, New York, 1975.

[40] A.M. Hasofer. Upcrossings of random fields. In R.L. Tweedie, editor, *Proceedings of the Conference on Spatial Patterns and Processes, Supplement to Adv. Appl. Prob.*, volume 10, pages 14–21, 1978.

[41] A.P. Holmes, R.C. Blair, J.D.G. Watson, and I. Ford. Non-parametric analysis of statistic images from functional mapping experiments. *Journal of Cerebral Blood Flow and Metabolism*, 16(1):7–22, 1996.

[42] K.K. Kwong, J.W. Belliveau, D.A. Chesler, I.E. Goldberg, R.M. Weisskoff, B.P. Poncelet, D.N. Kennedy, B.E. Hoppel, M.S. Cohen, R. Turner, H. Cheng, T.J. Brady, and B.R Rosen. Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proceedings of the National Academy of Science*, 89:5675–5679, 1992.

[43] N. Lange. Statistical approaches to human brain mapping by functional magnetic resonance imaging. *Statistics in Medicine*, 15(4):389–428, 1996.

[44] N. Lange and S.L. Zeger. Non-linear Fourier analysis of magnetic resonance functional neuroimage time series. *Applied Statistics*, 46(1), 1997. In Press.

[45] C. Lawrence, J.L. Zhou, and A.L. Tits. User's guide for CFSQP version 2.3: A C code for solving (large scale) constrained nonlinear (minimax) optimization problems, generating iterates satisfying all inequality constraints. Technical Report Institute for Systems Research TR-94-16r1, University of Maryland, College Park, MD 20742, USA, 1994.

[46] A.T. Lee, G.H. Glover, and C.H. Meyer. Discrimination of large venous vessels in time-course spiral blood-oxygen-level-dependent magnetic-resonance functional neuroimaging. *Magnetic Resonance in Medicine*, 33(6):745–754, 1995.

[47] MATLAB. Mathworks inc. Sherborn MA, USA.

[48] G. McCarthy, M. Spicer, A. Adrignolo, M. Luby, J. Gore, and T. Allison. Brain activation associated with visual motion studied by functional magnetic resonance imaging in humans. *Human Brain Mapping*, 2:234–243, 1995.

[49] A.R. McIntosh and F. Gonzalez-Lima. Structural equation modeling and its application to network analysis in functional brain imaging. *Human Brain Mapping*, 2(1–2):2–22, 1994.

[50] R.S. Menon, S. Ogawa, S.-G. Kim, J.M. Ellermann, H. Merkle, D.W. Tank, and K. Ugurbil. Functional brain mapping using magnetic resonance imaging: Signal changes accompanying visual stimulation. *Investigative Radiology*, 27(suppl. 2):S47–S53, 1992.

[51] V.P. Nosko. The characteristics of excursions of gaussian homogeneous random fields above a high level. In *Proceedings of the USSR-Japan Symposium on Probability*, Harbarovsk, pages 216–222, Novosibiriak, 1969.

[52] V.P. Nosko. Local structure of Gaussian random fields in the vicinity of high level shines. *Sov. Math. Doklady*, 10:1481–1484, 1969.

[53] V.P. Nosko. On shines of gaussian random fields (in russian). *Vestn. Moscov. Univ. Ser. I. Mat. Meh.*, pages 18–22, 1970.

[54] S. Ogawa, D.W. Tank., R. Menon, J.M. Ellerman, S.G. Kim, H. Merkle, and K. Ugurbil. Intrinsic signal changes accompanying sensory stimulation: Functional brain mapping with magnetic resonance imaging. *Proceedings of the National Academy of Science*, 89:5951–5955, 1992.

[55] J.-B. Poline and B.M. Mazoyer. Analysis of individual positron emission tomography activation maps by detection of high signal-to-noise-ratio pixel clusters. *Journal of Cerebral Blood Flow and Metabolism*, 13(3):425–437, May 1993.

[56] M.I. Sereno, A.M. Dale, J.B. Reppas, K.K. Kwong, J.W. Belliveau, T.J. Brady, and R.B.H. Tootell. Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science*, 268:889–893, May 1995.

[57] D.O. Siegmund and K.J. Worsley. Testing for a signal with unknown location and scale in a stationary Gaussian random field. *Annals of Statistics*, 23(2):608–639, April 1995.

[58] G.W. Snedecor and W.G. Cochran. *Statistical Methods*. Iowa State University Press, 7th edition, 1982.

[59] SPM statistical parametric mapping software. K.J. Friston and others, 1994. Wellcome Department of Cognitive Neurology.

[60] G.S. Watson. Serial correlation in regression analysis. i. *Biometrika*, 42:327–341, 1955.

[61] J.B. Weaver. Efficient calculation of the principal components of imaging data. *Journal of Cerebral Blood Flow and Metabolism*, 15:892–894, 1995.

[62] P.H. Westfall and S.S. Young. *Resampling-based multiple testing*. Wiley, New York, 1993.

[63] R.P. Woods, S.R. Cherry, and J.C. Mazziotta. Rapid automated algorithm for aligning and reslicing PET images. *Journal of Computer Assisted Tomography*, 16(4):620–633, July/August 1992.

[64] K.J. Worsley. Local maxima and the expected Euler characteristic of excursion sets of χ^2, F, and t fields. *Advances in Applied Probability*, 26:13–42, 1994.

[65] K.J. Worsley. Estimating the number of peaks in a random field using the Hadwiger characteristic of excursion sets, with applications to medical images. *Annals of Statistics*, 23(2):640–669, April 1995.

[66] K.J. Worsley, A.C. Evans, S. Marrett, and P. Neelin. A three-dimensional statistical analysis for CBF activation studies in human brain. *Journal of Cerebral Blood Flow and Metabolism*, 12(6):900–918, November 1992.

[67] K.J. Worsley and K.J. Friston. Analysis of fMRI time-series revisited – again. *NeuroImage*, 2(3):173–181, September 1995.

[68] K.J. Worsley, J.-B. Poline, and K.J. Friston. Tests for distributed, non-focal brain activations. *NeuroImage*, 2(3):183–194, September 1995.

[69] J.H. Zar. *Biostatistical Analysis*. Prentice-Hall, 2nd edition, 1984.