Distribution Agreement

Signature:

Gordana Derado Date

Methods for Addressing Spatial Correlations in Functional Neuroimaging Data

By

Gordana Derado

Doctor of Philosophy

Biostatistics

F. DuBois Bowman, Ph.D.
Advisor

Ying Guo, Ph.D.
Committee Member

Clinton D. Kilts, Ph.D.
Committee Member

Amita K. Manatunga, Ph.D.
Committee Member

Lance A. Waller, Ph.D.
Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date

Methods for Addressing Spatial Correlations in Functional Neuroimaging Data

By

Gordana Derado
M.S., Emory University, 2009
M.S., Georgia State University, 2001
B.S., University of Zagreb, Croatia, 1992

Advisor: F. DuBois Bowman, Ph.D.

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Biostatistics
2010

Abstract

Methods for Addressing Spatial Correlations in Functional Neuroimaging Data

By

Gordana Derado

Neuroimaging studies yield massive data sets that pose challenges for statistical analyses due, in part, to the intricate anatomical and functional properties of neurons. Our main objective is to uncover aspects of the complex spatial relationships present in functional neuroimaging data and to develop statistical methods that either evaluate or leverage those correlations. We propose the following methods to achieve our research goal: (i) a novel statistical approach to model the complex spatio-temporal structure of neuroimaging data, (ii) a method to evaluate the level of connectivity within functionally defined neural processing networks and (iii) a novel prediction framework for neuroimaging data based on a hierarchical Bayesian spatial model.

To date, there has been limited research on simultaneously modeling spatial correlations between the neural activity in distinct brain locations and temporal correlations between repeated neural activity measurements. We propose a spatio-temporal, autoregressive model which simultaneously accounts for spatial dependencies between voxels within the same anatomical region and for temporal dependencies between a subject's estimates from multiple sessions. We illustrate the application of our method using fMRI data from a cocaine addiction study.

Data-driven statistical approaches, such as ICA and cluster analysis, help to identify neural processing networks exhibiting similar patterns of activity. These approaches, however, do not quantify or statistically test the strength of the within-network relatedness between voxels. We adapt Moran's I statistic for applicability to our neuroimaging analyses to measure the degree of functional autocorrelation within identified neural processing networks and to evaluate the statistical significance of the observed associations. We illustrate the use of our methodology with data from an fMRI resting-state study of unipolar depression and a PET study of working memory among individuals with schizophrenia.

Recently there has been growing interest in the use of neuroimaging data as a tool for classification and prediction. We propose a novel Bayesian hierarchical framework for predicting follow-up neural activity based on the baseline functional neuroimaging data. The proposed model is multivariate and captures the correlations between brain activity at different scanning sessions. We illustrate the use of our proposed methodology with PET data from a study of Alzheimer's disease.

Methods for Addressing Spatial Correlations in Functional Neuroimaging Data

By

Gordana Derado
M.S., Emory University, 2009
M.S., Georgia State University, 2001
B.S., University of Zagreb, Croatia, 1992

Advisor: F. DuBois Bowman, Ph.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Biostatistics
2010

Acknowledgement

There are many people whose help has been invaluable to finishing this dissertation.

First and foremost, I want to thank my advisor and committee chair, Professor DuBois Bowman, the best advisor I could have imagined to have. Professor Bowman has guided me throughout my graduate study, since my first year in the Department of Biostatistics and Bioinformatics at Emory. During these years, I have learned so much from him and am deeply indebted to him for many hours of insightful discussions, constant encouragement, and for his endless patience. The breath and depth of his scientific knowledge, and his ideas, enthusiasm, and support made my Ph.D. experience productive and stimulating. It has been an honor to be Professor Bowman's student.

My special thanks go to the members of my dissertation committee, Professor Ying Guo, Professor Amita Manatunga, and our department chair, Professor Lance Waller, for their guidance and helpful discussions. Their door has always been open for me when I needed advice or guidance, while their comments and suggestions have greatly improved the presentation of my dissertation.

I am extremely grateful to Professor Clint Kilts from the Department of Psychiatry and Behavioral Sciences at Emory, to whom I owe most of my knowledge about brain functioning and brain disorders. Professor Kilts has been a great advisor and collaborator - his help and guidance throughout my research in the area of neuroimaging has been invaluable. He is extremely knowledgeable, a well-known expert in his area, and I have been very fortunate to have had an opportunity to work with him and his research team. Professor Kilts is now a director of the Brain Imaging Research Center at the University of Arkansas.

I am also extremely grateful and fortunate to have had the opportunity to be enrolled in, and supported by, the neuroimaging track of the Biostatistics in Genetics, Immunology, and Neuroimaging (BGIN) training program at Emory. The program has given me a great opportunity to pursue my true interests during my dissertation research. It has also been a great experience (which I hope will continue) to be a member of the Center for Biomedical Imaging statistics (CBIS) at Emory. Our weekly meetings have been very constructive, inspirational and fun.

Many thanks to my colleagues and friends, and the department faculty and staff, for making my time at Emory University memorable.

Most importantly, none of this would have been possible without the love and patience of my family. My special gratitude is due to my parents Ana and Juraj and to my brother Ivan for their loving support throughout my education and my life. Last but not least, my most heartfelt gratitude goes to my husband Josip and my wonderful children Luka and Tony. Without their love, encouragement, patience, and belief in me it would have been impossible for me to complete this work.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Overview

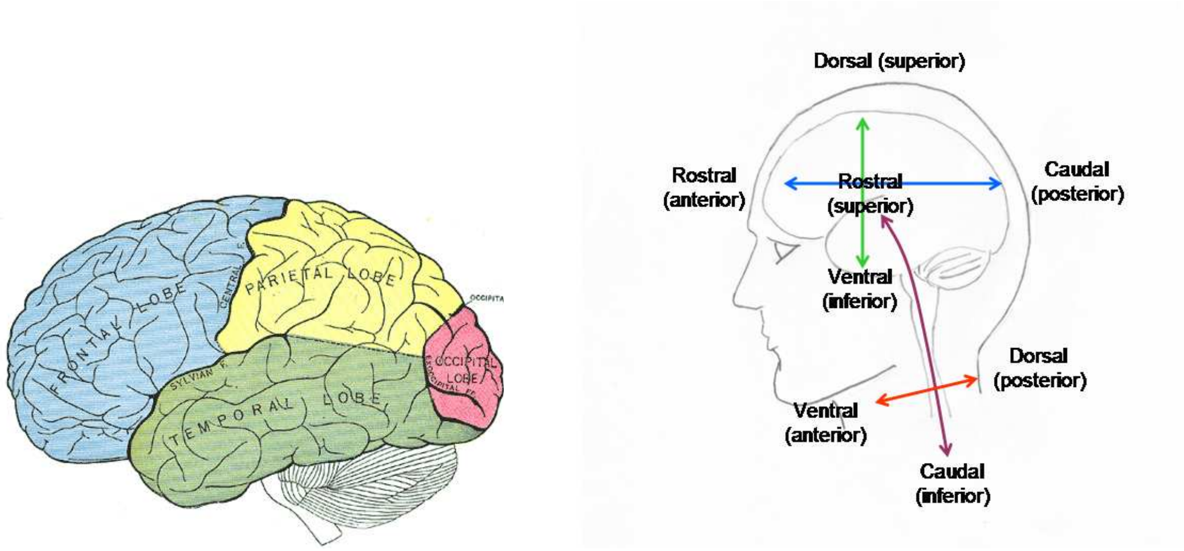
Advances in neuroimaging technology such as functional magnetic resonance imaging (fMRI), positron emission tomography (PET), and diffusion tensor imaging (DTI), have revolutionized our understanding of the human brain. The development of these neuroimaging techniques has given researchers an opportunity for in-depth investigations into both the anatomy and the function of the human brain. Neuroimaging plays an increasingly important role in the diagnosis, therapeutic monitoring, and scientific investigation of different neurologic and psychiatric disorders. Neuroimaging studies yield massive data sets that pose challenges for statistical analyses due, in part, to the intricate brain neurophysiology. The functional organization of the brain is based on information processing properties of specialized cells (neurons) and the complex and malleable modes of their interconnectivity. Our main objective was to study aspects of the complex spatial relationships present in functional neuroimaging data and to develop new statistical methods that provide accurate inference in the presence of associated patterns and correlations.

The dissertation is organized as follows: The remainder of Chapter 1 provides background info on the human brain, various neuroimaging technologies, and standard methods of statistical analyses and also outlines our proposed research objectives. Chapter 2 presents a novel statistical approach to model the complex spatio-temporal structure of neuroimaging data. Chapter 3 presents methods to evaluate the level of connectivity within functionally defined neural networks, and Chapter 4 presents a novel prediction framework for neuroimaging data.

1.2 An Introduction to Organization of the Human Brain

Basic Neuroanatomy. The brain is the center of the human nervous system and constantly receives, interprets, and directs sensory information and organizes motor and other responses throughout the body. The human brain has intricate structure and function and has been the target of scientific inquiry for centuries. The largest and uppermost section of the brain, operating at the highest functional level, is called the cerebrum. The cerebrum consists of two hemispheres that communicate with each other by a thick bundle of fibers called the corpus callosum. The right hemisphere controls the left side of the body and vice-versa. Each hemisphere is covered by a thick layer (1.5mm to 5mm) of gray matter called the cerebral cortex. The cerebral cortex is heavily folded, resulting in ridges called gyri and grooves called sulci. A cerebral hemisphere can be divided into four lobes: frontal, parietal, temporal, and occipital (Figure 1.1 (a)), each of which specializes in different functions. The cerebral cortex is connected to numerous subcortical structures such as the thalamus, the basal ganglia, hypothalamus, cerebellum, and the brainstem.

Neuroscientists use several terms to describe orientation, some of which are illustrated in Figure 1.1 (b). The following terms are used to describe direction in the brain: rostral/anterior means front or head end; caudal/posterior means hind



(a) (b)

Figure 1.1: (a) Principal fissures and lobes of the cerebrum viewed laterally, from *Anatomy of the Human Body*, the 1918 edition, available on-line at `www.bartleby.com/107` (Gray, 1918). (b) The axes of the central nervous system.

or tail end; dorsal means back or top side; ventral means belly or bottom side; lateral means away from the midline, medial means toward the midline. We typically view neuroimaging scans in three orthogonal planes referred to as axial (transverse or horizontal), coronal, and sagittal (see Figure 1.2).

The functional cellular unit of the central nervous system is a nerve cell or *neuron*. Brain tissue is divided into two types: gray matter and white matter. Grey matter contains neural cell bodies, in contrast to white matter, which mostly contains myelinated axons (myelin is a collection of lipid fats and proteins that sheaths the axons). The color difference arises mainly from the white color of myelin.

Brain Atlases and Coordinate Systems. When group studies using functional imaging data are performed, the individual brain images are usually transformed into a common coordinate space to accomodate the variation. The Talairach space and the Montreal Neurological Institute (MNI) space are the two most widely used atlas spaces in the neuroscience community. The Talairach coordinate system is based on a stereotaxic atlas of the human cerebral cortex published by Talairach and Tournoux (Talairach and Tournoux, 1988). It is related to anatomical landmarks and specifies locations relative to their distance from the midpoint of a brain white matter structure called the anterior commissure (AC) - a small but easy to spot region. Each location is described by three numbers which describe the distance in millimeters from the AC: x is the left/right dimension, y is the posterior/anterior dimension, and z is the ventral/dorsal dimension. In this atlas the axial slices are referred to by their z coordinate, coronal by their y coordinate, and sagittal by their x coordinate. The atlas is based on one brain, a brain of a 60-year-old French woman.

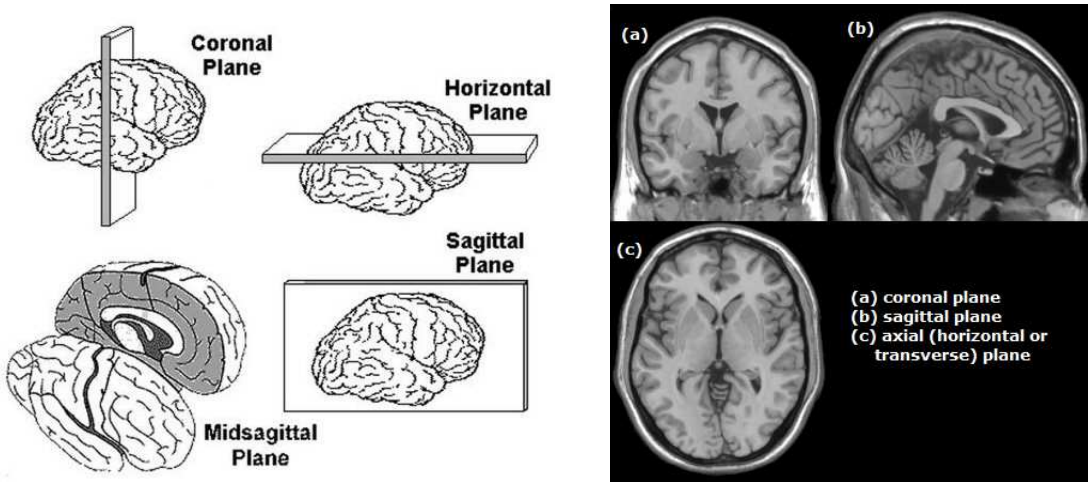


Figure 1.2: Planes of Brain Sections.

The MNI defined a new standard brain by using a large series of MRI scans of healthy normal controls (Evans et al., 1993). These atlases differ in shape and size. For example, the Talairach brain is considerably smaller than the average brain by up to 10 millimeters in each dimension. Software packages are available to approximately convert one coordinate system to the other.

1.3 Functional Neuroimaging

Neuroimaging is traditionally divided into *structural* and *functional* imaging. Structural imaging maps the brain anatomy and includes computed tomography (CT) and MRI. Functional techniques seek to examine the physiological properties of the brain, either at rest or during task-induced activation. A variety of methods allow mapping of human brain functioning. PET and fMRI measure localized changes in cerebral blood flow (referred to as activations) that accompany neuronal activity with relatively high spatial resolution ($\leq$ 1mm), but with temporal resolution (2-20 sec) limited by the much slower rate of brain regional blood flow and blood oxygenation. In contrast, techniques such as electroencephalography (EEG) and magnetoencephalography (MEG) map the underlying electrical activity of the brain cortex. These methods allow high temporal resolution of neural processes, but have poor spatial resolution (over 1 cm).

While each modality is interesting in its own right, in this dissertation we focus and introduce new statistical methods applied to fMRI and PET data. In the following two sections, we give a brief description of the main principles on which these two neuroimaging techniques are based. In this dissertation we concentrate on how they are applied to *human* brain neuroimaging.

1.3.1 Functional Magnetic Resonance Imaging (fMRI)

As opposed to PET, which requires the use of radioactive markers, fMRI is a relatively safe and non-invasive technique for generating maps of and studying brain activity. fMRI data consist of a 3-D sequence of individual magnetic resonance (MR) images - a “movie” of a subject’s brain activity. To understand the principles of fMRI, we need to understand how individual MR images are acquired. Below, we present a brief overview of the principles of MRI.

1.3.1.1 Basic Principles of MRI.

“MRI is based on ability of organic tissue, placed in a magnetic field, to absorb energy from radio waves, and then re-emit it, in proportion to mobile hydrogen ion concentration” (Higgins, 1996). The basis of this process is an interaction between radio waves and atomic nuclei called *nuclear magnetic resonance* (NMR). To obtain an MR image, a subject is placed in a field of a large electromagnet (generally from 1.5 to 4.0 Tesla). The magnetic field aligns the magnetization of hydrogen atoms in the brain. (Since much of the human brain is water, hydrogen nuclei are abundant, which makes them ideal candidates for brain imaging studies.) A hydrogen nucleus whose spin is oriented parallel to the applied magnetic field is said to be relaxed or in the low energy state, while a nucleus whose spin is oriented against the magnetic field is is said to be in an excited or high energy state. The scanner injects a pulse of radio frequency (RF) to excite the nuclei and raise them out of their low energy states. Upon removal of the RF pulse, the hydrogen nuclei return to their original aligned position (i.e., lower energy states), and in doing so, emit RF energy measured by the scanner. If the frequency of the input radio waves equals the natural resonance frequency (the Larmor frequency) of the element in the magnetic field, the tissue will absorb the radiation and when radio waves are turned off, it will emit a return signal of the same frequency, but lesser amplitude. If the frequency does not match the natural resonance of the element, the element will neither resonate nor send a signal. The emission of this RF signal is the MRI signal that is recorded in fMRI studies. Devices (coils) that can generate additional small local magnetic field gradients are used to slightly shift the Larmor frequency across brain areas and the scanner to recognize the location of the signal by its frequency.

By adding RF or gradient pulses and carefully choosing their timing, it is possible to highlight different characteristics of the imaged tissue. In general, MRI maps the water distribution in the brain. However, the useful contrast in MR images comes not only from spatial variation in the density of water, but also from differences in nuclear magnetic properties known as relaxation. They are characterized by distinct rates or *relaxation times*, used in MRI to distinguish between tissue types. Three relaxation times are of primary interest in MRI: T_1, T_2 and T_2^*. T_1 effects measure recovery of longitudinal magnetization (parallel to the main magnetic field), while T_2 refers to decay of transverse magnetization (perpendicular to the main magnetic field). T_1 time refers to interval where 63% of longitudinal magnetization is recovered, and T_2 time refers to the interval where only 37% of original transverse magnetization is present. T_2 decay results from inherent, random variations in the precession of individual protons. However, large-scale inhomogeneities in the magnetic field can also produce differences in precession frequency, resulting in signal reduction. When T_2 dephasing (which occurs as the protons gradually fall out of phase in their precession) is due to to one or more localisable sources, it is referred to as as T_2^*. So, T_2 has only one reason for dephasing and is intrinsic to tissue type (e.g. different T_2 of cerebral spinal fluid and fat), while T_2^* dephasing includes true T_2 as well as field inhomogeneity and tissue susceptibility. (The magnetic field inhomogeneity can result from the microscopic magnetic field gradients resulting from an increase in blood oxygenation. More detail will be given in the next section.) fMRI is a T_2^* image.

The raw data obtained from an MRI scanner are collected in the frequency domain. The inverse Fourier transform is then used to transfer the data into image space, where data analysis is performed.

1.3.1.2 Principles of fMRI

fMRI relies on hemodynamic correlates of neural activity, not the activity itself. It is most commonly performed using blood oxygenation level dependent (BOLD) contrast (Ogawa et al., 1990). The activation signal measured with BOLD fMRI is indirectly measuring changes in the concentration of paramagnetic deoxyhemoglobin

(hemoglobin without bound oxygen molecules) which arise from an increase in blood oxygen utilization in the vicinity of neuronal firing.

fMRI is possible because of two basic phenomena. The first is the fact that regional changes in brain activity are associated with local changes in blood flow and blood oxygenation in the brain (collectively known as *hemodynamics*). The second phenomenon is the ability of paramagnetic agents to produce contrast in the MR signal. (Paramagnetism is the ability of an otherwise nonmagnetic material to exhibit magnetic properties in the presence of magnetic field.)

The physiological basis of BOLD fMRI. When a subject performs a behavioral task, for example a repetitive finger-tapping task, neuronal networks involving multiple brain regions are activated. These collections of neurons require large amounts of energy as adenosine triphosphate (ATP) to sustain their metabolic activity. Since the brain does not store its own energy, increased blood flow is required to deliver the necessary glucose and oxygen (which is bound to hemoglobin) to meet this metabolic demand. There is a disproportionate increase in regional cerebral blood flow (CBF) relative to the cerebral oxygen metabolic rate in response to neuronal activation (e.g., 30% vs. 5%). This results in a decreased oxygen extraction fraction (more oxygen supply than consumption) and hence local 'hyperoxygenation' – a surfeit of oxygenated blood.

Biophysics of BOLD signal. The reduced oxygen extraction leads to an increase in the ratio of oxygenated hemoglobin (oxy-Hb) to deoxygenated Hb (deoxy-Hb) in the region of neural activation. Deoxy-Hb is slightly paramagnetic relative to brain tissue while oxy-Hb is dimagnetic. Because of its paramagnetic properties, deoxy-Hb has the effect of suppressing the MR signal, while oxy-Hb does not. Presence of deoxy-Hb causes microscopic magnetic field inhomogeneities (gradients) which leads to distractive interference from signal within the tissue element called *voxel* (a unit of 3D volume). This process tends to shorten the T_2^* relaxation time. As oxygen extraction falls with increased local blood flow in a region of greater neuronal activity, the T_2^* becomes longer and the MRI signal intensity increases relative to the baseline state. The image intensity that varies with deoxy-Hb content has been termed Blood Oxygenation Level Dependent (BOLD) and was suggested for potential use in functional study of the brain by Ogawa et al. (1990).

Basic parameters of image acquisition. There are many imaging design parameters (fixed at the beginning of the study) which control the spatial and temporal resolution of the data and image quality in general. The most important intrinsic parameters, which affect the signal measured at each voxel, are the repetition time (TR) and echo time (TE). TR is the time, in milliseconds (ms), between successive applications of RF pulses to a particular volume of tissue. It is impossible to measure the signal immediately after the RF is applied, due to hardware limitations. The short waiting time (also measured in ms's) during which the peak signal is obtained is called TE.

The most common imaging sequence used in fMRI is the fast method of echo planar imaging (EPI), which allows collection of whole brain data in a few seconds or less.

1.3.2 Positron Emission Tomography (PET) Imaging

PET is a nuclear imaging technique for mapping brain function, or other molecular processes in the body. It measures emissions related to positron decay from radioactively labeled chemicals that have been injected into the bloodstream and generates images of the distribution of the radio chemicals throughout the brain and body.

PET is also not an exact measure of brain function. It depends upon certain assumption about what happens when an area of the brain becomes active: 1) cerebral metabolism requires glucose metabolism, which requires oxygen from blood flow, i.e. in parts of the brain that are more active there is more blood flow; 2) cerebral blood flow varies locally with corresponding local variations in neural activity.

In PET, images of the brain's activity are generated by tracking the brain's use of a radioactively tagged compound, such as glucose, oxygen or a particular drug. A machine called a cyclotron (used to accelerate charged particles to high energies) is used to tag the specific substance with small amounts of radioactivity. Before scanning, the labeled compound (called a radiotracer) is injected into the subject's bloodstream (or inhaled). When the tracer gets into the bloodstream, it goes to areas of the brain that use it, while the subject is engaged in some type of mental activity. So, oxygen and glucose accumulate in brain areas that are metabolically active. When the radioactive material decays, it gives off a neutron and a positron. When a positron hits an electron, both are destroyed and two gamma rays are released approximately 180° opposed to each other. Sensors in the PET scanner detect and record the brain area where the gamma rays are emitted. By reconstructing the sites of the positron-electron collisions, the location of active regions can be imaged. A computer uses the data gathered by the sensors to construct multicolored 2- or 3-dimensional images that show where the compound is in the brain. Different colors or degrees of brightness on a PET image represent different levels of tissue or organ function. This method hence provides a functional view of the brain.

PET has very high biochemical sensitivity and selectivity (which allow probing the neurochemical processes at the molecular level), but its temporal and spatial resolution are inferior to that of fMRI. The mean free path of the positrons in brain tissue limits the spatial resolution of PET scanning to about 4 mm. However, PET images can be superimposed on subject's MRI images, providing detailed information about specific brain areas involved in a wide variety of functions. Spatial resolution of PET data depends upon several other factors: the size and type of the crystal used in the scanner scintillator to detect the gamma radiation emission, the energy of the positron emitted etc. Temporal resolution depends mainly upon half-life of the isotope. Safety regulations require to wait 5 half-lives between injecting the radioactive tracers.

Experimental Design. In a typical PET experiment, a brain scan is taken during a control task (e.g., resting with eyes closed). The control task is compared to brain scan taken while the subject is exposed to the experimental treatment or performing the experimental task. To determine the brain activity that can be attributed to the experimental condition, the difference between the PET scans is calculated.

1.4 Analysis of Functional Neuroimaging Data

Typical objectives in the analysis of functional neuroimaging data are sought through the following: (i) activation studies attempt to localize particular brain regions that are source(s) of task-related neural processing, (ii) connectivity studies seek to identify what brain areas show similar patterns of activity over time, yielding distributed networks of brain function, and (iii) prediction studies try to use neuroimaging scans to predict future behavioral, psychological, or cognitive outcomes, or to classify subgroups of individuals (e.g. for diagnostic purposes).

Data from a functional neuroimaging study consist of a series of 3-D images, typically obtained while the subject performs a certain cognitive, behavioral or emotional task, or while at rest. In an fMRI study, typically hundreds of such 3-D images are obtained, taken approximately 2-4 seconds apart. In a PET study, the number of scans is significantly smaller (the maximum number of scans per subject is limited due to the total isotope dose allowed), and the brain images are obtained much less frequently. Each 3-D image comprises of a large number ($>$ 100,000 in an fMRI study; usually smaller for PET) of voxels. In addition, the experiment may be repeated for the same subject, as well as for multiple subjects. Because of the neurophysiology of the network organization of the brain, spatial correlations are very likely. Also, temporal correlations (both within and between scanning sessions) are present due to repeated scanning.

To reduce artifacts caused by the scanner and subject movement, and to map individual brains to a standard brain atlas (when performing group analysis and making population inference), the functional scans undergo several preprocessing steps before statistical analysis.

1.4.1 Preprocessing Pipeline

The main steps involved in fMRI preprocessing are: slice timing correction (shifting each voxel's time course so that we can assume they were measured simultaneously), realignment (to correct for motion; usually a rigid body transformation using 6 variable parameters is used), coregistration of structural and functional images (to map the results obtained from functional data to a high resolution structural image for presentation purposes), normalization (important in group analysis, attempts to register each subject's brain to a standard template brain (see Section 1.2), it is commonly done using non-linear transformations to match local features), and spatial smoothing (typically done by convolving the functional images with a Gaussian kernel). Spatial smoothing is done for several reasons: it may improve inter-subject registration, it ensures that the assumptions of the random field theory (often used for multiple testing correction) are satisfied, and to increase signal to noise ratio. However, spatial smoothing causes a loss of acquired data and may introduce artificial spatial correlations between nearby voxels. Preprocessing steps and the order in which they are performed are important since they affect both the spatial and temporal correlation structure of the data.

We refer to Woods et al. (1998)a and Woods et al. (1998)b for the details on the preprocessing steps for PET data.

1.4.2 Statistical Modeling for Activation Studies

The general objective of activation studies is to identify brain locations that are involved in the neural processing associated with tasks that subjects perform while in the scanner and possibly to compare these neural processing traits between tasks or between subgroups of individuals. Commonly, a two-stage statistical modeling procedure is used, consisting of single subject analysis at stage 1 and a group level model at stage 2. We present the modeling procedure in context of fMRI data, and PET analyses proceed using a similar framework. To set the notation, assume that the data consist of a brain volume of V voxels (indexed by $v = 1 \ldots, V$), repeatedly measured at T time points (indexed by $t = 1, \ldots, T$). In addition, assume that the experiment is repeated for K subjects (indexed by $i = 1, \ldots, K$).

1.4.2.1 Single-Subject Analysis

The first stage of the model characterizes distributed brain activity associated with various experimental conditions, separately for each individual. A single subject model for fMRI activation in a single voxel can be expressed as

$$y_{iv}(t) = \sum_{l=1}^{L} x_{ivl}(t) B_{ivl} + \sum_{m=1}^{M} h_{ivm}(t) \nu_{ivm} + \varepsilon_{iv}(t), \qquad (1.1)$$

for $v = 1, \ldots, V$, $i = 1, \ldots, K$ and $t = 1, \ldots, T$. Here $y_{iv}(t)$ represents the measured fMRI signal at time t for voxel v and subject i; x_{vil} represents the task-related BOLD response (the signal of interest) corresponding to l-th condition at time t; and h_{ivm} represents the contribution of the nuisance parameters (such as high-pass filtering parameters - to remove the drift component typically present due to scanner instabilities, periodic fluctuations due to heart rate and respiration, and head motion) at time t. The terms B_{ivl} and ν_{ivm} represent the unknown amplitude of x_{ivl} and h_{ivm} and $\varepsilon_{iv}(t)$ represents the noise process. Typically, the relationship between stimuli and BOLD response is modeled using a linear time invariant system where the stimulus acts as the input and the hemodynamic response function (HRF) as the impulse response function. In a linear system framework the signal $x(t)$ is modeled as the convolution of a stimulus function $s(t)$ and the hemodynamic response $f(t)$, i.e. $x(t) = (s * f)(t) = \int_0^\infty s(u)f(t-u)\,\mathrm{d}u$.

In summary, (1.1) can, for most fMRI experiments, be represented in matrix form as

$$\mathbf{y}_i(v) = \mathbf{X}_{iv}\mathbf{B}_i(v) + \mathbf{H}_{iv}\boldsymbol{\nu}_i(v) + \boldsymbol{\varepsilon}_i(v), \tag{1.2}$$

where $\boldsymbol{\varepsilon}_i(v)$ is usually assumed to follow an AR(2) process. Here we arranged the serial BOLD fMRI responses for subject i in the vector $\mathbf{y}_i(v) = (y_{iv1}, \ldots, y_{ivT})'$ (at a given voxel location). $\mathbf{X}_{iv}$ is a $T \times L$ matrix with columns corresponding to predicted BOLD response for each condition l, and $\mathbf{B}_i(v) = (B_{iv1}, \ldots, \beta_{ivL})'$. If, for example, the nuisance matrix represents a drift, modeled using a pth order polynomial function, then $\boldsymbol{\nu}_i(v) = (\nu_{vi1}, \ldots, \nu_{vip})'$. $\mathbf{H}_{iv}$ is a $T \times p$ matrix with columns corresponding to the polynomial functions.

The *general linear model* (GLM)(Worsley and Friston, 1995; Friston et al., 2002) is the most popular way to analyze the single subject fMRI data. The GLM assumes that the HRF is known apriori. If, in addition, the stimulus function $v_l(t)$ is known and equivalent to experimental paradigm (for example: a vector of zero's and one's, representing times when the signal is 'off' and 'on', respectively), then (1.2) becomes a standard multiple regression model with known signal components and unknown amplitudes. HRF is usually modeled using a canonical HRF, typically a gamma function or a difference between two gamma functions. The BOLD response is summarized in the design matrix $\mathbf{X}_{iv}$, containing separate columns for each of the l predictors. The parameters $\mathbf{B}_i(v)$ and $\boldsymbol{\nu}_i(v)$ represent subject specific effects corresponding to $\mathbf{X}_{iv}$ and $\mathbf{H}_{iv}$, respectively. The vector $\boldsymbol{\varepsilon}_i(v)$ denotes random errors representing characteristics of the measurement process that are unaccounted for by the model, and we assume that the errors follow a zero-mean multivariate normal distribution $\boldsymbol{\varepsilon}_i(v) \sim \mathrm{N}(0, \sigma_i^2(v)\mathbf{I})$, where $\sigma_i^2(v)$ is the error variance at voxel v.

After fitting the model, a contrast vector $\mathbf{c}$ can be defined to test for an effect $\theta_i = \mathbf{c}'\mathbf{B}_i$. For example, we can test for the increased brain activity during the 'on' signal versus during the 'off' signal. Hypothesis testing can be performed in a usual way: individual model parameters can be tested using a t-test and subset of parameters using a partial F-test.

1.4.2.2 Multi-Subject (Group Level) Analysis

Typically, multi-subject analysis of fMRI data from an activation study is performed using hierarchical models, which provide a framework for performing mixed-effects analysis.

The stage 2 model combines the subject-specific effects $\mathbf{B}_i(v)$ from equation (1.2) to estimate the associated group level or population parameters. The second level model can then be written as

$$\mathbf{B}_i(v) = \mathbf{W}_{iv}\boldsymbol{\beta}(v) + \mathbf{e}_i(v) \tag{1.3}$$

where $\mathbf{W}_{iv}$ is the second-stage design matrix that often consists of subject-specific characteristics and $\boldsymbol{\beta}(v)$ contains the group level parameters representing the effects related to different sessions or the effects associated with different subpopulations (e.g. treatment groups). Often, the interest is in making inferences about the differences between sessions/groups. A common approach is to obtain θ_i from the single subject analysis and then model it in terms of the corresponding group level parameter.

1.4.3 Data-driven Descriptive Analysis Methods

Data-driven approaches to the analysis of functional neuroimaging data employ pattern-based exploratory methods to "characterize" the nature of the signal present in the data. A lot of recent work regarding the analysis of brain imaging data has focused on examining functional connectivity of the human brain. Functional connectivity has been defined as the temporal correlation between spatially remote neurophysiologic events. For example, Independent Component Analysis (ICA) and cluster analysis attempt to divide the brain into different functional networks involved in performing certain tasks or characterizing a "default mode" of the resting-state brain.

1.4.3.1 Clustering

Cluster analysis uses PET or fMRI data to help identify dissociable networks or clusters, each consisting of voxels that show correlated patterns of measured brain activity. These clustering solutions do not define the underlying neuroanatomical connections, but instead identify functional associations between voxels. Ideally, intracluster voxels should exhibit high functional (or spatial) autocorrelation, validating that the neural responses within clusters are functionally related.

Clustering procedures classify the V voxels in an image into G groups, with each cluster consisting of V_g voxels, where $g = 1, \ldots, G$ and $V = \sum_{g=1}^{G} V_g$. Most methods classify voxels by measuring the distance (dissimilarity) between the activity time courses for every pair of voxels and combining voxels with small distances. Two typical dissimilarity measures are Euclidean and Mahalonobis distances, which are easy to compute and naturally interpretable.

Numerous investigators have proposed the use of clustering methods for neuroimaging data including Balslev et al. (2002), Baumgartner et al. (2000), Bowman and Patel (2004), and Cordes et al. (2002). Clustering algorithms generally fall into one of two categories, namely hierarchical clustering algorithms or partitioning algorithms. Hierarchical clustering performs a series of nested merges or divisions, ranging from a solution with one large cluster containing all voxels to a solution in which every voxel represents a separate cluster. Among the available hierarchical clustering procedures are Ward's, beta-flexible, centroid linkage, variable linkage, median linkage, single linkage, and complete linkage algorithms (Bowman and Patel, 2004). Partitioning algorithms, such as K-means and fuzzy K-means, specify the number of clusters at the outset and sequentially reallocate voxels to clusters until obtaining a final clustering solution (Fadili et al., 2001). Descriptions of particular clustering algorithms are available in Rencher (2002), and Bowman et al. (2004) highlight several of these algorithms for neuroimaging applications.

1.4.3.2 Independent Component Analysis (ICA)

Among the data-driven techniques, ICA has become a very popular method, successfully employed to decompose functional neuroimaging data into sets of spatial maps and associated time-courses. ICA is an application of a blind source separation that attempts to decompose the data set into components that are as statistically independent from each other as possible (Common, 1994; Herault and Jutten, 1986). It is possible to pursue either the temporal or the spatial independence of the target components. One method of performing ICA minimizes the mutual information between components (Bell and Sejnowski, 1995; McKeown et al., 1997, 1998; Calhoun and Pekar, 2000). The fixed-point algorithm (Hyvärinen, 1999) pursues the same goal of minimizing the mutual information, but uses the concept of normalized differential entropy or negentropy (Common, 1994). One interpretation of negentropy is as a measure of non-normality, so maximizing the negentropy finds directions of maximal-non-normality in the data.

Classical ICA. Let $\mathbf{Y}$ denote a $S \times V$ matrix of the observed voxel time courses, where S is the number of scans, and V is the number of voxels included in analysis; $\mathbf{C}$ is a

$N \times V$ random matrix whose rows C_i are to be filled with the unknown realizations of the N spatial components (images, reshaped in 1-D vectors), and $\mathbf{A}$ is a $S \times N$ mixing matrix, whose columns contain the associated time-courses of the N components. The spatial ICA problem for fMRI time series can be formulated as an estimation of the following model for the data:

$$\mathbf{Y} = \mathbf{AC} \tag{1.4}$$

There are no assumptions about the mixing matrix $\mathbf{A}$, while the constraint on the spatial processes C_i is that they are (ideally) mutually statistically independent. The amount of statistical dependence within a fixed number of spatial components can be quantified by means of their mutual information, an important function in information theory (Common, 1994). Thus, the ICA decomposition of $\mathbf{Y}$ can be defined (up to a multiplicative constant and to the sign) as an invertible transformation: $\mathbf{C} = \mathbf{WY}$, where the matrix $\mathbf{W}$ (so-called unmixing matrix) is determined such that the mutual information of the target components C_i is minimized (i.e., such that C_i's are "as independent as possible"). Matrix $\mathbf{A}$ is the pseudoinverse of $\mathbf{W}$. In McKeown et al. (1998), the sources are estimated by iteratively optimizing the unmixing matrix $\mathbf{W}$ so that $\mathbf{C} = \mathbf{WY}$ contains mutually independent rows, using the 'infomax' algorithm.

Probabilistic ICA. Beckmann and Smith (2004) propose a probabilistic ICA (PICA) model aimed at solving the problem of overfitting in classical ICA applied to fMRI data, by including a Gaussian noise term in the classical ICA decomposition. The PICA model, which extends model (3.2), is formulated as a generative linear latent variables model. The model is "characterized by assuming that the S-variate vector of observations is generated from a set of q statistically independent non-Gaussian sources via a linear instantaneous mixing process, corrupted by additive noise $\boldsymbol{\eta}(t)$"(Beckmann and Smith, 2004):

$$\mathbf{y}_v = \mathbf{A}\mathbf{c}_v + \boldsymbol{\mu} + \boldsymbol{\eta}_v, \quad \forall v \in \{1, \ldots, V\}$$

where $\mathbf{y}_v$ denotes the S-dimensional column vector of individual measurements at voxel location v, $\mathbf{c}_v$ denotes the q-dimensional column vector of non-Gaussian source signals contained in the data and $\boldsymbol{\eta}_v$ denotes Gaussian noise $\boldsymbol{\eta}_v \sim N(0, \sigma^2 \boldsymbol{\Sigma}_v)$. The number of source processes is assumed to be smaller than the number of observations in time ($q < S$). The model is closely related to Factor Analysis (Bartholomew, 1987). The individual component maps are obtained using a modified fixed-point iteration scheme to optimize for non-Gaussian source estimates via maximizing the negentropy. The maps are then separately transformed to spatial Z-scores. These Z-score maps depend on the amount of variability explained by the entire decomposition at each voxel location relative to the residual noise. Next, Gaussian/Gamma mixture models are fitted to the individual Z-maps in order to infer voxel locations that are significantly modulated by the associated time-course. For more details on PICA, we refer the reader to Beckmann and Smith (2004).

Group ICA. Several methods for extending the single-session probabilistic ICA model to higher dimensions and making group inferences have been proposed (Calhoun et al., 2001; Beckmann and Smith, 2005; Guo and Pagnoni, 2008). In Beckmann and Smith (2005), the single-session PICA has been extended to higher dimensions allowing for a model-free analysis of multi-subject or multi-session fMRI data. For group ICA, either tensorial ICA (where the data is decomposed into spatial maps, time courses and subject/session modes), or a simpler temporal concatenation approach (a single 2-D ICA is run on the concatenated data matrix obtained by stacking all 2-D data matrices of every single data set on top of each other) are possible. The latter approach is recommended to use when the interest is in finding a common spatial pattern, but we cannot assume that the associated temporal response is consistent between subjects, such as in the analysis of data acquired without stimulation (e.g. resting-state). The technique is derived from the Parallel Factor Analysis (Harshman, 1970; Harshman and Lundy, 1994). [1]

1.4.4 Prediction and Classification

In recent years there has been growing interest in the use of neuroimaging data as a tool for classification (e.g. classification of mental and behavioral disorders) and prediction (e.g., predicting an early onset of Alzheimer's disease). Typical prediction objectives in functional neuroimaging studies include: (i) predicting human experiences and behaviors from brain imaging data, (ii) predicting brain activity patterns, e.g. following a treatment intervention, using pre-treatment data and patient characteristics, and (iii) predicting treatment response and relapse. Many attempts have been made to address (i), e.g. Pugh et al. (1996), Hoeft et al. (2007), Giessing et al. (2007). There have also been attempts to address (ii). For example, Guo et al. (2008) introduced a Bayesian hierarchical model for PET and fMRI data to forecast brain activity in schizophrenic patients following a specific treatment. Functional neuroimaging has recently begun to show promise as a clinical tool in the prediction of treatment response. For example, Evans et al. (2006) demonstated functional neuroimaging to be promising as a clinical tool in the prediction of treatment response in psychiatric disorders such as major depression and obsessive-compulsive disorder. The findings they reviewed suggest that treatment outcome may be predicted by patterns of pre-treatment brain activity in psychiatric patients. They conclude, however, that the actual clinical utility of such tests remains to be shown.

Popular classification methods in functional neuroimaging include Fisher's Linear (and Quadratic) Discriminant Analysis (LDA) (Fisher, 1936), logistic regression (LR), Support-Vector Machines (SVM) (Vapnik, 1995), Neural Network classifiers (NN) (Hertz et al., 1991), Partial Least Squares (PLS) analysis (Wold, 1966; McIntosh

---

[1] The research by Beckmann and Smith described above has been implemented as MELODIC (Multivariate Exploratory Linear Optimized Decomposition into Independent Components - a stand alone C++ program). It is freely available as part of FSL (FMRIB's Software Library).

et al., 1996), and Bayesian classifiers (e.g. naive Bayes classifier).

Pattern recognition methods (which include SVM, NN, LDA) have been successfully applied in functional neuroimaging studies (LaConte et al., 2005; Pessoa et al., 2007). These methods can be used to infer cognitive states (so-called brain decoding). Using such approaches, it is possible to predict the mental state of a subject or a stimulus class by analyzing the spatial distribution of neural responses. The SVM is one of the most popular methods used to carry out this type of analysis. SVMs arise from the Statistical Learning Theory of Vapnik (Vapnik, 1995) and their formulation was motivated to deal with small sample sizes and high dimensional inputs, which match the situation involved for temporally predictive modeling of fMRI data. Cox and Savoy (2003) used statistical pattern recognition algorithms including LDA and SVM to separate brain activation maps from an fMRI experiment in which participants viewed images of objects (baskets, birds, butterflies, chairs, cows, etc.). The objects belonged to various categories, both of similar and differing forms. A notable finding of their study was that it was possible, with small amounts of fMRI data, to determine what the participant was viewing at levels well above chance.

Hanson et al. (2004) implemented a variety of feed-forward, neural network architectures with both linear and nonlinear decision rules. PLS, introduced to neuroimaging community by McIntosh et al. (1996), is a useful method for forming prediction equations when there are a large number of explanatory variables, particularly when the random error variance is large. Giessing et al. (2007) applied PLS analysis to study behavioral effects of nicotine in nonsmokers and found that neural data under placebo can be used to predict individual behavioral effects of nicotine. Mitchell et al. (2004) present case studies in which they have successfully trained three classifiers: a Gaussian Naive Bayes classifier, k-nearest neighbor, and linear SVMs, to distinguish cognitive states such as whether the human subject is looking at a picture or a sentence or whether the word the subject is viewing is a word describing food, people, buildings, etc.

Methodology for statistical prediction based on functional neuroimaging data represents an important area for future research, and preliminary work in this area provides a promising outlook for the potential usefulness of functional neuroimaging data in a clinical setting. For the functional neuroimaging modality, however, to be useful in defining diagnostic categories or monitoring treatment success, it needs to be shown that the technology has the ability to define clinically relevant information on a single-subject basis.

1.5 Motivating Examples

1.5.1 An fMRI Study on Inhibitory Control in Cocaine Addicts

A common characteristic of drug addiction is an impairment in the ability to exert inhibitory control over drug-related behaviors, in spite of adverse consequences. This inhibitory control deficit can be elicited using response inhibition tasks outside of the drug seeking context. The inhibitory control task, referred to as the *STOP-signal* task, was designed to evaluate the ability to cancel a prepotent motor response (Aron and Poldrack, 2006). The sample included 12 cocaine addicts who received an extensive outpatient behavioral treatment program and 15 healthy controls. The cocaine addicts were scanned while performing an inhibitory control task in two separate sessions, before and after treatment, and similarly control subjects had baseline and follow-up scans.

The aim of our analysis is to identify brain locations where there are inhibitory control-related increases in brain activity following treatment for cocaine addiction. In our analysis we take into account that there are both spatial correlations present between neighboring voxels in the brain, as well as temporal correlations between repeated brain scanning sessions.

1.5.2 A PET Study on Working Memory in Schizophrenia Patients

Schizophrenia is a devastating mental illness that may cause severe disturbances in cognition, social behavior, and emotion. One important brain function that often goes awry in patients with schizophrenia is working memory, characterized by the ability to manipulate information in short-term memory. Our data are from a parametric study of working memory in an executive function task in individuals with schizophrenia. There were 16 subjects in the study, each having a total of 8 PET scans obtained using a blood flow tracer. The aim is to examine the functional connectivity related to working memory tasks and to evaluate the validity of the neural processing networks identified by functional clustering.

1.5.3 An fMRI Resting-state Study of Depression

This data consist of fMRI resting-state scans of seven women with a history of major depression. Functional data were acquired while the subjects were in the scanner with their eyes open and while looking at a visual fixation cross. A group independent component analysis (ICA) was applied to investigate the resting-state functional connectivity of a group of depressed patients. There is substantial interest in determining associations between the brain activity characteristics of different regions while subjects are in a resting state. For example, a set of regions has been consistently identified in these investigations and has been labeled as the *default mode network* (DMN). Our goal is to introduce a measure for the degree of functional autocorrelation within neural processing networks identified by ICA and to evaluate the statistical significance of the observed associations.

1.5.4 A PET Study of Alzheimer's Disease

Alzheimer's Disease Neuroimaging Initiative (ADNI)
(`http://www.loni.ucla.edu/ADNI/`) is a large national project with the goal to develop biomarkers of Alzheimer's Disease (AD) in elderly subjects, to define the rate of progress of mild cognitive impairment and Alzheimer's disease, and to provide a large database which will improve design of treatment trials. We analyzed PET data from 73 healthy controls and 73 Alzheimer's disease patients, obtained from the ADNI database. Our goal is to predict subject's follow-up (6 month) brain activity, based on the baseline activity.

1.6 Proposed Research

Functional neuroimaging studies yield large data sets characterized by complex dependence structures driven by highly sophisticated neurophysiology and neuroanatomy, and aspects of the experimental designs. These complex dependence structures pose analytical challenges for statistical modeling. To date, there is a paucity of methods that incorporate spatial considerations. Most analyses are performed in a univariate setting, where each voxel is modeled independently of the others ("massive univariate approach"). Ignoring correlations affects the precision of estimates of model parameters and consequently may lead to inaccurate statistical tests. Spatio-temporal modeling may mitigate these shortcomings by incorporating more physiologically plausible assumptions and by borrowing strength across related measures of neural activity.

1.6.1 Simultaneous Spatio-temporal Modeling of fMRI data

fMRI data sets contain temporal correlations from repeated scanning (within and between scanning sessions) and complex spatial correlations. Typical analyses investigating task-related changes in measured brain activity use a two-stage procedure in which the first stage involves subject-specific models and the second-stage specifies group (or population) level parameters. Customarily, the first-level accounts for temporal correlations between the serial scans acquired during one scanning session. Despite accounting for these correlations, fMRI studies often include multiple sessions and temporal dependencies may persist between the corresponding estimates of mean neural activity. Further, spatial correlations between brain activity measurements in different locations are often unaccounted for in statistical modeling and estimation. Bowman (2005) proposed an extended two-stage model for the estimation and testing of localized activity in which the second stage accounts for spatial dependencies between voxels within the same neural processing cluster (defined by a data-driven cluster analysis). This model, however, did not account for repeated measures type associations between the multiple experimental effects for each subject. We propose a two-stage, spatio-temporal, autoregressive model which simultaneously accounts for spatial dependencies between voxels within the same anatomical region and for temporal dependencies between a subject's estimates from multiple sessions. We develop an algorithm that leverages the special structure of our covariance model, enabling relatively fast and efficient estimation.

1.6.2 Functional Autocorrelation within Neural Processing Networks

Data-driven statistical approaches, such as cluster analysis or independent component analysis, applied to in vivo functional neuroimaging data help to identify neural processing networks that exhibit similar task-related or resting-state patterns of activity. Ideally, the measured brain activity for voxels within such networks should exhibit high autocorrelation. An important limitation of the existing approaches is that they do not quantify or statistically test the strength or nature of the within network relatedness between voxels. To extend the results given by such data-driven analyses, we propose the use of Moran's I statistic to measure the degree of functional autocorrelation within identified neural processing networks and to evaluate the statistical significance of the observed associations. We adapt the conventional definition of Moran's I, for applicability to neuroimaging analyses, by defining the global autocorrelation index using network-based neighborhoods. Also, we compute network-specific contributions to the overall autocorrelation.

1.6.3 A Novel Spatial Prediction Model

Neuroimaging has been used in clinical practice for over 30 years, but it is still perceived as rarely offering clinicians much help in direct patient management. One of the important directions in neuroimaging is increasing its clinical applicability in terms of diagnostic purposes and to help predict future patients' outcomes.

We propose a novel Bayesian hierarchical framework for predicting follow-up neural activity based on the baseline functional neuroimaging data that attempts to overcome some shortcomings of the presently used modeling methods (such as small sample sizes, heterogeneity, as well as recording of brain activity under resting-state conditions) by borrowing strength from the spatial correlations present in the data. The spatial correlations are incorporated in the model in two ways: the short-range correlations between neighboring voxels are incorporated through a multivariate conditional autoregressive (CAR) prior of the spatial parameters, while the long-term correlations between anatomical brain regions are incorporated through the covariance matrix of the random effect parameters.

Chapter 2

Modeling the spatial and temporal dependence in fMRI data

2.1 Introduction

Functional magnetic resonance imaging (fMRI) studies yield large data sets that contain temporal correlations from repeated scanning (within and between scanning sessions) and complex spatial correlations. Ignoring correlations affects the precision of estimates of model parameters and consequently may lead to inaccurate statistical tests. Spatio-temporal modeling may mitigate these shortcomings by incorporating more physiologically plausible assumptions and by borrowing strength across related measures of neural activity. One session of a typical fMRI neuroactivation study acquires 3-D scans every 2-3 seconds while the subject performs different experimental tasks. fMRI studies may also involve multiple sessions (e.g corresponding to pre- and post-treatment periods). Typical analyses investigating task-related changes in measured brain activity use a two-stage procedure in which the first stage involves subject-specific and voxel-specific models relating neural processing to experimental tasks. The second stage specifies voxel-specific models for group (or population) level parameters. The stage I analysis accounts for temporal correlations between serial scans within one session using variants of autoregressive models (Bullmore et al., 1996; Purdon et al., 2001; Friston et al., 2002). Despite addressing scan-to-scan correlations within a session, the stage I estimates for multiple sessions may also exhibit correlations, typically not accounted for at the second stage. Our focus is on modeling these latter correlations.

Some attempts have been made at modeling the repeated measures correlations and spatial correlations in fMRI data, though typically not in the same statistical model. Worsley et al. (2002) introduce a random effects analysis for combining sessions and the method removes trends in effects over time. Bowman and Kilts (2003) fit repeated measures covariance structures within a linear model to address between-session correlations in positron emission tomography (PET) neuroimaging data. Neither of these two approaches addresses the spatial correlations between localized brain activity measurements, and the method by Bowman and Kilts (2003) involves long computations for PET data, making it less suitable for higher dimensional fMRI data. Bowman (2005) presents a two-stage model in which the second stage accounts for spatial dependencies within brain regions or networks, and Bowman et al. (2008) give a more flexible Bayesian model to capture correlations both within and between brain regions. Neither of these approaches, however, accounts for temporal or repeated measures associations between the multiple experimental effects for a subject. There have been a few attempts to model both temporal and spatial correlations simultaneously, but these approaches are generally hampered by costly computation. The computations generally involve inverting large covariance matrices, and separable models (over space and time) are often considered (Benali et al., 1997; Hartvig, 2002). Gössl et al. (2001) use a Bayesian approach to fit a semi-parametric spatial and temporal model for fMRI data. Their approach requires extensive computations, limiting its applicability in practice. Woolrich et al. (2004)

propose a Bayesian framework to model the noise via a non-separable space-time, simultaneous autoregressive model. This approach is very time consuming for the fMRI datasets, taking roughly 6 hours for processing a single slice of 3-D fMRI data.

We propose a two-stage model that accounts for both spatial and temporal correlations in fMRI data, and our model leads to fast parameter estimation. In the second stage, we construct a simultaneous autoregressive model to capture spatio-temporal correlations between the multiple (session) effects at a given location and between pairs of voxels within defined anatomical regions. We use maximum likelihood (ML) methods to estimate parameters from our spatio-temporal model. We overcome computational challenges involved with estimation by deriving an algorithm that simplifies the calculations of inverses and determinants of large matrices, leading to fast estimation of the model parameters. Our spatio-temporal model provides a unified framework for both voxel-level and region-level inferences. Using our proposed model, we analyze fMRI data from a study of inhibitory control in cocaine addicts to evaluate the effects of behavioral therapy on neural processing related to inhibitory control. To further delineate benefits of our spatio-temporal model, we conduct two simulation studies: one to validate the accuracy of our estimation methods and the other to evaluate the relative efficiency of our proposed spatio-temporal model compared to the general linear model (GLM (not to be confused with *generalized linear models* as in McCullagh and Nelder (1989))).

2.2 Experimental Data

To illustrate the use of our proposed model and to give motivation for its development, we apply it to data from an fMRI study evaluating the impact of cocaine addiction and treatment-related abstinence on neural responses to motor inhibition tasks. The sample included 12 cocaine addicts enrolled in an intensive outpatient behavioral treatment program and 15 healthy controls, matched by age, sex (all males), race, handedness, education and early life adversity. All the subjects were scanned while performing an inhibitory control task used to model a characteristic hallmark of drug addiction: an impairment in the ability to inhibit behaviors (e.g. drug-seeking behaviors). The cocaine addicts were scanned in two separate sessions, before and after treatment, and similarly control subjects had baseline and follow-up scans.

Experimental conditions: The inhibitory control task, referred to as the *STOP-signal* task, was designed to evaluate the ability to cancel a prepotent motor response (Aron and Poldrack, 2006). In this task, subjects are presented with visual *GO-stimuli*, consisting of uppercase alphabetical letters appearing on the screen for 0.5 seconds with an inter-stimulus interval of 2.3 seconds. They are instructed to respond to this signal by pressing a button as quickly as possible. An auditory STOP-signal (also lasting 0.5 seconds) appears randomly in 16% of the trials. The occurrence of the STOP-signal following a GO-stimulus is an indicator to refrain from pushing the button (the prepotent response). Hence, a successful performance requires the inhibition of a prepotent behavior. The aim of our analysis is to identify brain locations where there are inhibitory control-related increases in brain activity following treatment for cocaine addiction.

2.3 Methods

Data from a single scan in our study are represented as a $53 \times 63 \times 46$ 3-D rectangular lattice, comprised of a large number of voxels, indexed by a mapping of coordinates (x, y, z) to $v = 1, \dots, V$ ($V > 150,000$ in our study). Each voxel contains intensity information corresponding to measures of localized brain activity (BOLD fMRI responses). Since the number of intracranial voxels is too large to estimate a global correlation matrix including all voxel pairs, we consider an approach that partitions the voxels into mutually exclusive subregions based on known neuroanatomical divisions. Specifically, we define anatomical regions using a parcellation of the brain based on Brodmann areas (BA) (Brodmann, 1909a) and additionally separate BA's in the left and right hemispheres. To provide broader coverage, we also add several subcortical regions (left and right caudate, thalamus and putamen) to our parcellation, which were obtained from the Automated Anatomical Labeling (AAL) map based on (Tzourio-Mazoyer et al., 2002). Other such parcellations exist, including the Freesurfer software (Fischl et al., 2002, 2004) and LONI (Laboratory of Neuro Imaging, UCLA) Probabilistic Brain Atlas (Shattuck et al., 2008). Building our model based on well-established anatomical parcellations, rather than data-driven functional clusters aids interpretability.

2.3.1 Statistical Model

Following the conventional two-stage modeling approach for fMRI data, we fit a GLM at the first stage for each individual's vector of serial BOLD responses, separately for each voxel. The GLM regresses each voxel's BOLD responses for subject i ($i = 1, \ldots, K$) on within-subject design variables (e.g. session or stimulus indicators) and on covariates that are not of intrinsic interest, such as high-pass filtering variables. The model accounts for short-range scan-to-scan correlations within a session using a first-order autoregressive process with white noise (Purdon et al., 2001; Friston et al., 2002). The regression coefficient, $B_{igs}^{(p)}(v)$, for voxel v represents a summary measure of an individual's (mean) neural activity associated with session or stimulus p. We add a subscript g to denote the neuroanatomic region to which voxel v belongs ($g = 1, \ldots, G$), with the region consisting of V_g voxels in total.

At the second stage, we propose a spatio-temporal (ST) autoregressive model to capture temporal correlations between the multiple sessions and spatial correlations between pairs of intra-regional voxels. We express the spatio-temporal model as follows:

$$\begin{aligned} B_{igs}^{(p)}(v) &= \mathbf{x}_i'\boldsymbol{\beta}_{gs}^{(p)}(v) + \rho_{gs}\,\frac{1}{V_g - 1}\sum_{v^*\in\mathcal{N}_v}\Big(B_{igs}^{(p)}(v^*) - \mathbf{x}_i'\boldsymbol{\beta}_{gs}^{(p)}(v^*)\Big) \\ &\quad + \xi_{gs}\,\frac{1}{q-1}\sum_{p^*\neq p}\Big(B_{igs}^{(p^*)}(v) - \mathbf{x}_i'\boldsymbol{\beta}_{gs}^{(p^*)}(v)\Big) + e_{igs}(v)\,, \end{aligned} \tag{2.1}$$

where $\mathbf{x}_i$ is a vector of between-subject design variables, $\boldsymbol{\beta}_{gs}^{(p)}(v)$ are the group-level parameters, $e_{igs}(v) \sim N(0, \phi_{gs}^2)$ (ϕ_{gs}^2 is the between-subject variation), ρ_{gs} is a spatial dependence parameter, ξ_{gs} reflects the temporal dependence between sessions, $\mathcal{N}_v$ is the set of all voxels within the same neuroanatomic region as v, and s denotes a subject group (e.g., patients, controls). When $\rho_{gs} = \xi_{gs} = 0$, model (2.1) resembles a typical stage II GLM from a random effects analysis of brain imaging data, with constant variances within each region.

2.3.2 Parametric Covariance Model

Our model involves three parameters defining the variance-covariance structure: the spatial (ρ) and temporal (ξ) dependence parameters and the variance parameter (ϕ^2). In this section, we present details about the assumed covariance/correlation model. Let $\mathbf{B}_{igs} = \Big(\mathbf{B}_{igs}^{(1)'}, \ldots, \mathbf{B}_{igs}^{(q)'}\Big)'$, where each $\mathbf{B}_{igs}^{(p)}$ contains individualized parameters for all voxels in region g, and similarly $\boldsymbol{\beta}_{gs} = \Big(\boldsymbol{\beta}_{gs}^{(1)'}, \ldots, \boldsymbol{\beta}_{gs}^{(q)'}\Big)'$. Model (2.1) implies that $\mathbf{B}_{igs}$ has the following multivariate normal distribution $\mathbf{B}_{igs} \sim \mathrm{MVN}\Big(\boldsymbol{\beta}_{gs}, \boldsymbol{\Omega}_{gs}\Big)$ (see Appendix 5.1.1), where $\boldsymbol{\Omega}_{gs} = \phi_{gs}^2\boldsymbol{\Psi}_{gs}\boldsymbol{\Psi}_{gs}$, and $\boldsymbol{\Psi}_{gs}$ is a $qV_g \times qV_g$ matrix with

$$\boldsymbol{\Psi}_{gs}^{-1} = \begin{bmatrix} \mathbf{I}_{V_g} - \rho_{gs}\mathbf{W}_g & -\frac{\xi_{gs}}{q-1}\mathbf{I}_{V_g} & \cdots & -\frac{\xi_{gs}}{q-1}\mathbf{I}_{V_g} \\ -\frac{\xi_{gs}}{q-1}\mathbf{I}_{V_g} & \mathbf{I}_{V_g} - \rho_{gs}\mathbf{W}_g & \cdots & -\frac{\xi_{gs}}{q-1}\mathbf{I}_{V_g} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{\xi_{gs}}{q-1}\mathbf{I}_{V_g} & -\frac{\xi_{gs}}{q-1}\mathbf{I}_{V_g} & \cdots & \mathbf{I}_{V_g} - \rho_{gs}\mathbf{W}_g \end{bmatrix}_{qV_g\times qV_g} \tag{2.2}$$

($\mathbf{I}_{V_g}$ and $\mathbf{J}_{V_g}$ denote identity and unity matrices of size V_g). In (2.2), $\mathbf{W}_g = [1/(V_g - 1)](\mathbf{J}_{V_g} - \mathbf{I}_{V_g})$ specifies the neighborhood structure consisting of all voxels that fall in the same anatomical region. The blocks along the main diagonal of $\boldsymbol{\Psi}_{gs}^{-1}$ allow for spatial correlations between measures of task-related brain activity for voxels in region g, i.e. between elements of $\mathbf{B}_{igs}^{(p)}$. The off-diagonal blocks of $\boldsymbol{\Psi}_{gs}^{-1}$ capture correlations between the summary measures of brain activity associated with various scanning sessions. Our model assumes an exchangeable covariance structure between voxels in the same neuroanatomic region. Although the complexity of human brain function is likely to render departures from this assumption, the exchangeable structure provides an improvement over the often used independence assumption and seems reasonable for statistical modeling purposes based on descriptive empirical results (not shown here; see Appendix 5.1.2).

Calculating the matrix $\boldsymbol{\Psi}_{gs}$. Calculating $\boldsymbol{\Psi}_{gs}$ involves inverting a matrix of size $qV_g \times qV_g$, which can become unwieldy for large regions, and this calculation is performed iteratively during estimation of the model. We derive an algorithm to facilitate calculations of inverses and determinants of our large highly structured covariance matrices, thereby enabling estimation of our model for fMRI applications. By recursive calculation on q, the number of blocks in $\boldsymbol{\Psi}_{gs}^{-1}$, we represent $\boldsymbol{\Psi}_{gs}$ as follows

$$\boldsymbol{\Psi}_{gs} = \begin{bmatrix} d\mathbf{I}_{V_g} + f\mathbf{J}_{V_g} & u\mathbf{I}_{V_g} + z\mathbf{J}_{V_g} & \cdots & u\mathbf{I}_{V_g} + z\mathbf{J}_{V_g} \\ u\mathbf{I}_{V_g} + z\mathbf{J}_{V_g} & d\mathbf{I}_{V_g} + f\mathbf{J}_{V_g} & \cdots & u\mathbf{I}_{V_g} + z\mathbf{J}_{V_g} \\ \vdots & \vdots & \vdots & \ddots \\ u\mathbf{I}_{V_g} + z\mathbf{J}_{V_g} & u\mathbf{I}_{V_g} + z\mathbf{J}_{V_g} & \cdots & d\mathbf{I}_{V_g} + f\mathbf{J}_{V_g} \end{bmatrix}_{qV_g \times qV_g} . \tag{2.3}$$

Hence, $\boldsymbol{\Psi}_{gs}$ is determined by four functions: $d(\rho_{gs}, \xi_{gs}, V_g, q)$, $f(\rho_{gs}, \xi_{gs}, V_g, q)$, $u(\rho_{gs}, \xi_{gs}, V_g, q)$, and $z(\rho_{gs}, \xi_{gs}, V_g, q)$. The explicit formulas for d, f, u and z are quite lengthy and are given in the Appendix 5.1.3. We employ (2.3) to circumvent issues of long computing times and excessive memory consumption.

Estimated spatial and temporal parameters are not directly interpretable as spatial and temporal correlations. The spatial and temporal correlations, denoted η_ρ and η_ξ, respectively, require calculation of the covariance matrix $\mathbf{\Omega}_{gs}$, i.e. of $\mathbf{\Psi}_{gs}\boldsymbol{\Psi}_{gs}$. Leveraging the special structure of $\mathbf{\Psi}_{gs}$, we obtain $\mathbf{\Psi}_{gs}\mathbf{\Psi}_{gs}$ as a partitioned matrix, with equal diagonal blocks $A\mathbf{I}_{V_g} + B\mathbf{J}_{V_g}$, and off-diagonal blocks $C\mathbf{I}_{V_g} + D\mathbf{J}_{V_g}$, where

$$\begin{aligned} A &= d^2 + (q-1)u^2 \qquad & B &= 2df + V_g f^2 + (q-1)(2uz + V_g z^2) \\ C &= 2du + (q-2)u^2 \qquad & D &= 2(fu + dz + V_g fz) + (q-2)(2uz + V_g z^2)\,. \end{aligned} \tag{2.4}$$

The expressions for the spatial correlation, η_ρ, and the temporal correlation, η_ξ, are then given by $\eta_\rho = B/(A+B)$ and $\eta_\xi = (C+D)/(A+B)$.

2.3.3 Estimation

We perform estimation of the second stage spatio-temporal model using ML methods. The log-likelihood function for model (2.1) is given by

$$\begin{aligned} \mathcal{L}(\boldsymbol{\phi}, \boldsymbol{\rho}, \boldsymbol{\xi}, \boldsymbol{\beta} | \mathbf{B}) \propto \sum_{s=1}^{2} \sum_{i \in K_s} \sum_{g=1}^{G} \Bigg\{ & -\tfrac{qV_g}{2} \ln(\phi_{gs}^2) + \ln\big(\text{abs}|\mathbf{\Psi}_{gs}^{-1}|\big) \\ & -\tfrac{1}{2\phi_{gs}^2} (\mathbf{B}_{igs} - \mathbf{X}_i \beta_{gs})' \mathbf{\Psi}_{gs}^{-1} \mathbf{\Psi}_{gs}^{-1} (\mathbf{B}_{igs} - \mathbf{X}_i \beta_{gs}) \Bigg\}\,, \end{aligned} \tag{2.5}$$

where K_s denotes the number of subjects in subgroup s (cocaine addicts and controls). The ML estimator of the mean parameter vector $\boldsymbol{\beta}_{gs}$ is given by $\hat{\boldsymbol{\beta}}_{gs} = \frac{1}{K_s} \sum_{i=1}^{K_s} \boldsymbol{B}_{igs},\ g = 1, \ldots, G$. $\hat{\boldsymbol{\beta}}_{gs}$ is unbiased (as is $\hat{\boldsymbol{\beta}}_{gs}$ from a GLM analysis) and does not depend on the covariance parameters. Estimation of the covariance parameters, therefore, proceeds using the partially maximized likelihood function $\mathcal{L}(\boldsymbol{\phi}, \boldsymbol{\rho}, \boldsymbol{\xi}, \hat{\boldsymbol{\beta}} | \mathbf{B})$. We use the Fisher scoring algorithm to estimate the covariance parameters $\boldsymbol{\rho}_s = (\rho_{1s}, \ldots, \rho_{Gs})'$, $\boldsymbol{\xi}_s = (\xi_{1s}, \ldots, \xi_{Gs})'$, and $\boldsymbol{\phi}_g = (\phi_{1s}, \ldots, \phi_{Gs})'$.

Note on the parameter space. To apply our model, the matrix $\mathbf{\Psi}_{gs}^{-1}$ must be nonsingular, or equivalently $\mathbf{\Omega}_{gs}$ must be positive definite. In the context of our fMRI data, with $V_g > 20$, for all g, and $q = 2$, the resulting explicit parameter constraints are as follows:

$$\begin{aligned} \rho_{gs} + (V_g - 1)(1 + \xi_{gs}) &\neq 0 & \qquad \rho_{gs} + (V_g - 1)(1 - \xi_{gs}) &\neq 0 \\ \rho_{gs} - \xi_{gs} &\neq 1 & \qquad \rho_{gs} + \xi_{gs} &\neq 1 \, . \end{aligned}$$

We monitor these boundary constraints during our iterative estimation procedure.

Estimation involves iterative calculations of the determinant and the inverse of a $qV_g \times qV_g$ matrix (up to roughly 5,800×5,800 in our data application). Employing the simplifications discussed in Section 2.3.2 enables fast and efficient computations. Appendix 5.1.4 contains details regarding the score functions necessary for estimation.

2.3.4 Inferences

Inferences that are commonly sought in functional neuroimaging studies target the identification of differences in neural processing between experimental tasks, sessions (e.g. treatment periods), or groups. One may pursue these inferences at an extremely localized (voxel) level or at a more spatially coarse regional level. In either case, estimation and inferences in this setting produce maps of distributed brain activity and the corresponding thresholded maps showing statistically significant (or highly probable, in a Bayesian context) differences. We seek inferences about linear functions of $\boldsymbol{\beta}_{gs}^{(p)}$, where for the inhibitory control study in cocaine addicts, $p = 1$ (baseline or pre-treatment) or $p = 2$ (follow-up or post-treatment) and $s = 1$ (cocaine addicts) or $s = 2$ (controls).

Voxel-level inferences. Following estimation using our spatio-temporal model, we obtain t-statistic images from the voxel-specific contrast estimates, then threshold the t-statistics to determine voxels exhibiting statistically significant changes (or differences) in brain activity. To adjust for typical changes observed with repeated scanning sessions in the inhibitory control study, we identify voxels for which the changes in brain activity following treatment are larger in cocaine addicts than the corresponding changes between follow-up and baseline activity in control subjects. Specifically, we estimate and test hypotheses about $\boldsymbol{\theta}_g^* = \mathbf{C}_g^* \boldsymbol{\beta}_g$, where $\mathbf{C}_g^* = [-\mathbf{I}_{V_g}\ \mathbf{I}_{V_g}\ \mathbf{I}_{V_g}\ \ -\mathbf{I}_{V_g}]_{(V_g \times 4V_g)}$ and $\boldsymbol{\beta}_g = [\boldsymbol{\beta}_{g1}^{(1)'}, \boldsymbol{\beta}_{g1}^{(2)'}, \boldsymbol{\beta}_{g2}^{(1)'}, \boldsymbol{\beta}_{g2}^{(2)'}]'_{(4V_g \times 1)}$. Each element of $\boldsymbol{\theta}_g^*$ represents a voxel-specific parameter. We construct Wald-type statistics for hypothesis testing, with $\mathrm{Var}(\hat{\boldsymbol{\theta}}_g^*) = \mathbf{C}_g^* \,\mathrm{Var}(\hat{\boldsymbol{\beta}}_g)\, \mathbf{C}_g^{*'}$ and $df = 2(n_p + n_c - 2)$, where n_p is the number of addicts, and n_c is the number of controls. We apply the stringent threshold of $\alpha = 0.005$, which is one approach that has a strong precedent in the neuroimaging literature, but other thresholding approaches, such as false discovery rate (Benjamini and Hochberg, 1995) and random field theory (Friston et al., 1995) are also available.

Region-level inferences. Our spatio-temporal model also enables analyses targeting an entire anatomical brain region corresponding to the underlying anatomical parcellation. For inferences in region g, define $\theta_g = \mathbf{C}_g \boldsymbol{\beta}_g$, with $\mathbf{C}_g = \frac{1}{V_g}[-\mathbf{1}'_{V_g}\ \mathbf{1}'_{V_g}\ \mathbf{1}'_{V_g}\ \ -\mathbf{1}'_{V_g}]_{(1 \times 4V_g)}$, and construct the t-statistic using $\mathrm{Var}(\hat{\theta}_g) = \mathbf{C}_g \,\mathrm{Var}(\hat{\boldsymbol{\beta}}_g)\, \mathbf{C}'_g$ and degrees of freedom $df_r = 2V_g(n_p + n_c - 2)$. Combining both sessions for subgroups s, i.e., $\boldsymbol{\beta}_{gs} = [\boldsymbol{\beta}_{gs}^{(1)'}, \boldsymbol{\beta}_{gs}^{(2)'}]'$, the variance of $\hat{\boldsymbol{\beta}}_{gs}$ is given by $\mathrm{Var}(\hat{\boldsymbol{\beta}}_{gs}) = \frac{1}{n_p+n_c} \phi_{gs}^2 \boldsymbol{\Psi}_{gs} \boldsymbol{\Psi}_{gs}$. A notable advantage of modeling spatial correlations using our model is that the resulting regional level analyses account for spatial dependencies between intra-regional voxel pairs, rather than implicitly assuming independence.

Statistical significance of spatial and temporal correlations. It is often informative to examine the magnitudes of spatial and temporal correlations as discussed in Section 3.2. Given the complexity of the estimators of these quantities, it is difficult to derive analytical expressions for the variances of these estimators. If one has inherent interest in testing hypotheses about these parameters, then standard bootstrapping methods are applicable. However, this approach may be computationally demanding. Given that testing hypotheses about correlations was not of chief interest in the inhibitory control study, we present estimates of correlations mainly for descriptive purposes. We estimate standard errors of the correlations using 30 bootstrap samples drawn (with replacement) from the subjects in our data, separately for controls and patients. We apply our model and estimate the spatial and temporal correlations for each sample, and then calculate the bootstrap standard errors as the standard deviations of the 30 estimated values for each parameter, which should provide reasonable estimates of variability (Efron and Tibshirani, 1998).

2.4 Application to Inhibitory Control in Cocaine Addicts Study

We used our spatio-temporal model to analyze fMRI data from the inhibitory control study in cocaine addicts. Spatial and temporal parameters were estimated separately for cocaine addicts and controls. The aim of our analysis was to evaluate whether the inhibitory control-related changes in measured brain activity following treatment for the addicts were larger than the corresponding between-session changes in the control subjects.

All of the image pre-processing (slice timing, realigning, normalizing, and smoothing [6mm (2 voxels) full width at half maximum (FWHM) Gaussian kernel], as well as the first level, single subject analyses, were carried out in SPM5[1]. To account for inter-subject neuroanatomic differences that may persist after spatial normalization, we chose to apply very focal spatial smoothing, rather than the more spatially expan-

---

[1] SPM5 is is a MATLAB software package implementing Statistical Parametric Mapping for neuroimaging data available for download from the Wellcome Trust Centre for Neuroimaging web page `http://www.fil.ion.ucl.ac.uk/spm/software/spm5/`.

sive smoothing often applied in practice. Our smoothing should only have a small effect on subsequent spatial modeling, since most of the Brodmann areas (within which we model spatial correlations) are quite large, relative to the small size of the smoothing kernel. The second level analysis, implementing our proposed spatio-temporal model, was carried out using MATLAB. Our analysis includes 38 BA's in each hemisphere, after excluding those with fewer than 20 voxels, and we supplement the BA's with six subcortical regions, resulting in a parcellation consisting of 82 brain regions. We perform exploratory analyses, computing crude estimates of spatial and temporal correlations (Appendix 5.1.2), which provide support for the existence of correlations from both sources as well as for the exchangeability assumption in our covariance model.

2.4.1 Voxel-level inferences

Figure 2.1(a) shows voxels that reveal statistically significant increases in inhibitory control-related brain activity following treatment at $\alpha = 0.005$. The axial slices correspond to 5 mm, 14 mm, 44 mm, and 65 mm above the anterior-posterior commissural plane. The significant areas include the right frontopolar cortex (BA 10), left middle temporal gyrus (BA 21), and retrosubicular area (BA 48) in the 5 mm slice; left and right thalamus and the right inferior frontal cortex (BA 45) at 14 mm; left visual association cortex V3 (BA 19) and somatosensory association cortex (BA 7) and the left angular gyrus (BA 39) at 44 mm; and the right pre-supplementary motor area (pre-SMA, BA 6) at 65 mm. Other regions (not shown) exhibiting significant treatment-related increases in brain activity include the right inferior prefrontal gyrus (BA 47), left and right fusiforum gyrus (BA 37), and left primary auditory cortex (BA 41).

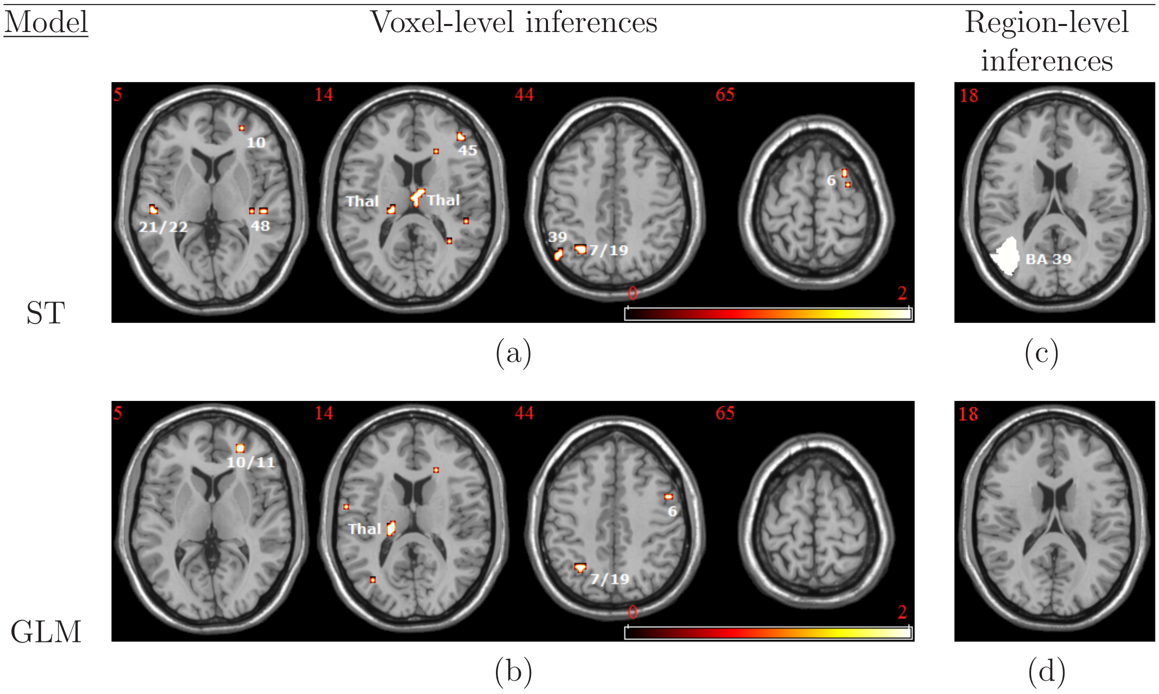
Figure 2.1(b) shows results from the corresponding GLM–based analysis commonly used in the neuroimaging literature. Many of the areas showing statistical significance agree with the results from our approach (e.g. Brodmann areas 6, 7, 10, and 19 emerged in both analyses), but there are some differences. For example, our method identifies increased treatment-related brain activity in BA 48 and BA's 21 and 22 (5 mm), which do not emerge from the GLM analysis. Also, both methods identify areas in the thalamus, but our method detects a large right thalamic activation not revealed by the GLM. Similarly, both methods identify voxels in BA's 6 (pre-SMA) and 39, but our model yields more spatially extensive activations. Overall, our ST model produces more statistically significant voxels than the GLM. However, neither method produces significant voxels when applying more conservative family-wise error or false discovery rate multiple testing procedures. The differences between the two models become more apparent at larger significance levels (e.g. $\alpha = 0.01$, not shown), suggesting that with a larger sample size and stronger activation signal, greater differences are likely to emerge. We provide interpretive remarks about our results in Section 2.4.4.

2.4.2 Region-level inferences

Our region-level analysis of treatment-emergent changes in neural processing, using methods described in Section 2.3.4, revealed only one region (left BA 39, angular gyrus) achieving statistical significance at $\alpha = 0.05$, uncorrected (see Figure 2.1(c)). The comparative region-level GLM analysis (that implicitly assumes spatial and temporal independence) did not yield any statistically significant differences. Neither method reveals significant differences at more stringent thresholds such as $\alpha = 0.005$ (uncorrected).

2.4.3 Spatial and temporal correlations

Figure 2.2 shows the model-based estimates of spatial and temporal correlations for both the cocaine addicts and control subjects in one axial slice of the brain. The



(a) (c)

(b) (d)

Figure 2.1: Voxel-level and region-level inferences for two models: the proposed spatio-temporal (ST) model (top) and the GLM (bottom). (a): voxels that achieve statistical significance in the ST model analysis at $\alpha = 0.005$ in favor of a one-sided alternative hypothesis of increased activity following treatment for cocaine addiction. Significant increases occur in the right frontopolar area (BA 10), left middle and superior temporal gyri (BA's 21/22), right retrosubicular area (BA 48), right inferior frontal cortex (BA 45), left and right thalamus (Thal), left somatosensory and visual association cortices (BA's 7 and 19), left angular gyrus (BA 39), and right BA 6 (pre-SMA); (b): voxel-level results from a GLM analysis; (c): region, left BA 39 (angular gyrus), that achieves significance in our ST model analysis at $\alpha = 0.05$; (d): corresponding region-level results from the GLM. Slice labels denote the distance, in mm, above the anterior-posterior commissural plane.

color scale indicates the strength of correlations, with brighter shades indicating correlations of larger magnitude. For temporal correlations, positive values are shown in one color (top colorbar) and negative values are shown in a separate color (bottom colorbar). For clearer distinction, we add stripes to regions with negative correlations. All of the correlations in Figure 2.2 would have the value zero in a GLM analysis, signaling the need to account for them in our model.

Table 2.1 gives the model-based estimates, along with bootstrap standard errors (in parentheses), of both temporal and spatial correlations for select regions in the

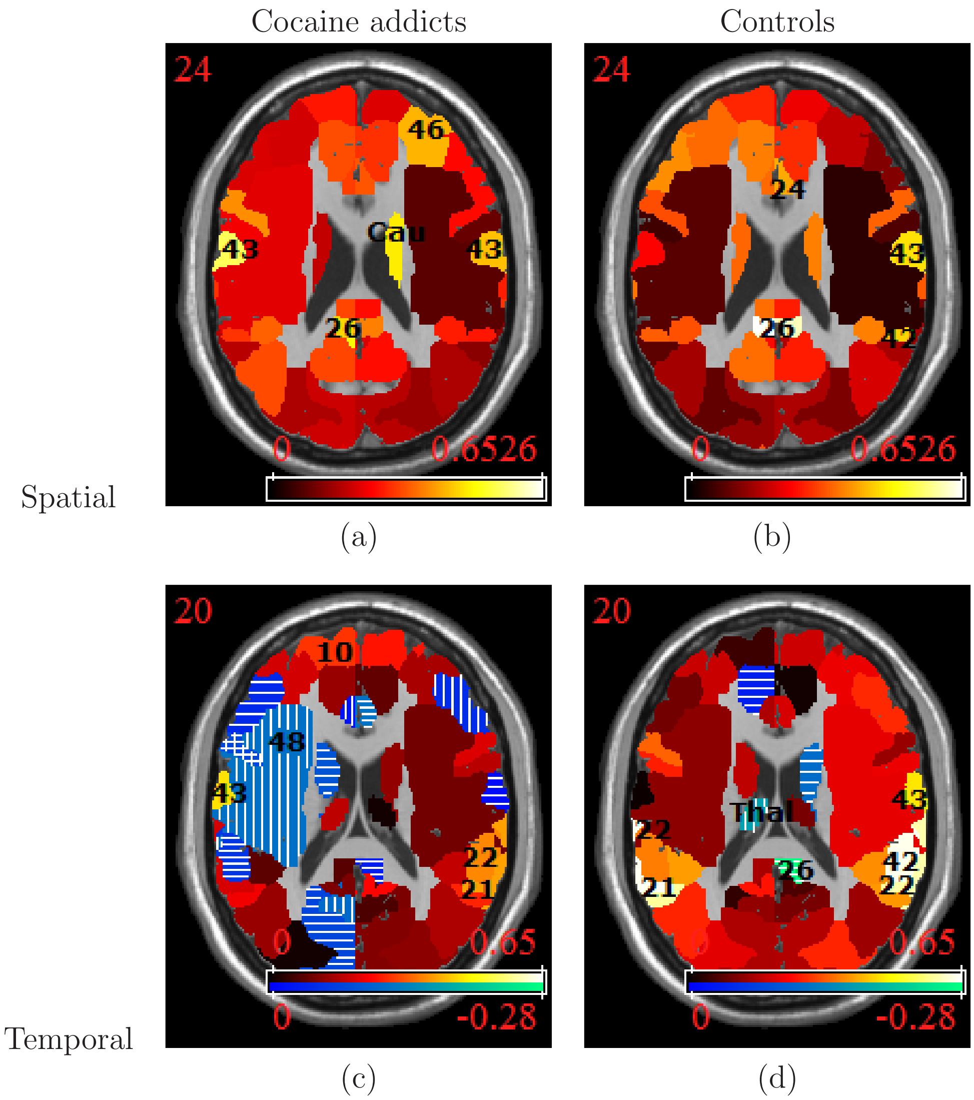


Figure 2.2: Model-based estimates of the spatial and temporal correlations for the inhibitory control study (a) addicts: spatial correlations, (b) controls: spatial correlations, (c) addicts: temporal correlations, (d) controls: temporal correlations. For (a) and (b), the darker to lighter intensities in the colorbar represent lower to higher correlations. For (c) and (d), the top colorbar represents positive correlations, again with lighter intensities reflecting higher correlations, and the bottom colorbar represents negative correlations (stripes added to maps) with brighter intensities indicating stronger negative correlations. The temporal correlations for control subjects range from -0.28 to 0.64 and are in general larger than those for the cocaine addicts, which range from -0.24 to 0.45. The regional spatial correlations are not consistently higher for either group, ranging from 0.08 to 0.54 among cocaine addicts and from 0.04 to 0.65 for healthy controls.

Table 2.1: Model-based estimates of spatial and temporal correlations and estimated bootstrap standard errors (in parentheses).

Group	Region	Temporal correlations (std. err.)	Region	Spatial correlations (std. err.)
Addicts				
	BA 43 (L)	0.45 (0.19)	BA 43 (L)	0.54 (0.10)
	BA 22 (R)	0.39 (0.15)	BA 29 (L)	0.50 (0.12)
	BA 42 (R)	0.37 (0.18)	BA 3 (L)	0.47 (0.10)
	BA 3 (L)	0.37 (0.17)	caudate (R)	0.47 (0.10)
	BA 9 (R)	0.34 (0.28)	BA 43 (R)	0.43 (0.07)
	BA 35 (R)	-0.24 (0.06)	BA 48 (R)	0.08 (0.07)
Controls				
	BA 22 (L)	0.64 (0.07)	BA 26 (L)	0.65 (0.05)
	BA 42 (R)	0.63 (0.09)	BA 26 (R)	0.60 (0.08)
	BA 22 (R)	0.60 (0.09)	BA 5 (R)	0.50 (0.12)
	BA 21 (L)	0.58 (0.10)	putamen (L,R)	0.50 (0.08, 0.06)
	BA 43 (R)	0.46 (0.10)	BA 38 (L)	0.47 (0.10)
	BA 26 (R)	-0.28 (0.11)	BA 48 (R)	0.04 (0.02)

inhibitory control study. The associated spatial (ρ) and temporal (ξ) dependence parameters achieve statistical significance for all regions at $\alpha = 0.01$, except for BA 29 (L), which is significant at $\alpha = 0.05$. The selected regions in Table 1 are those that have correlations with the largest magnitude. The temporal correlations for cocaine addicts range from -0.24 to 0.45 and tend to be larger in control subjects for whom they range from -0.28 to 0.64. The regional spatial correlations are not consistently higher for either group, ranging from 0.08 to 0.54 among cocaine addicts and from 0.04 to 0.65 for healthy controls.

The brain regions with highest temporal correlations differ between the two groups. For the controls, the highest temporal correlations (~ 0.6) are in right BA 42, partially covering the auditory association cortex, and in left BA 21, left BA 22, and right BA 22, regions in the temporal lobe which are also revealed by our voxel-level analysis. For the cocaine addicts, the highest temporal correlations are somewhat smaller (~ 0.4) and appear in left BA 43 and in right BA 22. Spatial correlations for the control group are highest (≥ 0.6) in left and right BA 26, a retrosplenial region in the cingulate cortex. For the cocaine addicts, the highest estimated spatial correlations (~ 0.5) are in left BA 29, also in the retrosplenial region, and in left BA 43. For both groups, the region with the lowest estimated spatial correlation (≤ 0.08) is the right BA 48, which is the second largest region (2,875 voxels). This suggests that it is the least spatially homogeneous region and that subdividing large BA's may be warranted.

2.4.4 Implications of Results

The study on inhibitory control in cocaine addicts that is the subject of this novel image analysis approach is, to the best of our knowledge, the first to examine the changes in brain activity associated with an intensive addiction behavioral therapy. Inhibiting prepotent or automatic responses is critical to the organization of successful goal-directed behaviors. Aron et al. (007a) state that it is "... likely that the simple Stop signal task taps into a control circuit . . . [and] variation (or damage) to key nodes in this control circuitry (or to their connections) could produce important individual differences, for example in liability toward and recovery from addiction (Garavan and Hester (2007)." Aron et al. (2007) demonstrated that the ability to stop motor responses depends critically on a network of structures including the right inferior frontal cortex and the subthalamic nucleus (STN), both of which are connected to the pre-SMA.

Our results agree, in several aspects, with the findings from these and other previous neuroimaging studies. Greater activation of the right orbitofrontal cortex is necessary for behavioral inhibition in impulsive individuals (Horn et al., 2003). Additionally, the right ventral prefrontal cortex has been selectively implicated in the neuropathophysiology of drug addiction (Goldstein and Volkow, 2002; Bolla et al., 2003; Volkow et al., 2005). Voxel-level analysis of our data found that the major brain effect of treatment is represented by an enabling of the right inferior prefrontal gyrus (BA 47) and right orbitofrontal cortex (BA 11) responses to a demand for inhibition of habitual motor responses. Our findings are promising since they suggest a positive effect of the behavioral therapy on the neural processing deficits associated with cocaine-addiction, though we do not strictly regard these effects to be causal.

Our analysis was not able to explore voxels in the STN because this brainstem area is not represented in the Brodmann area parcellation. This limitation is not an inherent shortcoming of our model, and one can easily include additional regions such as the STN in the analysis, if desired. The pre-SMA is identified as a significant treatment-related area in our voxel-level analysis. The caudal pre-SMA is critical to controlled action selection (Nachev et al., 2007) and is functionally compromised in cocaine addicts (Kaufman et al., 2003). Enhancing the neural response of the pre-SMA to a demand for a shift from habitual responses to controlled response inhibition represents a plausible neural correlate of drug refusal skills acquired in behavioral therapies targeting relapse prevention. Our region level analysis identifies the angular gyrus (BA 39), which is important in visuospatial attention (Cattaneo et al., 2009).

In areas exhibiting significant voxel-level post-treatment increases in task-related brain activity, we observe distinct patterns of intra-regional correlations between patients and controls. Based on the estimates of the spatial correlations, the controls exhibit higher spatial coherence in, for example, right BA's 11 $\big($controls = 0.24(0.11), addicts = 0.14(0.04)$\big)$ and 47 $\big($controls = 0.28(0.13), addicts = 0.18(0.10)$\big)$, which are frontal areas involved in executive functioning, planning and decision making, as well as in both left and right thalamus $\big($e.g. in right thalamus, controls = 0.32(0.09), addicts = 0.17(0.07)$\big)$, which play a central role in the flow of information to the cortex (Sherman and Guillery, 2002). Cocaine addicts showed more coherent functioning in left BA's 19 $\big($addicts = 0.16(0.02), controls = 0.09(0.07)$\big)$ and 39 $\big($addicts

= 0.31(0.10), controls = 0.16(0.10)), which are involved in visual attention. The latter finding is particularly interesting since BA 39 was identified in our region-level analysis.

2.5 Simulations

We performed two simulation studies. The goal of the first simulation study was to evaluate the accuracy of our estimation procedure. We selected regions (Brodmann areas) with voxel sizes ranging from 100 to 2000. To generate data for the simulations, we used our experimental data from the inhibitory control fMRI study to specify *true* values for $\boldsymbol{\beta}_{gs} = (\boldsymbol{\beta}_{gs}^{(1)}, \boldsymbol{\beta}_{gs}^{(2)})$ (where the superscripts denote sessions 1 and 2), ρ_{gs}, ξ_{gs} and ϕ_{gs}. We drew 200 samples from the multivariate normal probability distribution $\text{MVN}(\boldsymbol{\beta}_{gs}, \phi_{gs}^2 \boldsymbol{\Psi}_{gs} \boldsymbol{\Psi}_{gs})$, and we applied our spatio-temporal model to estimate the true parameters $\boldsymbol{\beta}_{gs}$, ρ_{gs}, ξ_{gs}, and ϕ_{gs}.

In the second simulation study, we evaluated the relative efficiency of our spatio-temporal model compared to the GLM, both for voxel-level and region level estimators of secondary parameters of interest (θ_g and $\boldsymbol{\theta}_g^*$). We selected two regions, region A with 100 randomly selected voxels from BA 24, and region B with 1,000 randomly selected voxels from BA 19. We set the true parameters for ρ_{gs}, ξ_{gs}, ϕ_{gs} and drew 200 samples from $\text{MVN}(\boldsymbol{\beta}_{gs}, \phi_{gs}^2 \boldsymbol{\Psi}_{gs} \boldsymbol{\Psi}_{gs})$, where $\boldsymbol{\beta}_{gs}, \phi_{gs}^2$ and $\boldsymbol{\Psi}_{gs}$ were again obtained from our experimental data. In addition to estimating the true parameters, we estimated the secondary parameters of interest $\boldsymbol{\theta}_g^*$ (voxel-level) and θ_g (region-level), defined to address the stated objective of our study (see Section 2.3.4). We calculated the estimated variances for $\hat{\boldsymbol{\theta}}_g^*$ and $\hat{\theta}_g$, both applying the GLM and our spatio-temporal model, and compared the voxel-level and region-level relative efficiencies (spatio-temporal model vs. GLM), averaged over 200 samples in each case.

Simulation results. The first simulation study revealed extremely small biases

Table 2.2: Estimation accuracy of the spatio-temporal model obtained from 200 simulated data sets for the cocaine addicts and the healthy control subjects, for four selected regions. The difference between the median of the estimated values from the generated samples and the true value is shown.

Group	Est. param.	True value	Region* (no. voxels)			
			1 (100)	2 (500)	3 (1000)	4 (2000)
			Difference (s.e.)			
Addicts	ρ	**0.5**	-0.0097 (0.0621)	-0.0001 (0.0646)	-0.0015 (0.0581)	-0.0097 (0.0650)
	ξ	**0.2**	0.0001 (0.0146)	0.0001 (0.0063)	-0.0001 (0.0044)	0.0001 (0.0032)
	ϕ^2	**1.8**	-0.1592 (0.0490)	-0.1537 (0.0213)	-0.1483 (0.0148)	-0.1501 (0.0120)
Controls	ρ	**0.5**	-0.0075 (0.0531)	-0.0011 (0.0547)	-0.0034 (0.0574)	0.0009 (0.0517)
	ξ	**0.2**	0.0012 (0.0111)	0.0009 (0.0056)	-0.0005 (0.0039)	0.0000 (0.0028)
	ϕ^2	**1.8**	-0.1137 (0.0461)	-0.1209 (0.0191)	-0.1196 (0.0140)	-0.1191 (0.0107)

* 1=BA 24(L), 2=BA 32(R), 3=BA 19(L), 4=BA 48(R). Corresponding region sizes are 100, 500, 1000, 2000 voxels, respectively. Starting value for ρ: 0.7, starting value for ξ: 0.1.

in the estimator of the mean model parameters $\boldsymbol{\beta}_{gs}$ (not shown). Table 2.2 shows the simulation results for the covariance parameters for the generated data. The following regions were selected: BA 24 (L), BA 32 (R), BA 19 (L) and BA 48 (R). The differences between the median values of the estimated parameters and the true values, from 200 generated samples are shown for each of the regions (1)-(4).

The second simulation study revealed similar results for regions A and B, and we only present the region B results for brevity. To summarize the voxel-level relative efficiency results, we plot the distribution of relative efficiencies across 1,000 voxels from region B (Figure 2.3). We see that the GLM-based voxel-level variance estimates are on average higher than the corresponding estimates obtained from our spatio-temporal model. The simulation-based estimate of the region-level relative efficiency is 1.72 for region B. This demonstrates that, on the region level, our model also leads to more precise estimates compared to GLM, and therefore may lead increased statistical power.

We found that estimates for both ρ_{gs} and ξ_{gs} (spatial and temporal parameters) are quite accurate, while the between subject variability ϕ^2_{gs} was slightly underestimated (roughly 6%), which likely stems from the well-known downward bias of ML estimates of variance components (Laird and Ware, 1998). Restricted maximum likelihood (REML) estimation can be used for our model and would presumably mitigate the observed bias. We favor ML estimation in our context, however, because it substantially facilitates computations.

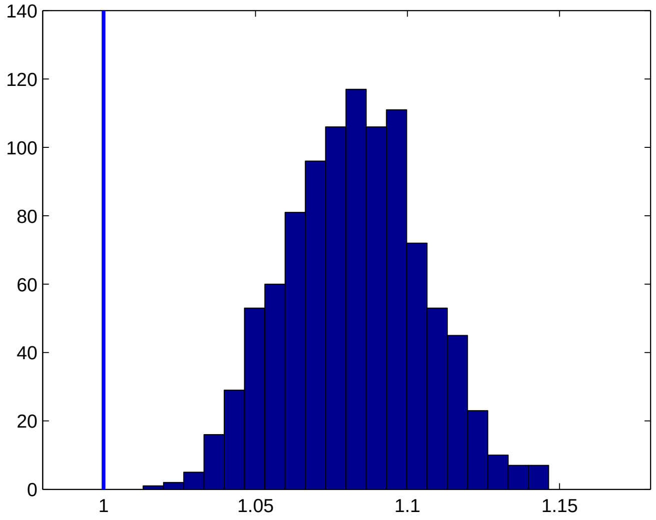


Figure 2.3: The distribution of relative efficiency values from 1,000 voxels randomly selected from left BA 19. Each voxel-level quantity is calculated as the mean relative efficiency across 200 simulated data sets. The relative efficiency was calculated based on the estimates from our model, relative to GLM.

2.6 Discussion

We propose a novel spatio-temporal modeling framework for functional neuroimaging data, overcoming unsupported assumptions of independence between the multiple summary statistics for each subject (e.g. from different sessions or stimuli) and between different brain locations within neuroanatomic regions. We used Brodmann area templates, which are popular in the neuroimaging community, to define neuroanatomic regions, but other maps may be applicable. Our proposed framework is targeted to fMRI data; however, the general ideas extend easily to PET brain imaging data measuring distributed patterns of regional cerebral blood flow. Our approach provides a unified framework for voxel-level and region-level inferences. An important contribution of our work is that we derive efficient computational solutions, e.g. enabling the inversion of large covariance matrices, to facilitate implementation. Our simulation studies demonstrated that our method is quite accurate and that the estimated standard errors are on average smaller compared to GLM, leading to better precision of estimates of model parameters, and more accurate statistical tests by borrowing the strength across related measures of neural activity.

One limitation of our model is that it does not account for correlations *between* regions. Doing so within our current framework would lead to substantially increased, perhaps prohibitive, computations. Bowman et al. (2008) establish a Bayesian framework that models between-region correlations, in addition to within-region correlations, but the number of regions included in the analysis is consequently constrained by the sample size. One advantage of our proposed spatio-temporal model is that the number of regions included is unconstrained. Our current maximum likelihood estimation procedure requires close monitoring to ensure convergence, confinement within the parameter space, and avoidance of the other specified boundary conditions.

In summary, our proposed spatio-temporal model provides an appealing, computationally efficient alternative to standard GLM-based methods for analyzing fMRI data. Our model is based on assumptions that are more neurophysiologically plausible, capturing correlations between different brain locations and between estimates of neural activity at different scanning sessions. These correlations lead to interpretive advantages over the GLM, e.g. revealing information about the degree of coherence in brain activity within defined neuroanatomic regions. Our simulation studies demonstrate that our model estimates are quite accurate and that the standard errors associated with estimates of our mean model parameters are on average smaller than those from a GLM. This increase in efficiency will often lead to more powerful statistical tests and the detection of more statistically significant voxels.

Chapter 3

Evaluating Functional Autocorrelation within Spatially Distributed Neural Processing Networks

3.1 Introduction

Data-driven statistical approaches, such as cluster analysis or independent component analysis, applied to in vivo functional neuroimaging data help to identify neural processing networks that exhibit similar task-related or resting-state patterns of activity. Ideally, the measured brain activity for voxels within such networks should exhibit high autocorrelation. An important limitation is that the algorithms do not typically quantify or statistically test the strength or nature of the within-network relatedness between voxels.

There have been numerous applications of clustering analysis to functional imaging data (Bowman et al., 2004). Despite the successful application of clustering methods to neuroimaging data, statistical procedures for testing the validity of the resulting cluster structure are still lacking. Although a clustering solution suggests the existence of networks of functionally related voxels, the clusters exhibit unknown levels of autocorrelation, which is a desirable characteristic of behaviorally relevant neural circuits. Typical cluster analyses neither quantify the strength or nature of the relatedness between within-cluster voxels using an easily interpretable measure nor apply formal tests of hypotheses to evaluate the statistical significance of the level of autocorrelation. These limitations are particularly noteworthy since clustering procedures always produce a solution, even in the absence of true functional connections in the data. Therefore, a valuable addition to cluster analyses applied in functional neuroimaging studies is to measure the degree of functional autocorrelation within a clustering solution and to evaluate the statistical significance of the observed associations.

Similarly, even though ICA has proved promising, there is a need to study the properties of ICA as applied to fMRI data. One major issue in application of ICA is that the reliability of the estimated independent components (ICs) is unknown. The results from ICA may vary considerably with different algorithm starting points, sampling of subjects, preprocessing steps or type of decomposition algorithm (Calhoun et al., 2003). Therefore, validation of estimated ICs has become important for correct interpretation of ICA results. There has been previous work related to evaluating the results of ICA by testing mutual independence between the extracted source signals (Murata, 2001; Shimizu and Kano, 2001; Chiu et al., 2003; Stogbauer et al., 2004; Wu et al., 2009). There have also been several measures proposed to evaluate the fit of the estimated ICs to the data (McKeown et al., 1998; Esposito et al., 2002; Himberg et al., 2004) and some attempts to investigate the reliability of the ICA estimates (Himberg and Hyvärinen, 2003; Himberg et al., 2004). Still, it may remain unclear how much autocorrelation is present in the identified components.

We introduce a global index to evaluate the validity of neural processing networks identified by functional clustering or by ICA. Specifically, we propose the use of Moran's I statistic (Cliff and Ord, 1973) to quantify and test the autocorrelation present within computed functional networks. We tailor our definition of the global index to neuroimaging applications through the use of network-based neighborhoods, and we compute network-specific contributions to the overall autocorrelation. Consequently, our methods target autocorrelation exhibited by voxels within defined neural processing networks, rather than between voxels globally throughout the entire brain. Thus, we make no assumptions (or statements) about statistical dependence between voxels on a global level. Hypothesis testing results from our framework augment the descriptive findings of data-driven analysis by determining the presence, direction, and statistical significance of functional autocorrelation within neural processing networks. For applications of the proposed methods to resting-state fMRI data, we present an explicit temporal component. We illustrate the use of our methodology with data from two studies: a PET study of regional cerebral blood flow (rCBF) correlates of parametrically manipulated working memory among individuals with schizophrenia and an fMRI resting-state study of depression. Moreover, we provide empirical support for the use of the proposed methods using a bootstrap analysis and a simulation study.

3.2 Experimental Data

PET Data on Working Memory in Schizophrenia Patients. We illustrate cluster analysis using PET data from a study of schizophrenia, but the cluster analysis is also applicable to fMRI data. Our data are from a parametric study of working memory in an executive function task in individuals with schizophrenia. We use data from $N = 16$ subjects, each having a total of 8 PET scans obtained using the blood flow tracer

$[^{15}O]H_2$. After aligning and re-slicing the PET images for each subject (Woods et al., 1998), we spatially normalize each subject's scans to a population-representative PET atlas (Woods et al., 1998) centered in Talairach stereotaxic coordinates (Talairach and Tournoux, 1988). We refrain from spatial smoothing of the PET data during preprocessing to avoid artificially inducing correlations between nearby voxels. We analyze data from 9,919 voxels, consisting primarily of gray-matter voxels that exhibit at least a 1.6% change throughout the study. The study design encompasses 2 replicate sessions of 4 scans, where the scans in each session represent four experimental conditions that parametrically vary working memory load-the ability to manipulate information in short–term memory. Specifically, the conditions are defined as digit shadowing (minimal working memory load) and low, moderate, and high working memory loads in the form of serial addition tasks. In the digit shadowing condition, the subjects simply repeat a number that they receive by an auditory presentation, which imposes a negligible load on working memory. In the other three experimental conditions, the subjects receive auditory presentations of a series of positive integers and are instructed to provide the sum of the current number and the preceding number, which requires that they store the previous number in short-term memory and mentally suppress the previous sum. The range of the integers in the working memory load presentations distinguishes the conditions. Both numbers in the low load condition fall between 1 and 3, both numbers in the moderate load condition are between 1 and 5, and the high load condition includes integers between 1 and 9. On average, the sums for the three active load conditions during our experiment are 4.23 (low), 7.22 (moderate), and 10.09 (high).

FMRI Data from a Study on Depression. We use data from a second study intended to characterize the impact of childhood abuse/early trauma, and its developmental course, on brain structure and connectivity. For our purpose, we analyze fMRI resting-state scans of seven women with a history of major depression. While in the scanner, subjects had their eyes open and were looking at a visual fixation cross. Functional data were acquired on a Siemens 3T HRRT (high resolution research tomography) scanner, in 3D mode. During a 7.5 min acquisition, a series of 210 scans were acquired with TR=2sec, 20 axial slices, $3.4 \times 3.4 \times 4$ mm^3 resolution. The functional runs were collected with a Z-saga sequence to avoid orbitofrontal signal ablation. A group ICA was applied to investigate the resting-state functional connectivity of the depressed patients. There is substantial interest in determining associations between the brain activity characteristics of different regions while subjects are in a resting state. For example, a set of regions has been consistently identified in these investigations and has been labeled as the *default mode network* (DMN).

The first step in the analysis (preprocessing and the group ICA) of the fMRI data was done in FSL (FMRIB's Software Library, `www.fmrib.ox.ac.uk/fsl`) software package. The preprocessing and the first level, individual subject analysis of the PET data was done in SPM (Statistical Parametric Mapping, Welcome Department of Cognitive Neurology, London, UK, `www.fil.ion.ucl.ac.uk/spm/`). The rest of the analyses were carried out in Matlab.

3.3 Methods

Notation and terminology. The framework we present allows estimation (and testing) for the presence of autocorrelation separately for each experimental condition p ($p = 1, \dots, P$) or scan s ($s = 1, \dots, S$). Data from a single scan are represented as a 3-D rectangular lattice, comprised of a large number of voxels, indexed by a mapping of coordinates (x, y, z) to $v = 1, \dots, V$. Each voxel contains intensity information corresponding to measures of localized brain activity for fMRI data, or rCBF in case of PET data.

3.3.1 Clustering

Summary statistics, such as means or regression coefficients, are often inputs for clustering procedures applied to task-related fMRI or PET data (Simon et al., 2004). We develop notation for this case, but the methodology extends to settings that cluster the original response data directly, e.g. in resting-state fMRI studies. We denote rCBF, as a proxy for brain activity, by $\mathbf{Y}_k(v) = \big(Y_{k1}(v), \ldots, Y_{kS}(v)\big)'$, representing localized blood flow measurements from all S scans for subject k, $k = 1, \ldots, K$ (at voxel v). For clustering, we typically consider statistics $\mathbf{T}(v) = \big(T_1(v), \ldots, T_P(v)\big)' = f\big(\mathbf{Y}_1(v), \ldots, \mathbf{Y}_K(v)\big)$, $P \leq S$, that summarize data from all individuals, e.g. where $T_p(v)$ is the mean or estimated effect associated with experimental condition p. By selecting the pth element of the summary vector from every voxel, we also define the vector $\mathbf{T}_p = \big(T_p(1), \ldots, T_p(V)\big)'$, which will facilitate our upcoming discussion of functional autocorrelation.

Clustering procedures classify the V voxels in an image into G groups, with each cluster consisting of V_g voxels, where $g = 1, \ldots, G$ and $V = \sum_{g=1}^{G} V_g$. Most methods classify voxels by measuring the distance (dissimilarity) between the activity time courses for every pair of voxels and combining voxels with small distances. For example, one popular measure of distance between the activity in voxels i and j is the Euclidean distance given by

$$d\big(\mathbf{T}(v_i), \mathbf{T}(v_j)\big) = \Big[\big(\mathbf{T}(v_i) - \mathbf{T}(v_j)\big)'\big(\mathbf{T}(v_i) - \mathbf{T}(v_j)\big)\Big]^{1/2}. \tag{3.1}$$

Descriptions of particular clustering algorithms are available in Rencher (2002), and Bowman et al. (2004) highlight several of these algorithms for neuroimaging applications.

Some clustering criteria and stopping rules for hierarchical clustering methods quantify measures based on within-cluster similarity, e.g. proportional to the change in within-cluster variation, relative to between-cluster variation. However, these quantities are limited because they are relative measures, e.g comparing two consecutive levels of a clustering hierarchy. Also, these measures often lack easy interpretability and are not amenable to formal hypothesis testing. The probability distributions of clustering criteria are difficult to derive theoretically (Hartigan, 1977, 1978) and thus do not provide a basis for formal hypothesis testing using a conventional Neyman-Pearson framework. Furthermore, even if one successfully derives the distribution of a particular clustering criterion, the result may be somewhat limited in practice in cases where an alternative clustering algorithm provides better performance. We evaluate statistical significance of the functional autocorrelation present in the final clustering solution, regardless of the clustering algorithm employed to obtain the solution and without relying on specific distributional assumptions of the data.

3.3.1.1 ICA

Classical ICA. Let $\mathbf{Y}$ denote an $S \times V$ matrix of the observed voxel time courses, where S is the total number of scans, and V is the number of voxels included in the analysis; $\mathbf{C}$ is an $N \times V$ random matrix whose rows C_i are to be filled with the unknown realizations of the N spatial components (images, reshaped in 1-D vectors), and $\mathbf{A}$ is an $S \times N$ mixing matrix, whose columns contain the associated time-courses of the N components. The spatial ICA problem for fMRI time series can be formulated as an estimation of the following model:

$$\mathbf{Y} = \mathbf{A}\mathbf{C}\,. \tag{3.2}$$

There are no assumptions about the mixing matrix $\mathbf{A}$, while the constraint on the spatial processes C_i is that they are (ideally) mutually statistically independent. The amount of statistical dependence within a fixed number of spatial components can be quantified by means of their mutual information, an important function in information theory (Common, 1994). Thus, the ICA decomposition of $\mathbf{Y}$ can be defined (up to a multiplicative constant and to the sign) as an invertible transformation: $\mathbf{C} = \mathbf{UY}$, where the matrix $\mathbf{U}$ (so-called unmixing matrix) is determined such that the mutual information of the target components C_i is minimized (i.e., such that C_i's are "as independent as possible"). The matrix $\mathbf{A}$ is the (pseudo)inverse of $\mathbf{U}$. In McKeown et al. (1998), the sources are estimated by iteratively optimizing the unmixing matrix $\mathbf{U}$ so that $\mathbf{C} = \mathbf{UY}$ contains mutually independent rows, using the 'infomax' algorithm.

Probabilistic ICA (PICA). The model of equation (3.2) does not include random noise. The PICA model, which extends model (3.2), is formulated as a generative linear latent variables model. The model "is characterized by assuming that the S-variate vector of observations is generated from a set of q statistically independent non-Gaussian sources via a linear instantaneous mixing process corrupted by additive noise $\boldsymbol{\eta}(t)$" (Beckmann and Smith, 2004):

$$\mathbf{y}_v = \mathbf{A}\mathbf{c}_v + \boldsymbol{\mu} + \boldsymbol{\eta}_v, \quad \forall v \in \{1, \ldots, V\}.$$

where $\mathbf{y}_v$ denotes the S-dimensional column vector of individual measurements at voxel location v, $\mathbf{c}_v$ denotes the q-dimensional column vector of non-Gaussian source signals contained in the data and $\boldsymbol{\eta}_v$ denotes Gaussian noise $\boldsymbol{\eta}_v \sim N(0, \sigma^2\boldsymbol{\Sigma}_v)$. The number of source processes is assumed to be smaller than the number of observations in time ($q < S$). The model is closely related to factor analysis. The individual component maps are obtained using a modified fixed-point iteration scheme to optimize for non-Gaussian source estimates via maximizing the negentropy. The maps are then separately transformed to spatial Z-scores. These Z-score maps depend on the amount of variability explained by the entire decomposition at each voxel location relative to the residual noise. Next, Gaussian/Gamma mixture models are fitted to the individual Z-maps in order to infer voxel locations that are significantly modulated by the associated time-course. For more details on PICA, we refer the reader to Beckmann and Smith (2004).

Group ICA. In Beckmann and Smith (2005), the single-session PICA has been extended to higher dimensions allowing for a model-free analysis of multi-subject or multi-session fMRI data. For group ICA, either tensorial ICA (where the data is decomposed into spatial maps, time courses and subject/session modes), or a simpler temporal concatenation approach (a single 2D ICA is run on the concatenated data matrix obtained by stacking all 2D data matrices of every single data set on top of each other) are possible. The latter approach is recommended to use when the interest is in finding a common spatial pattern, but we cannot assume that the associated temporal response is consistent between subjects, such as in the analysis of data acquired without stimulation (e.g. resting-state data). We hence apply this approach for IC analysis of our fMRI data set. The technique is derived from parallel factor analysis (Harshman, 1970; Harshman and Lundy, 1994). [1]

We define $\mathbf{Y}(v) = \big(\mathbf{Y}_1(v)', \ldots, \mathbf{Y}_K(v)'\big)'$, representing concatenated measures of localized brain activity at voxel v, across all subjects. In our upcoming fMRI example, $\mathbf{Y}(v)$ is a $(210 * 7) \times 1$ vector. We also define the mean vector of brain activity, across all voxels in all ICs as $\overline{\mathbf{T}} = \frac{1}{V}\sum_{v=1}^{V}\mathbf{Y}(v)$.

3.3.2 Functional Autocorrelation Statistic

Functional autocorrelation measures the extent to which voxels within neural processing networks exhibit similar patterns of brain activity. For any selected element of the summary vector (indexed by p), we measure the functional autocorrelation of

---

[1] The research by Beckmann and Smith described above has been implemented as MELODIC (Multivariate Exploratory Linear Optimized Decomposition into Independent Components - a stand alone C++ program). It is freely available as part of FSL (FMRIB's Software Library).

a clustering, or ICA solution using Moran's I (Moran, 1950), defined as

$$I_p = \frac{V}{S_0} \frac{\mathbf{Z}_p' \mathbf{W} \mathbf{Z}_p}{\mathbf{Z}_p' \mathbf{Z}_p} \,. \tag{3.3}$$

In the case of clustering, $\mathbf{Z}_p = [\mathbf{T}_p - (\overline{T}_p \mathbf{1}_V)]$, with $\mathbf{1}_V$ denoting a $(V \times 1)$ vector of ones and $\overline{T}_p$ denoting the mean summary element across all voxels (i.e., $\overline{T}_p = \frac{1}{V}\sum_{v=1}^{V} T_p(v))$; $S_0 = \sum_{g=1}^{G}(V_g - 1)V_g$ and $\mathbf{W}$ is the symmetric spatial connectivity matrix representing the clustering solution. The matrix $\mathbf{W}$ defines all pairs of voxels that are in the same cluster, i.e., that exhibit functional connectivity, by including a nonzero value in the appropriate off-diagonal element. Specifically, we define $\mathbf{W}$ such that the diagonal elements are $W_{vv} = 0$, and the (u, v)th off-diagonal element is either $W_{vu} = 1$, if voxels u and v are in the same cluster, or $W_{vu} = 0$, otherwise.

When applied to group ICA results, we define $\mathbf{Z} = [\big(\mathbf{Y}(1)', \ldots, \mathbf{Y}(V)'\big)' - (\mathbf{1}_V \otimes \overline{\mathbf{T}})]$, where $\otimes$ denotes the Kronecker product, or the direct product operator. Here we drop the index p for convenience, since in our fMRI data example there is only one experimental condition. $\mathbf{W}$ is now a block matrix: $\mathbf{W} = \mathbf{W}^1 \otimes \mathbf{I}_{(K*S)\times(K*S)}$, where $\mathbf{W}^1$ is a $V \times V$ connectivity matrix (i.e. $W^1_{vv} = 0$ and $W^1_{vu} = 1$, if voxels u and v are in the same IC and 0 otherwise). V and S_0 are the same as defined above. Connectivity matrices in other areas of statistical application, e.g. geostatistics, typically specify connections between locations that are physically adjacent, e.g. neighboring cities. We adapt Moran's I by tailoring our definition of adjacency to reflect the distributed neural processing networks, which typically contain voxels that are not all spatially contiguous. Since our application defines the proximity of voxels based on measures of brain function, rather than physical or spatial distance, we refer to Moran's I as a *functional* autocorrelation index. However, the phrase *spatial* autocorrelation is also appropriate since we view functional networks as spatially dissociable regions on a brain map.

The massive amount of data collected in functional neuroimaging studies yields high dimensional spatial connectivity matrices, which may lead to computational difficulties. In the case of clustering (and similarly for ICA), we can re-express I_p, using the following, perhaps more computationally convenient, expression

$$I_p = \frac{V}{S_0} \frac{\sum_{v=1}^{V} \sum_{u=1}^{V} W_{vu}[T_p(v) - \overline{T}_p][T_p(u) - \overline{T}_p]}{\sum_{v=1}^{V} [T_p(v) - \overline{T}_p]^2} . \tag{3.4}$$

A similar formula can be obtained when (3.3) is applied to an ICA solution from fMRI data. In (3.4), the numerator of Moran's I calculates the products of the mean-centered summary statistics for within-cluster voxel pairs and computes the sum of all such products. The denominator gives a measure of variation about the overall mean. These interpretations reveal the conceptual similarity of Moran's I to the Pearson correlation coefficient, with Moran's I representing a spatially weighted version of Pearson's measure. Moran's I generally ranges between -1 and 1, but the precise limits depend on the eigenvalues of $\mathbf{W}$. A large positive value of I_p indicates the presence of functional clusters in which there is high within-cluster similarity and a negative value indicates clusters exhibiting dissimilarity between voxels. Randomly assigning voxels to clusters typically results in uncorrelated rCBF between the voxels within each group.

3.3.3 Hypothesis Testing

The randomization scheme, characterized by the random allocation of voxels to neural processing networks, serves as the basis for establishing the null hypothesis and related distributional properties for evaluating the statistical significance of any observed value of Moran's I. To conduct a test to determine if statistically significant functional autocorrelation is present in a clustering solution, compute the test statistic

$$Z_p = \frac{I_p - E(I_p)}{\sqrt{V(I_p)}} \, . \tag{3.5}$$

where $E(I_p)$ and $V(I_p)$ represent the theoretical mean and variance of Moran's I under randomization, respectively (we provide expressions below). Next, calculate $\Pr(|Z_p| > z_{\alpha/2})$, the probability that the magnitude of the observed statistic Z_p exceeds the critical value from a standard normal distribution corresponding to type-I error at level α. The test rejects the null hypothesis of uncorrelated networks, e.g. no true cluster structure, when the absolute value of Z_p is sufficiently large. Therefore, a statistically significant test result indicates that the observed functional autocorrelation exhibited by a clustering or ICA solution is large relative to the expected value of the autocorrelation under randomization (Goodchild, 1986).

Calculating the test statistic Z_p requires expressions for the expected value and the variance of Moran's I under the random assignment of voxels to functional networks. The mean and the variance of Moran's I under randomization are given by following formulas (Moran, 1948, 1950; Cliff and Ord, 1973, 1981):

$$\mathrm{E}(I_p) = -\frac{1}{V-1} \, , \tag{3.6}$$

$$\begin{aligned} \mathrm{V}(I_p) = {} & \frac{V\big[(V^2 - 3V + 3)S_1 - VS_2 + 3S_0^2\big]}{(V-1)(V-2)(V-3)S_0^2} \\ & - \frac{m_p\big[(V(V-1)S_1 - 2VS_2 + 6S_0^2\big]}{(V-1)(V-2)(V-3)S_0^2} - \mathrm{E}^2(I_p) \end{aligned} \tag{3.7}$$

where $S_1 = 2S_0$, $S_2 = \sum_{g=1}^{G} 4V_g(V_g - 1)^2$, and

$$m_p = \frac{V^{-1}\sum_{v=1}^{V}\big[T_p(v) - \overline{T}_p\big]^4}{\big\{V^{-1}\sum_{v=1}^{V}\big[T_p(v) - \overline{T}_p\big]^2\big\}^2} \, . \tag{3.8}$$

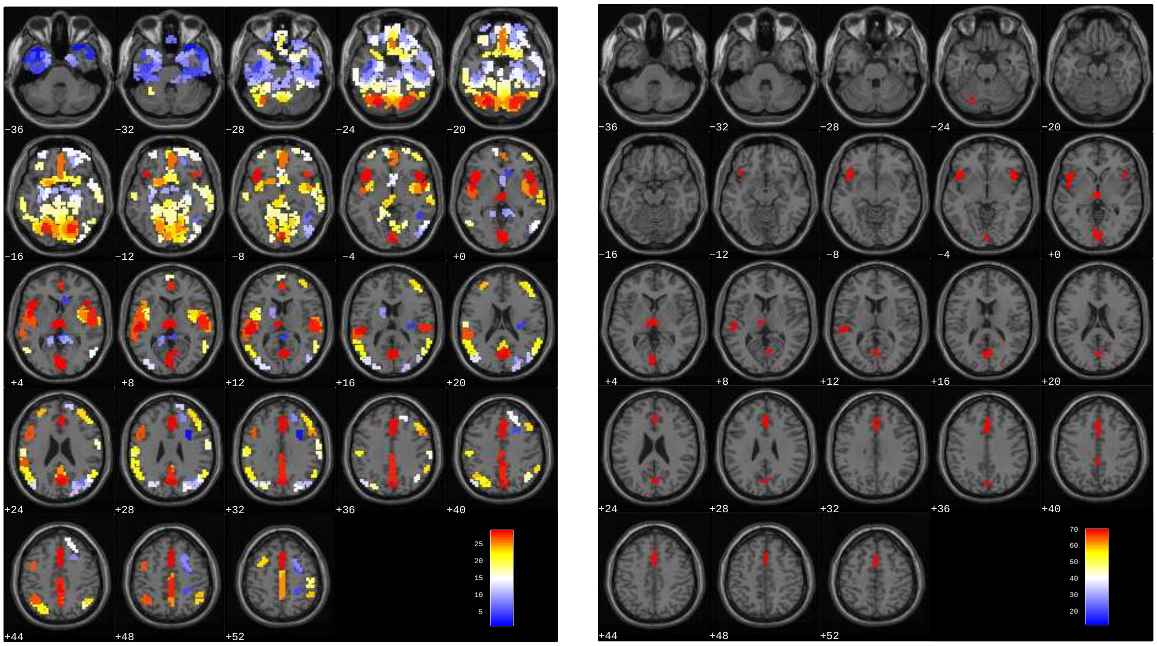
Using Monte Carlo simulations based on our PET data, we empirically validate the theoretical properties of Moran's I under randomization to protect against potential bias, e.g. caused by intrinsic spatial autocorrelation in PET data (see the Results section, Bootstrap analysis). Substituting the computed quantities (3.6)–(3.8) into equation (3.5) gives the observed value of the test statistic upon which to base hypothesis testing and statistical inferences regarding functional autocorrelation.

3.4 Results

Working Memory Data. We begin by fitting the working memory data using a general linear model with expected value parameters representing the four experimental conditions (digit shadowing and low, moderate, and high working memory loads) and a covariate adjustment for global cerebral blood flow (gCBF). We obtain ordinary least-squares estimates of the regression parameters and perform clustering using the vector of summary statistics $\mathbf{T}(v) = \hat{\boldsymbol{\beta}}(v)$ for each voxel. We perform hierarchical clustering using Ward's method (Rencher, 2002) and determine the number of clusters using the cubic clustering criterion (CCC) (Searle, 1983). Generally, CCC compares the square of the observed correlation coefficient (R^2) and an approximation of its expected value under specified regularity conditions. A plot of CCC ranging from 1 to 40 clusters (Figure 3.2) clearly identified $G = 29$ as the number of clusters present in the data. Therefore, the final clustering solution contains 29 clusters, and we depict the cluster map in Figure 3.1(a).

The cluster sizes range widely with the smallest cluster (1) containing only 8 voxels and the largest cluster (17) containing 773 voxels. The average number of voxels in a cluster is 342 with a standard deviation of 165 voxels. We calculated the mean normalized rCBF values for all voxels within a cluster (averaged across levels of working memory load), and arranged them in increasing order. Cluster 1 contains voxels with outlying observations and has a mean of only 14.0. The most

(a) Cluster Map (b) Cluster 29



(c) Cluster 23 (d) Cluster 20

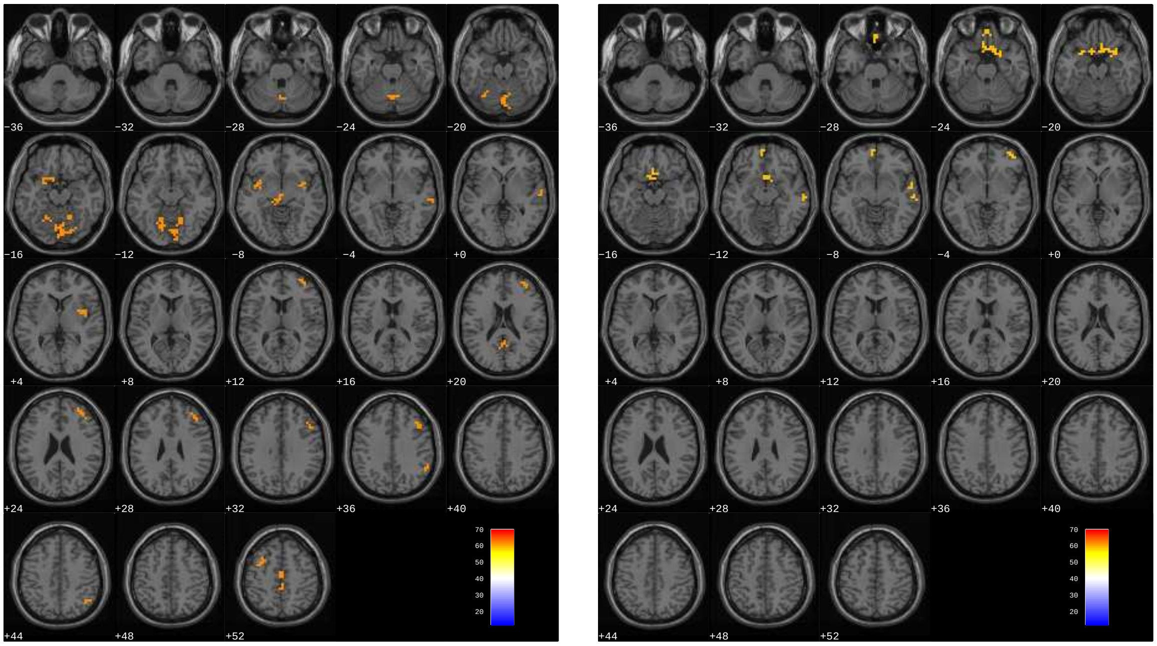


Figure 3.1: Cluster maps from Ward's method displayed on axial brain slices ranging from -36mm to +52mm relative to the anterior/posterior commissure plane (z). (a) View of the 29 clusters, with each cluster containing voxels that exhibit similar responses across varying loads on working memory. (b) Cluster 29 includes voxels in the inferior frontal gyrus (BA 47), thalamus, anterior and posterior cingulate (BA 30;32), and the lingual gyrus (BA 17;18). (c) Cluster 23 consists of voxels in the precentral and middle frontal gyri (BA 9), in the middle, medial, and superior frontal gyri (BA 6), and along the supramarginal gyrus (BA 40). (d) Cluster 20 spans portions of the right middle frontal gyrus (BA 10), the left medial frontal gyrus (BA 11), and the anterior cingulate (BA 25).

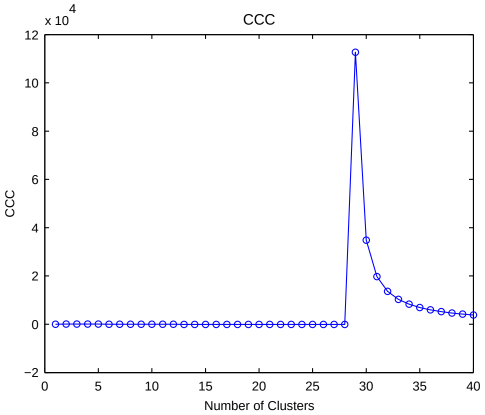
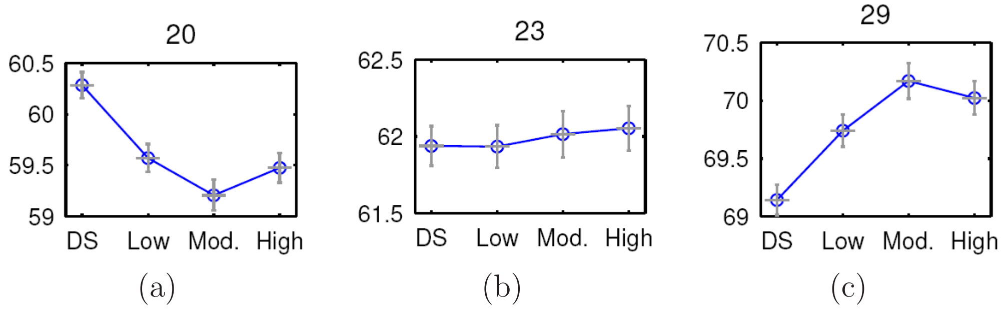


Figure 3.2: Plotted values of the cubic clustering criterion (CCC). The figure identifies 29 clusters present in the data.

active cluster (29) has a mean value of 69.8. Ten of the clusters identified by Ward's method exhibit brain activity that is lower than the global average of 50, while the other clusters show increased activity relative to the global average value.

Typical cluster analyses cease after establishing a cluster solution and perhaps quantifying associated descriptive statistics. While the descriptive cluster map provides some insights on functional associations in the brain, it does not quantify the similarity of voxels within the 29 clusters. Our methodology allows us to compute a global measure of the functional autocorrelation using Moran's I statistic and to evaluate the validity of the clustering solution by examining the statistical significance of the autocorrelation. Overall, the clusters exhibit very strong positive functional autocorrelation for the digit shadowing condition (0.80), the low working memory load condition (0.81), the moderate load condition (0.81), and the high load condition (0.81). The large positive values of the functional autocorrelation index suggest that there is generally strong coherence between the voxel profiles within the defined neural processing clusters. Furthermore, these functional autocorrelations are highly statistically significant with p-values all less than 0.0001, confirming the validity of the clustering solution relative to the expected associations under spatial randomness.

Our analysis identifies clusters containing voxels that, on average, reveal activ-



(a) (b) (c)

Figure 3.3: Mean rCBF profiles, along with standard errors, for 3 representative clusters across levels of working memory load ranging from digit shadowing (DS) to the high load. Many clusters (not all shown here) exhibit nonlinear trends over the parametrically varied memory loads.

ity profiles resembling particular polynomial curves as a function of varying working memory load levels. The representative mean profiles for three clusters appear in Figure 3.3; we display the associated standard errors using vertical bars. The plots reveal potentially important functional characteristics of the relationship of cluster-level activity to working memory. The nonlinear trends apparent in the plots provide insights on the nature of the stimulus-response relationship in the cluster and may assist in guiding subsequent statistical modeling, particularly in studies of parametrically varying experimental conditions.

Cluster 29 is the most active cluster, i.e. has the highest normalized rCBF, associated with the neural processing underlying working memory. The voxels within this cluster reveal a roughly quadratic trend in rCBF, exhibiting an increase in measured brain activity from digit shadowing through the moderate working memory load, followed by slight attenuation in the activity at the high load (see Figure 3.3(c)). Axial slices of the cluster map in Figure 3.1(b) show that this neural processing cluster consists of voxels localized bilaterally in the inferior frontal gyrus (Brodmann area (BA) 47) (Brodmann, 1909b), extending into the left superior temporal gyrus (BA 22); thalamus; anterior (BA 32) and posterior (BA 30 and 31) cingulate cortex; and the lingual gyrus (BA 17 and 18). The quadratic trend that these voxels display suggests that, in individuals with schizophrenia, the activity in this cluster increases to accommodate more substantial memory loads, but then diminishes when the working memory load exceeds capacity. There was not a precipitous fall in the performance of the serial addition task imposing a high working memory load, rather a near linear decrease in accuracy from the digit shadowing to high load conditions. Performance at high load levels may be subsumed by other clusters (e.g. cluster 20).

Cluster 20 defines a network of functionally related voxels spanning portions of the prefrontal cortex including the right middle frontal gyrus (BA 10), left ventromedial frontal gyrus (BA 11), and the subcallosal cortex (BA 25), all visible in axial slices $z = -12$ mm through -4 mm of Figure 3.1(d). As indicated by Figure 3.3(a), these voxels exhibit an approximate quadratic pattern of task-related activity with declining activity from digit shadowing to moderate working memory load and a slight increase at the high load level.

The full set of individual cluster maps identify functional associations related to the neural activity involved in performing parametrically varying working memory tasks. Here, our detailed inspection of two clusters reveals distinct profiles of task-related brain activity and reveals associations between brain regions that have been previously linked to working memory tasks. Numerous studies have detected the involvement of both dorsal lateral and ventral lateral regions of the prefrontal cortex in working memory functions (D'Esposito, 2001). In addition, several studies have implicated distributed activations in the posterior parietal cortex (BA 40/7), anterior cingulate cortex (BA 32), and the thalamus, as well as premotor areas (BA 6 and 8) (D'Esposito et al., 1998; Perlstein et al., 2003; Picard and Strick, 1996).

The global autocorrelation measures augment the descriptive findings of classification procedures by evaluating the strength and direction of the associations within the defined neural processing clusters. Local measures of autocorrelation do not extend readily to neuroimaging classification applications because they would require definitions of within-cluster adjacency or functional connections. To provide a crude

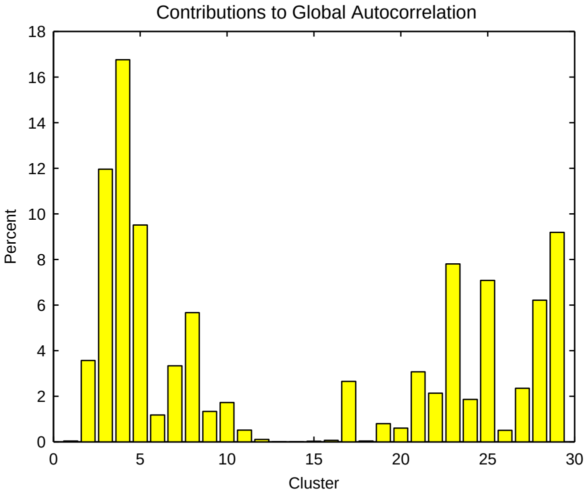


Figure 3.4: The cluster-specific contributions to the overall autocorrelation index.

measure of the contribution that each cluster has on the overall Moran's I autocorrelation statistic, we compute the percent contribution of each cluster. Figure 3.4 displays the results of this local measure. The plot reveals that several clusters with large or small mean rCBF values contribute more heavily to the global autocorrelation measure. Generally, a cluster's contribution to the overall level of global autocorrelation gives an indication of the degree of similarity of rCBF profiles from voxels within the cluster and of the spatial extent of the cluster. Both a high degree of within-cluster homogeneity in working memory-related rCBF values and a large cluster size lead to greater influence on the overall level of functional autocorrelation. Clusters 3, 4, 5, 23, and 29 have the largest impact on the overall functional coherence, attesting to the high degree of functional autocorrelation in these clusters.

Depression Data. We first perform a group ICA analysis of the fMRI data using PICA (Beckmann and Smith, 2004) as implemented in the FSL tool MELODIC (Multivariate Exploratory Linear Decomposition into Independent Components) Version 3.09. The following data pre-processing was applied to the input data: masking of non-brain voxels; voxel-wise de-meaning of the data; normalisation of the voxel-wise variance. Pre-processed data were whitened and projected into a 33-dimensional sub-

IC	% of global autocorrelation	# voxels	BA's
4	12.41	843	7, 9, 10, 11, 23, 30, 39, 40
5	6.88	1032	putamen, 3, 4, 6, 48
6	7.24	968	thalamus(l), 6, 9, 10, 46
17	10.65	1464	6, 9, 24, 32, 46
27	0	45	21(r), 22(r), thalamus (r)
29	4.24	827	40(r), 44(r), 45(r), 47

Table 3.1: Network-specific contributions to the overall autocorrelation index of selected ICs (as identified by ICA) of the fMRI data from the study on depression. (l and r denote the 'left' and 'right' hemisphere).

space using Principal Component Analysis. The whitened observations were decomposed into sets of vectors which describe signal variation across the temporal domain (time-courses), the session/subject domain and across the spatial domain (maps) by optimizing for non-Gaussian spatial source distributions using a fixed-point iteration technique (Hyvärinen, 1999). Estimated component maps were divided by the standard deviation of the residual noise and thresholded by fitting a mixture model to the histogram of intensity values (Beckmann and Smith, 2004). The optimal number of components determined by the PICA algorithm (Beckmann and Smith, 2004) is 33. Many brain voxels were allocated to more than one IC, and some voxels were not assigned to any of the ICs. Out of a total of 19,257 intracranial voxels, 17,197 were assigned to at least one of the ICs. We assign each of the latter voxels to a single IC, based on the largest (in magnitude) Z-score.

For this data set, the functional autocorrelation index is 0.17. Even though it is modest compared to the one for PET data, it is statistically significant with a p-value less than 0.0001. There are only 4 ICs for which the crude estimates of local autocorrelation are > 0.01, and Figure 3.5 shows three of them. IC 4 represents the DMN, and IC 6 contains voxels in several brain areas (e.g. dorsolateral prefrontal cortex (BA 9), anterior prefrontal cortex (BA 10), and left thalamus) relevant in

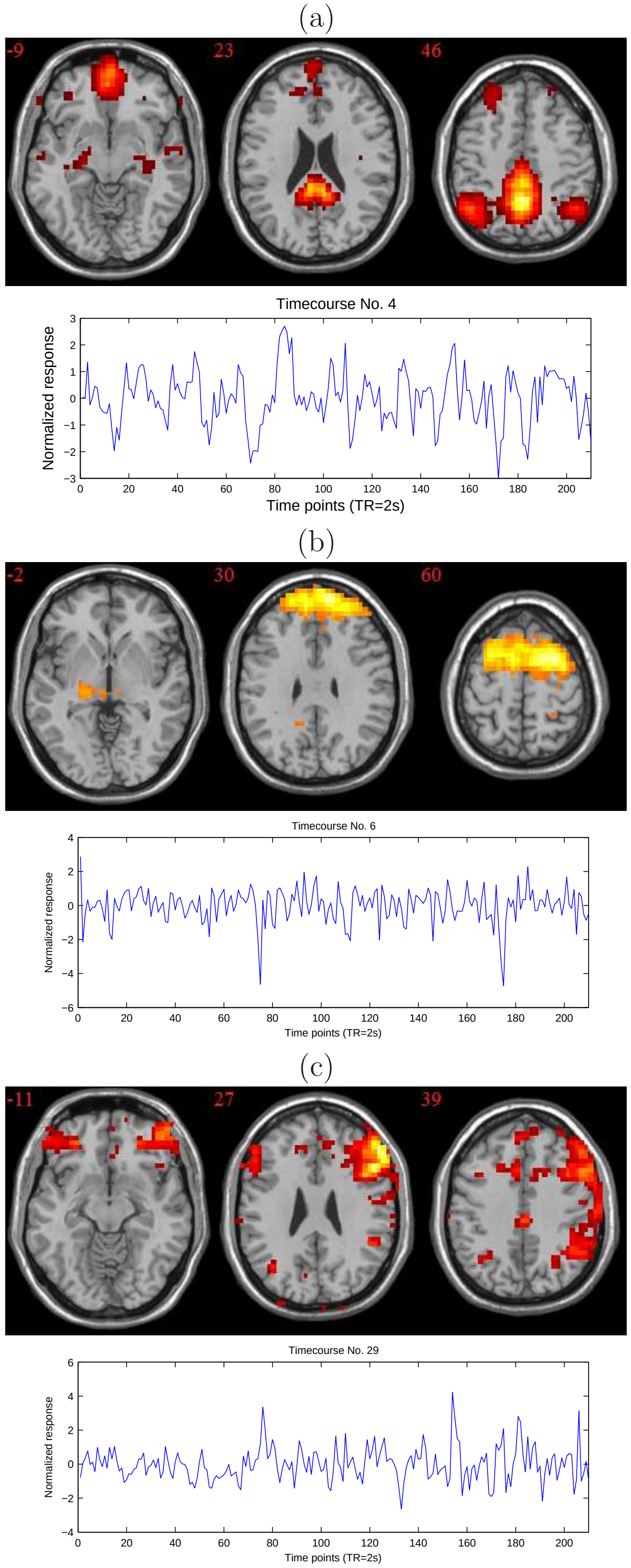


Figure 3.5: ICA results with optimal number of components (33); (a) IC 4 (with the corresponding time course): the DMN; (b) IC 6, (c) IC 29.

depression (Mayberg, 1997; Goldapple et al., 2004)). Table 3.1 lists the IC-specific contribution to the global autocorrelation index, number of voxels and Brodmann regions in which the IC's voxels lie for several ICs.

Since for this data set, we calculate the global Moran's I statistic based on the voxels' time courses (not the summary statistic vectors), the statistic I is small, mainly due to the fact that the denominator in (3.3) increases very fast since the number of time points is large (210). For this reason, we apply another approach to evaluate the global autocorrelation. We calculate the global Moran's I separately for each time point, this way obtaining a $(210 * 7 \times 1)$ vector of global indices of autocorrelation, for our ICA solution. The obtained values are of a larger scale. They range from 0.03 to 0.53, with the peak values within subjects reaching at least 0.3. The corresponding IC-specific contributions to the overall autocorrelation index agree with the previous results reported in Table 3.1. Namely, the ICs with the largest range of 'local' Moran's I statistics are still IC 4 (range of I's [0, 0.19]), IC 5 (range of I's [0, 0.14]), IC 6 (range of I's: [0, 0.15]), and IC 17 (range of I's [0, 0.26]). Therefore, both approaches bring us to similar conclusions about which of the 33 ICs are most functionally and spatially coherent neural networks. There are several possible reasons why, for this data set, the estimated global autocorrelation index is relatively small. First, the sample size is small for a group analysis (only 7 subjects). This is resting-state data, so the signal to noise ratio is likely to be lower than for task data (such as our PET data set). Also, even though the ICA analysis should ideally result in functionally correlated ICs, the approach is based on conceptually different criteria which results in ICs that are as statistically independent as possible.

Bootstrap Analysis We do not perform spatial smoothing of the PET data prior to conducting the cluster analysis to avoid artificially inflating the correlations between the rCBF measurements in nearby voxels. A related issue is whether the intrinsic spatial correlations in PET data, e.g. stemming from the acquisition and reconstruction

Mean	Working memory load			
Moran's I	digit shadowing (DS)	low	moderate	high
Theoretical	-0.000101	-0.000101	-0.000101	-0.000101
Bootstrap	-0.00033	-0.00025	- 0.00029	-0.00030
Bias	-2.29×10^{-4}	-1.49×10^{-4}	-1.89×10^{-4}	-1.99×10^{-4}
Variability	0.62×10^{-3}	0.62×10^{-3}	0.63×10^{-3}	0.63×10^{-3}

Table 3.2: Summary of the Bootstrap analysis designed to evaluate a potential bias of Moran's I (e.g. caused by intrinsic spatial autocorrelation) in PET data.

processes, may bias Moran's I toward higher values. If true, then the upward bias of Moran's I would call for modifications to the hypothesis testing framework outlined previously. We conduct a Monte Carlo simulation study to evaluate the expected value and the variability of Moran's I under the null hypothesis of random cluster assignments.

We conduct a bootstrap analysis that includes 500 samples drawn, with replacement, from the working memory data. For each sample, we compute summary statistics representing the four working memory load conditions, adjusted for gCBF. To establish the null distribution of Moran's I, we randomly allocate voxels to clusters in each bootstrap sample, with the cluster sizes matching those obtained in our original analysis, and then we calculate Moran's I for the resulting randomized clustering solutions. The 500 computed statistics represent the empirical distribution of Moran's I under the null hypothesis. The bootstrap distribution allows a comparison of empirical and theortical properties of Moran's I including both its mean and variance.

Table 3.2 summarizes the results of the bootstrap analysis. Theoretically, the mean of Moran's I is equal to -0.000101. The similarity of the theoretical and empirical mean values provides strong evidence of negligible bias in Moran's I for our PET application. We find the empirical distributions of the bootstrap biases, reflecting estimates of Moran's I obtained from the boostrap samples (under randomization) minus the theoretical value of the mean. Averaging over all the bootstrap samples gives the bootstrap estimate of bias, for each condition. Although slightly negative, the biases are all extremely small. Similarly, the bootstrap variance estimates exhibit close correspondence to the theoretical values. The magnitudes of all differences between the empirical and theoretical variances are less than 0.4224×10^{-6}. The accuracy of the theoretical mean and variance expressions relative to their empirical counterparts obtained from the bootstrap estimates supports the use of the hypothesis testing procedure outlined in the Methods section.

Simulated Data Example. We analyze simulated data based on the working memory PET study to illustrate further the utility of Moran's I for detecting the presence of functional autocorrelation in clustering solutions. We simulate data using a mixture model that takes a weighted average of the global mean and the mean of the voxel-specific summary values within the same neural processing cluster, with weights defined by a spatial dependence (correlation) parameter ρ. Specifically, we generate data for a particular voxel within cluster g using

$$\hat{T}_v = (1-\rho)\mu + \rho \overline{T}_{N_v} + \varepsilon_v, \tag{3.9}$$

where $\varepsilon_v \sim N(0, \sigma_v^2)$, μ is the overall mean, e.g. 50; and $\overline{T}_{N_v}$ is the mean of the summary statistics from the $V_g - 1$ voxels in the same cluster as voxel v, i.e., from the functional *neighborhood* of v, denoted by N_v. We consider six simulated data sets with the correlation parameter varying across the set of values $\rho \in \{0, 0.2, 0.4, 0.6, 0.8, 1\}$. We use the 29 clusters computed from the working memory data to define the cluster structure and the corresponding cluster sizes V_g. We utilize estimates of the variance parameters and the summary statistic vector from the working memory data.

The error terms from different voxels in model (3.9) are independent. However, the model yields simulated values that depend on the activity levels from other voxels within the same neural processing cluster, when $\rho > 0$. The activity in a given voxel is uncorrelated with the activity of the other voxels within the same cluster when

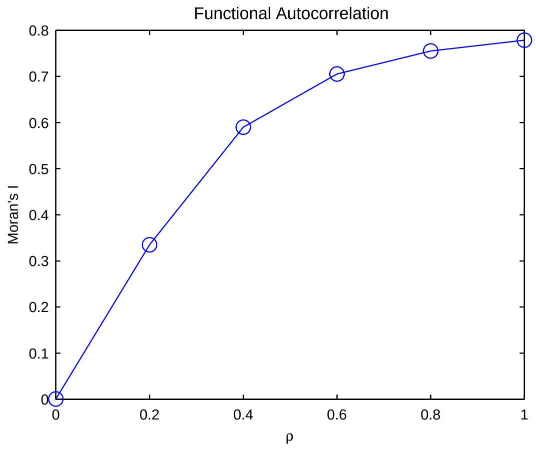


Figure 3.6: Plot of Moran's I versus the correlation parameter ρ for the simulated data. Data with increased functional autocorrelation yield larger values of Moran's I.

$\rho = 0$.

We compute Moran's I for the clustering solution corresponding to each simulated data set, and Figure 3.6 displays the relationship between Moran's I and the correlation parameter for the simulated data. The plot illustrates that Moran's I is effective for detecting and quantifying functional autocorrelation in clustered data. The value of Moran's I is zero in the absence of functional autocorrelation. In practice, the importance of revealing the absence of functional autocorrelations lies in the fact that a simple cluster analysis would yield misleading results by providing a clustering solution when no true underlying structure is present in the data. Moran's I increases as a function of ρ, reinforcing that the global index captures the functional autocorrelation inherent within the computed clusters. The simulated data do not achieve perfect correlation, even when $\rho = 1$, primarily due to the addition of random noise specified by our simulation model. Consequently, Moran's I is approximately equal to 0.8 when the correlation parameter for the simulated data equals 1.

3.5 Discussion

Data-driven approaches, such as cluster analysis and ICA, are extremely useful tools for describing the functional structure and patterns present in neuroimaging data. Cluster analyses help dissociate groups of voxels that exhibit similar patterns of brain activity over time or across experimental conditions. ICA is very useful for detecting resting-state neural networks. In this work, we extend the interpretations given by those two methods by introducing Moran's I for statistical estimation and hypothesis testing of the functional autocorrelation present in neural processing networks as identified by clustering or ICA. Applying these procedures helps to evaluate the validity of the typological descriptions given by a clustering or an ICA solution. Computing Moran's I is fast and easy to program using standard statistical software such as R or SAS. We use MATLAB to calculate the functional autocorrelation of neural processing networks of (functional clusters or ICA components) from PET working memory data, fMRI depression data, and simulated data.

We propose a definition of Moran's I that estimates the functional autocorrelation separately for each element of the summary statistic vector, e.g. pertaining to a specified experimental condition such as high working memory load. For studies that cluster time courses measured under numerous experimental stimuli, it is perhaps more beneficial to aggregate the spatial index across the components of the summary statistic vector. One can easily modify Moran's I to define a more crude measure that "averages" across all elements of the summary vector. When computing the functional network-specific contributions to the global autocorrelation index, the network sizes impact the corresponding contributions. The data must provide extremely strong evidence that a small neural processing network (e.g. a small cluster) contributes substantially to the global autocorrelation for the network to weigh heavily in the global calculation. The number of networks may also affect estimates of the functional autocorrelation. Assigning voxels into a large number of groups will allow the solution to reflect strong resemblances among within-cluster voxels, whereas designating fewer groups will often lead to less coherence within clusters.

We demonstrate the use of Moran's I using PET neuroimaging data analyzed using cluster analysis and fMRI resting-state data analyzed using ICA. In addition, Moran's I may combine with results from related descriptive procedures such as principal component analysis (Friston et al., 1993). Structural equation modeling is a useful approach to examine the direct and indirect effects that one brain region has on another (McIntosh and Gonzalez-Lima, 1994), but it requires the specification of anatomical models and substantially simplifies spatial representations from a large number of voxels to a small number of regions. Both cluster analyses and ICA generally serve as exploratory tools, but often the ultimate objectives of neuroimaging studies require the use of inferential or confirmatory statistical procedures. In this light, conducting such analyses, followed by an assessment of the functional autocorrelation, helps provide insights about characteristics of neuroimaging data that may prove valuable for subsequent modeling, estimation, and hypothesis testing.

Chapter 4

Bayesian Hierarchical Spatial Model for Predicting Brain Activity

4.1 Introduction

Functional neuroimaging (such as PET and fMRI), which permit noninvasive measures of neurophysiology and neuroreceptor binding, are powerful tools which have an increasingly important role in defining the neural basis of illnesses and risk factors for major psychiatric disorders such as depression, schizophrenia and Alzheimer's disease (AD). The clinical capabilities of neuroimaging tools for guiding treatment decisions for such disorders, however, have not been fully established. There is emerging interest in using functional neuroimaging to guide treatment selections for individual patients and to predict the progression of the disease, prompting the need to develop statistical methodology that would provide clinicians with predictive information about patients' brain activity. In treatment-related studies, such methods would assist clinicians in making treatment decisions by forecasting post-treatment neural activity. In studying the progression of dementia, they can identify preclinical changes that may predict, for example, the onset of Alzheimer's disease.

Functional brain imaging has only recently been used to predict brain and clinical outcomes in individual patients. Guo et al. (2008) proposed a predictive statistical model for PET and fMRI data, using a Bayesian hierarchical framework, that uses patient's pretreatment scans, coupled with relevant patient characteristics, to predict brain activity in schizophrenic patients after a specified treatment regimen. This represents an important first step in attempting to help make treatment decisions through using functional neuroimaging data and provides a foundation on which future research, including that on predicting clinical symptom responses, can build. Stonnington et al. (2010) used relevance vector regression (RVR) to predict clinical scores from individual scans. In particular, they used individuals' MRI T1 weighted image to predict their performances on established tests used in the evaluation of Alzheimer's disease, in two independent data sets. Predicted and actual clinical scores were highly correlated. In their analysis they use only the structural MRI, rather than neural activity derived from fMRI or PET scans.

Recently, Gaussian Processes (GPs), based on Bayesian theory, emerged as an alternative to SVM. A Gaussian Process is a generalization of the multivariate Gaussian distribution to infinitely many dimensions, with a constraint that any finite number have a multivariate normal distribution. Marquand et al. (2010) evaluate the predictive capability of GP models for two types of quantitative prediction: multivariate regression and probabilistic classification, using whole-brain fMRI volumes from a study investigating subjective responses to thermal pain. They show that GP models predict subjective pain ratings without requiring anatomical hypotheses about functional localization of relevant brain processes, and that GP predictions were more accurate than any region previously demonstrated to encode pain intensity. They also show that GP regression models outperform support vector and relevance vector regression.

Although each of these methods represent an important contribution to predicting treatment outcome or brain activity based on neuroimaging data, none of the methods use the spatial information from the neighboring voxels/regions to improve the prediction accuracy. In addition to background spatial correlations inherent in neuroimaging data, functional neuroimaging data naturally exhibit spatial correlations due to underlying functional connectivity. We propose a model that borrows strength from such correlations, with the goal of improving prediction.

Bowman et al. (2008) propose a spatial Bayesian hierarchical model for analyzing functional neuroimaging data that provides a unified framework to obtain neuroactivation inferences as well as task-related functional connectivity inferences. The model combines whole-brain voxel-by-voxel modeling and ROI analyses within a unified framework. An unstructured variance/covariance matrix for regional mean parameters allows for the study of inter-regional (long-range) correlations, and the model employs an exchangeable correlation structure to capture intraregional (short-range) correlations. Although this model captures both long-range and short-range correlations, and provides a very useful framework that is applicable for both making inferences regarding task-related changes in brain activity and for identifying prominent task-related connectivity, it does not capture temporal correlations between the brain activity in repeated scanning sessions which would constitute a basis for the prediction.

We propose a novel Bayesian hierarchical framework for predicting follow-up neural activity based on the baseline functional neuroimaging data that attempts to overcome some shortcomings of the modeling methods used in other neuroimaging settings by borrowing strength from the spatial correlations present in the data. The top level of our hierarchical model specifies a multivariate normal likelihood function, where the mean is composed of a population-level mean parameter, a voxel-specific spatial parameter, an individualized (random effect) component and the covariates contribution. The spatial correlations are incorporated in the model in two ways: the short-range correlations between neighboring voxels are incorporated through a multivariate *conditional autoregressive* (CAR) prior of the spatial parameters, while the long-range correlations between anatomical brain regions are incorporated through the covariance matrix of the random effect parameters. Our proposed model is multivariate (bi-variate), allowing modeling correlations between brain activity at different scanning sessions, which are the bases for the prediction of follow-up brain activity for the new patients.

Estimation is performed using Markov Chain Monte Carlo (MCMC) techniques implemented via Gibbs sampling. We apply our Bayesian hierarchical spatial model to a PET data set from a study of Alzheimers disease. This bivariate application is concerned with modeling spatial patterns for bivariate brain activity: the baseline and the 6-month follow-up activity. An association between the baseline and follow-up brain activity for a given individual is certainly expected. Also, spatial associations across the brain locations (both between voxels and between regions) are expected.

Since our model incorporates a conditional autoregressive prior at the second level, we give a brief introduction to CAR and multivariate CAR models here and give references for more details.

4.1.1 Conditional Autoregressive (CAR) Models

Conditional autoregressive (CAR) models were introduced by Besag (1974) and are one of the most important and widely used models to represent spatial correlations (e.g., in disease mapping: Clayton and Kaldor (1987), Marshall (1991), Waller et al. (1997)). They are conveniently implemented using Gibbs samplier and more general MCMC methods for fitting certain classes of hierarchical spatial models.

Let $\boldsymbol{\phi} = (\phi_1, \phi_2, \ldots, \phi_n)$ represent a vector of univariate variables. The zero-centered CAR specification, following Besag (1974), gives

$$\phi_i|\phi_j, j \neq i, \tau_i^2 \sim N\Big(\sum_j b_{ij}\phi_j, \tau_i^2\Big), \quad i = 1, \dots, n. \tag{4.1}$$

Using Brook's Lemma (1964), it can be shown that

$$f(\boldsymbol{\phi}, \boldsymbol{\tau}) \propto \exp\Big\{-\tfrac{1}{2}\boldsymbol{\phi}^T D^{-1}(\mathbf{I} - B)\boldsymbol{\phi}\Big\}, \tag{4.2}$$

where $B = \{b_{ij}\}$ and D is diagonal with $D_{ii} = \tau_i^2$. (4.2) suggests a multivariate normal distribution, with zero mean, but in order for (4.2) to be a proper distribution, $D^{-1}(\mathbf{I}-B)$ must be non-singular and symmetric. This implies the following condition on b_{ij} and τ:

$$\frac{b_{ij}}{\tau_i^2} = \frac{b_{ji}}{\tau_j^2}, \text{ for all } i, j.$$

However, in spatial applications a symmetric proximity matrix W is usually created and one sets $b_{ij} = w_{ij}/w_{i+}$, where $w_{i+} = \sum_j w_{ij}$. If $\tau_i^2 = \sigma^2/w_{i+}$, we define D_W to be diagonal matrix with entries w_{i+} and have $D^{-1}(\mathbf{I} - B) = 1/\sigma^2(D_W - W)$. Since $(D_W - W)\mathbf{1} = 0$, (4.2) is an *improper* distribution. This problem is in practice ignored and the ϕ_i's are sampled using the full conditional distributions in (4.1) with a linear constraint imposed. Since $f(\boldsymbol{\phi}, \tau)$ in (4.2) cannot be used as a model for data (data cannot arise under an improper stochastic mechanism), it is usually attached to random spatial effects introduced at a second stage of a hierarchical specification. Besag et al. (1995) showed that an improper CAR prior can still result in a proper posterior distribution, provided the likelihood is well defined.

A suggested repair for the improperiety problem (Cressie, 1993) is to introduce a parameter ρ into the mean specification which results in the covariance matrix $D^{-1}(\mathbf{I} - \rho B)$ and choose ρ to make it nonsingular. This is guaranteed if $\rho \in (1/\lambda_{(1)}, 1/\lambda_{(n)})$, where $\lambda_{(1)}$ and $\lambda_{(n)}$ are the smallest and the largest eigenvalues respectively of $D_W^{-1/2} W D_W^{-1/2}$. Moreover, it is not difficult to show that $\lambda_{(n)} = 1$ and $\lambda_{(1)} < 0$. Parameter ρ can be interpreted as a coefficient which measures spatial association (see Gelfand and Vounatsou (2003) for an explanation of this interpretation).

Multivariate Conditional Autoregressive (MCAR) *Models.* Most of the applications of CAR modeling are done in the univariate case and employ an improper specification. Gelfand and Vounatsou (2003) move to multivariate conditional autoregressive models and provide rich, flexible classes which yield proper distributions. Their approach is to introduce spatial autoregression parameters. They first clarify what classes can be developed from the family of Mardia (1988) and then propose to employ these models as specifications for second-stage spatial effects in hierarchical models. For more details about the CAR and MCAR models we refer the reader to Carlin and Banerjee (2003), Gelfand and Vounatsou (2003), Banerjee et al. (2004), and Jin et al. (2005). The theoretical work underlying the development of the MCAR models is given in Mardia (1988).

In our model we adopt a MCAR prior with a single (scalar) spatial parameter (in our case, it means we assume the same level of spatial association for the baseline and the follow-up data), denoted in Gelfand and Vounatsou (2003) by MCAR(ρ, Σ).

4.2 Experimental Data

We consider data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) study (`http://www.loni.ucla.edu/ADNI/`). The goal of this national multi-center project is to develop biomarkers of Alzheimer's Disease (AD) in elderly subjects. Study participants receive [^{18}F]-2-fluoro-2-deoxy-2-glucose (FDG) positron emission tomography (PET) scans several times during the study: at baseline (screening), 6 months, 12 months and 24 months. FDG is an analogue of glucose, and in PET it yields concentrations of the injected tracer indicating tissue metabolic activity, in terms of regional glucose uptake. In our analysis, we used the baseline and month 6 scans. For more details about the ADNI, see Mueller et al. (2005).

Participants are classified as mild cognitive impairment (MCI) patients, Alzheimer's disease (AD) patients, or healthy controls (HC). The data from 40 AD and 40 HC subjects were used in the training step of the prediction model development. The prediction model is then applied for predicting the month 6 follow-up PET scan, based on the baseline scan, for an additional group of 33 AD and 33 HC subjects.

The PET scans being used in ADNI measure the brain's rate of glucose metabolism using the tracer [^{18}F] Fluorodeoxyglucose. Since the images are from many different PET scanner models and differ in resolution, orientation, voxel and image dimensions, count statistics, etc., we did not use the raw or "original" PET images, but the already processed PET image data available from the ADNI web site.

The processing steps are as follows. (1) Co-registration: In most cases, six five-minute frames are acquired 30 to 60 minutes post-injection. Each extracted frame is co-registered to the first extracted frame of the raw image file. (2) Averaging: the six five-minute frames of the co-registered image set are averaged to create a single 30 min PET image.(3) Standardizing image and voxel size: Each subject's co-registered, averaged image from their baseline PET scan is reoriented into a standard $160 \times 160 \times 96$ voxel image grid, having 1.5 mm cubic voxels. This standardized image then serves as a reference image for all PET scans on that subject. The individual frames from each PET scan (the baseline study as well as all subsequent studies (6-month scan, 12-month scan, etc.) are co-registered to this baseline reference image. 4) Spatial smoothing: Each image set is filtered with a scanner-specific filter function (can be a non-isotropic filter) to produce images of a uniform isotropic resolution of 8 mm FWHM, the approximate resolution of the lowest resolution scanners used in ADNI. In addition to the above preprocessing steps, we performed a spatial normalization to a standard $91 \times 109 \times 91$ MNI space (Tzourio-Mazoyer et al., 2002).

The following covariates were included in our analysis: the Alzheimers Disease Assessment Scale – cognitive subscale [ADAS Cog] and the subjects's age (years). ADAS was designed to measure the severity of the most important symptoms of AD. Its subscale ADAS-cog is the most popular cognitive testing instrument used in clinical trials of nootropics. It consists of 11 tasks measuring the disturbances of memory, language, praxis, attention and other cognitive abilities which are often referred to as the core symptoms of AD.

4.3 Methodology

In this work, we propose a novel Bayesian hierarchical framework for predicting follow-up (or post-treatment) neural activity based on the baseline (or pre-treatment) functional neuroimaging data that attempts to overcome some shortcomings of the presently used modeling methods by borrowing strength from the spatial correlations present in the data (both local, between-voxel, correlations and more long-range, between-region correlations). The proposed model builds on the proper multivariate conditional autoregressive model (MCAR(ρ, Σ)) proposed in Gelfand and Vounatsou (2003). Our proposed prediction algorithm is similar to the predictive method proposed in Guo et al. (2008), but our proposed model is incorporating spatial information, in addition to capturing correlations between the repeated scans. Also, Guo et al. (2008) use the expectation-maximization (EM) algorithm to estimate the model parameters, while we use the MCMC approach and the Gibbs sampling technique. Our model can also be seen as an extension of a hierarchical model for functional neuroimaging data proposed by Bowman et al. (2008). Using our proposed method, we analyze the PET data from a study of Alzheimer's disease described in Section 4.2.

4.3.1 Model and Estimation

We consider an anatomical parcellation of the brain consisting of $g = 1, \ldots, G$ regions, where we may set G to be as high as 116 (Tzourio-Mazoyer et al., 2002). Alternative anatomical parcellations are also available, such as those based on Brodmann regions (Brodmann, 1909b). Let $i = 1, \ldots, n$ denote subjects, $v = 1, \ldots, V$ voxels, and let V_g represent the number of voxels in a particular region indexed by g. We denote the regional glucose use, as a proxy for brain activity at voxel v, by $Y(v)$. In our experimental data (see Section 4.2), $Y(v)$ represents an average over several scans that were acquired on a same subject during one scanning session. Let $\mathbf{Y}_{ig}(v) = \Big(Y_{ig}(v)^{(1)}, Y_{ig}(v)^{(2)}\Big)^T$, where superscripts (1) and (2) denote the baseline and follow-up scans, respectively. Before processing, a local neighborhood of each voxel included in the analysis must be selected, for the voxel-level spatial modeling. We apply a three-dimensional (3D) neighborhood structure containing the 26 immediate neighboring voxels: 9 above, 9 below, and 8 adjacent. In addition, we consider only within-region neighbors. For each voxel included in the analysis, the neighbors are identified and this information is saved in a connectivity matrix W, before the estimation procedure is performed.

We propose a multivariate Bayesian hierarchical model that accounts for both spatial correlations between intra-regional voxels, and between regions. The model also accounts for correlations between baseline and follow-up regional glucose use. In particular, our proposed model has the following hierarchical structure:

$$
\begin{aligned}
\mathbf{Y}_{ig}(v)|\boldsymbol{\beta}_g, \boldsymbol{\phi}_g, \boldsymbol{\alpha}_{ig}, \boldsymbol{\gamma}_{gv}, \mathbf{Z}_g &\sim \mathrm{N}(\boldsymbol{\beta}_g(v) + \boldsymbol{\phi}_g(v) + \boldsymbol{\alpha}_{ig} + \mathbf{X}_{ig}\boldsymbol{\gamma}_g, \mathbf{Z}_g) \\
\boldsymbol{\phi}_v|\boldsymbol{\phi}_{v'}, v \neq v', \Sigma, v = 1\ldots, V &\sim \mathrm{N}\big(\rho \sum \tfrac{w_{vv'}}{w_{v+}} \mathbf{I}\,\boldsymbol{\phi}_{v'}, \tfrac{1}{w_{v+}}\Sigma\big) \quad \big(\mathrm{MCAR}(\rho, \Sigma)\big) \\
\boldsymbol{\beta}_{gj}|\lambda_{gj}^2 &\sim \mathrm{N}(\boldsymbol{\beta}_{0gj}, \lambda_{gj}^2 \mathbf{I}) \quad (\lambda_{vgj} = \lambda_{gj}, \forall v \in \text{region } g) \\
\mathbf{Z}_g^{-1} &\sim \mathrm{Wishart}\big((c_1\boldsymbol{\Omega}_1)^{-1}, c_1\big) \\
\Sigma^{-1} &\sim \mathrm{Wishart}\big((c_2\boldsymbol{\Omega}_2)^{-1}, c_2\big) \\
\boldsymbol{\alpha}_{ij}|\boldsymbol{\Gamma}_j &\sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Gamma}_j) \quad (\boldsymbol{\alpha}_{ij} = \boldsymbol{\alpha}_i^{(j)}) \\
(\boldsymbol{\Gamma}_j)^{-1} &\sim \mathrm{Wishart}\{(h_j H_j)^{-1}, h_j\} \;\; j = 1, 2 \\
\lambda_{gj}^{-2} &\sim \mathrm{Gamma}(a_j, b_j) \\
\gamma_{gjq}|\tau_{gjq}^2 &\sim \mathrm{N}(0, \tau_{gjq}^2) \quad q = 1, \ldots, Q \text{ (covariates)} \\
\tau_{gjq}^{-2} &\sim \mathrm{Gamma}(e_0, f_0) \\
\rho &\sim \mathrm{Uniform}(\{0, 0.05, 0.1, \ldots, 0.8, 0.81, \ldots, \\
&\qquad 0.9, 0.91, \ldots, 0.99\})
\end{aligned}
\tag{4.3}
$$

where j denotes the scanning session (i.e., $j = 1$ – baseline and $j = 2$ – follow-up).

For each voxel v, the subject-specific quantities $\mathbf{Y}_{ig}(v)$ are assumed to vary randomly about a mean determined by a population-level mean parameter $\boldsymbol{\beta}_g(v) = \big(\boldsymbol{\beta}_g(v)^{(1)}, \boldsymbol{\beta}_g(v)^{(2)}\big)^T$, a population-level spatial dependence parameter $\boldsymbol{\phi}_g(v) = \big(\phi_g(v)^{(1)}, \phi_g(v)^{(2)}\big)^T$, an individualized random effect component $\boldsymbol{\alpha}_{ig} = \big(\alpha_{ig}^{(1)}, \alpha_{ig}^{(2)}\big)^T$ and the covariates' parameters $\boldsymbol{\gamma}_g = \big(\gamma_{g1}^{(1)}, \gamma_{g1}^{(2)}, \ldots, \gamma_{gQ}^{(1)}, \gamma_{gQ}^{(2)}\big)^T$. $\mathbf{X}_{ig} = \begin{bmatrix} X_{ig1}^{(1)} & 0 & \cdots & X_{igQ}^{(1)} & 0 \\ 0 & X_{ig1}^{(2)} & \cdots & 0 & X_{igQ}^{(2)} \end{bmatrix}$ is a $2 \times 2Q$ matrix containing the mean-centered subject-specific covariates, where Q is the number of covariates in the model. The random effects $\boldsymbol{\alpha}_i^{(j)} = \big(\alpha_{i1}^{(j)}, \ldots, \alpha_{iG}^{(j)}\big)$, where j denotes the baseline and follow-up sessions, respectively, and $\boldsymbol{\alpha}_i = \big(\boldsymbol{\alpha}_{i1}^T, \ldots, \boldsymbol{\alpha}_{iG}^T\big)^T$. Given $\{\boldsymbol{\beta}_g\}$, $\{\boldsymbol{\phi}_g\}$, $\{\boldsymbol{\alpha}_{ig}\}$, $\{\boldsymbol{\gamma}_{gv}\}$ and $\mathbf{Z}_g$, the $\{\mathbf{Y}_{ig}\}$ are conditionally independent.

At the second level, the model expresses a prior belief that each voxels population mean (for the j-th session) arises from a normal distribution with a mean given by the overall region mean $\boldsymbol{\beta}_{0gj}$ and variance λ_{gj}^2. It represents a reasonable starting point to assume that voxels within anatomically defined regions exhibit brain activity that deviates around an overall mean for that region.

Spatial associations are introduced through random effects in the mean structure of the data. Bivariate spatial random effects at the voxel level call for a multivariate CAR (MCAR) specification, where ρ is a scalar parameter representing the overall degree of spatial dependence and Σ is the covariance matrix between $\boldsymbol{\phi}_v$ and $\boldsymbol{\phi}_{v'}$. The connectivity (proximity) matrix W used in the MCAR prior for the spatial random effects $\boldsymbol{\phi}$ is based on the 3D neighborhood structure described above. $W = \{w_{vv'}\}_{(V \times V)}$ is a symmetric matrix (i.e. $w_{vv'} = 1$ if voxel v' is in the defined 3D neighborhood of v), and w_{v+} is the sum of the elements in row v of W.

Parameter ρ determines the magnitude of the spatial neighborhood effect. We follow the suggestion in Gelfand and Vounatsou (2003) and specify a discrete uniform prior for the spatial autoregression parameter ρ. The authors suggest that "...discretization provides the most computationally feasible approach to update [ρ], avoiding Metropolis steps". Following the criteria in Gelfand and Vounatsou (2003), we assume $\rho < 1$ to ensure propriety, we do not allow $\rho < 0$ since this would violate the similarity of spatial neighbors which we seek, and we place prior mass which favors the upper range of ρ "...since even moderate spatial dependence requires values of ρ near 1 (see the reply to the discussion in Besag et al. (1991))". In particular, similar to Gelfand and Vounatsou (2003), we put equal mass on the following 36 values: 0, 0.05, 0.1, ..., 0.8, 0.81 0.82, ..., 0.90, 0.91, 0.92, ..., 0.99. We provide additional details regarding CAR and MCAR models in the Section 4.1.1.

The model also captures potential functional connections between anatomical brain regions through the covariance matrix $\boldsymbol{\Gamma}_j$. We specify a conjugate Wishart prior for $\mathbf{\Gamma}_j$, allowing a flexible unstructured covariance matrix. The multivariate structure enables us to compute the between-region connectivity matrix, separately for each session. Finally, hyperpriors on the MCAR parameters and other hyperparameters complete the model.

Using the vector notation, Kronecker product, and some simple algebraic manipulations, the model can we written at a region level. Some additional notation is introduced and region-level model is given in Appendix 5.2.1.

Estimation. Estimation is performed using Markov Chain Monte Carlo (MCMC) techniques implemented via Gibbs sampling. The Gibbs sampling algorithm is an iterative Markov Chain Monte Carlo algorithm to sample from the joint distribution of a vector of random variables, when only conditional distributions are available to sample from. To use the Gibbs sampling algorithm, we need the full conditional distributions of each of the parameters and random effects, given the rest.

Applying MCMC methods in our context is complicated by the massive amount of data, the large number of spatial locations, and the large number of parameters that need to be estimated. The Gibbs-friendly model specification facilitates estimation by providing substantial reductions in computing time and memory. We present the full conditionals required to run the Gibbs sampler below.

Iterating through simulations from the full conditionals, updated with the most recent simulated parameters, results in (following a burn-in period) draws from the joint posterior distribution of the parameters.

4.3.2 Full Conditional Distributions

Joint Posterior Distribution. Full conditional distributions are derived from the joint distribution of the variables. The joint conditional distribution of all the variables in our model can be represented by the following expression (we drop the variable indices for simplicity)

$$[\boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\alpha}, \boldsymbol{\gamma}, \mathbf{V}, \Sigma, \boldsymbol{\Gamma}, \boldsymbol{\lambda}, \boldsymbol{\tau}, \rho] \propto [\boldsymbol{\Gamma}]\,[\boldsymbol{\lambda}]\,[\boldsymbol{\tau}]\,[\rho]\,[\mathbf{V}]\,[\Sigma]\,[\boldsymbol{\beta}|\boldsymbol{\lambda}]\,[\boldsymbol{\alpha}|\boldsymbol{\Gamma}]\,[\boldsymbol{\phi}|\rho,\Sigma]\,[\boldsymbol{\gamma}|\boldsymbol{\tau}] \cdot [\mathbf{Y}|\boldsymbol{\beta}, \mathbf{V}, \boldsymbol{\phi}, \boldsymbol{\alpha}, \boldsymbol{\gamma}]$$

Full Conditional Expressions. Here we give the expressions for the full conditional distributions of the parameters in the model. We note that the expressions are written at a region level (i.e. for $\boldsymbol{\beta}_g$, instead of $\boldsymbol{\beta}_g(v)$). The details on how the formulas for the full conditionals are derived are given in Appendix 5.2.2. Some computational details and simplifications are also given in Appendix 5.2.2. The full conditionals are given by the following (we omit notation for conditioned variables for simplicity):

1. $\boldsymbol{\beta}_g \sim \mathrm{N}(\boldsymbol{\Omega}_g\mathbf{T}_g, \boldsymbol{\Omega}_g)$, where $\boldsymbol{\Omega}_g = \mathbf{I}_{V_g} \otimes (\Lambda_v^{-1} + n\mathbf{Z}_g^{-1})^{-1}$, $\mathbf{T}_g = \big(\mathbf{I}_{V_g} \otimes \Lambda_v^{-1}\big)\boldsymbol{\beta}_{0g} + \mathbf{V}_g^{-1}\sum_{i=1}^{n} \mathbf{r}_{ig}$, $\Lambda_v^{-1} = \begin{bmatrix} 1/\lambda_1^2 & 0 \\ 0 & 1/\lambda_2^2 \end{bmatrix}$, and $\mathbf{r}_{ig} = \mathbf{Y}_{ig} - \boldsymbol{\phi}_g - 1 \otimes \boldsymbol{\alpha}_{ig} - 1 \otimes \mathbf{X}_{ig}\boldsymbol{\gamma}_g$.

2. The full conditional distribution for $\boldsymbol{\phi}_v$ is a likelihood adjusted version of the conditional distribution for MCAR(ρ, Σ).

   $\boldsymbol{\phi}_v \sim \mathrm{N}(\mathbf{H}_v\mathbf{P}_v, \mathbf{H}_v)$, where $\mathbf{H}_v^{-1} = w_{v+}\Sigma^{-1} + n\mathbf{Z}_g^{-1}$, $\mathbf{P}_v = w_{v+}\rho\Sigma^{-1}\sum\limits_{v' \neq v} \frac{w_{vv'}}{w_{v+}}\boldsymbol{\phi}_{v'} + n\mathbf{Z}_g^{-1}(\bar{\mathbf{Y}}_g(v) - \boldsymbol{\beta}_g(v) - \bar{\boldsymbol{\alpha}}_g - \bar{\mathbf{X}}_g\boldsymbol{\gamma}_g)\big)$, and $\bar{\mathbf{Y}}_g(v) = \frac{1}{n}\sum\limits_{i=1}^{n}\mathbf{Y}_{ig}(v)$, $\bar{\boldsymbol{\alpha}}_g = \frac{1}{n}\sum\limits_{i=1}^{n}\boldsymbol{\alpha}_{ig}$, $\bar{\mathbf{X}}_g = \frac{1}{n}\sum\limits_{i=1}^{n}\mathbf{X}_{ig}$.

3. $\boldsymbol{\alpha}_i \sim \mathrm{N}(\boldsymbol{\Psi}\mathbf{P}, \boldsymbol{\Psi})$, where $\boldsymbol{\Psi} = \big(\boldsymbol{\Gamma}^{-1} + \boldsymbol{\Omega}_\alpha^{-1}\big)^{-1}$,

   $$\mathbf{P} = \begin{bmatrix} 1_{V_1} \otimes \mathbf{Z}_1^{-1} & \dots & \dots \\ \vdots & \ddots & \vdots \\ \dots & \dots & 1_{V_G} \otimes \mathbf{Z}_G^{-1} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{Y}_{i1} - \boldsymbol{\beta}_1 - \boldsymbol{\phi}_1 - 1_{V_1} \otimes \mathbf{X}_{i1}\boldsymbol{\gamma}_1 \\ \vdots \\ \mathbf{Y}_{iG} - \boldsymbol{\beta}_G - \boldsymbol{\phi}_G - 1_{V_G} \otimes \mathbf{X}_{iG}\boldsymbol{\gamma}_G \end{bmatrix}, \text{ and}$$

   $$\boldsymbol{\Gamma}^{-1} = \boldsymbol{\Gamma}_1^{-1} \otimes \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \boldsymbol{\Gamma}_2^{-1} \otimes \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \text{ while } \boldsymbol{\Omega}_\alpha^{-1} = \begin{bmatrix} V_1\mathbf{Z}_1^{-1} & \dots & \dots \\ \vdots & \ddots & \vdots \\ \dots & \dots & V_G\mathbf{Z}_G^{-1} \end{bmatrix}.$$

4. $\boldsymbol{\Gamma}_j^{-1} \sim \text{Wishart}\big((h_j H_j + \sum_{i=1}^n \boldsymbol{\alpha}_{ij}\boldsymbol{\alpha}_{ij}^T)^{-1}, h_j + n\big)$.

5. $\mathbf{Z}_g^{-1} \propto \text{Wishart}\big((c_1\Omega_1 + \sum_{i=1}^n \sum_{v\in g} \mathbf{u}_{igv}\mathbf{u}_{igv}^T)^{-1}, c_1 + nV_g\big)$, where $\mathbf{u}_{igv} = \mathbf{Y}_{ig}(v) - \boldsymbol{\beta}_g(v) - \boldsymbol{\phi}_g(v) - \boldsymbol{\alpha}_{ig} - \mathbf{X}_{ig}\boldsymbol{\gamma}_g$.

6. $\boldsymbol{\Sigma}^{-1} \sim \text{Wishart}\big((c_2\Omega_2 + \sum_{v=1}^V w_{v+}[\boldsymbol{\phi}_v - \rho\sum_{v'\neq v} b_{v'}\boldsymbol{\phi}_{v'}][\boldsymbol{\phi}_v - \rho\sum_{v'\neq v} b_{v'}\boldsymbol{\phi}_{v'}]^T), c_2 + V\big)$, where V is the total number of voxels included in the analysis (with at least one within-region neighbor).

7. $\lambda_{gj}^{-2} \sim \text{Gamma}\Big(a_j + V_g/4, \Big[\frac{1}{b_j} + \frac{(\boldsymbol{\beta}_{gj} - \mathbf{1}_{V_g}\boldsymbol{\beta}_{0gj})^T(\boldsymbol{\beta}_{gj} - \mathbf{1}_{V_g}\boldsymbol{\beta}_{0gj})}{2}\Big]^{-1}\Big)$, where $\boldsymbol{\beta}_{gj} = \boldsymbol{\beta}_g^{(j)}$ and $\boldsymbol{\beta}_{0gj} = \boldsymbol{\beta}_{0g}^{(j)}$, for $j = 1, 2$.

8. For the spatial parameter ρ, we set an array of values and take a uniform prior on that set. The full conditional distribution of the spatial parameter ρ is just the corresponding set of normalized likelihood weights. More details are given in Appendix 5.2.2.

9. $\boldsymbol{\gamma}_g \sim \text{N}(\boldsymbol{\Omega}_{\gamma_g}\mathbf{T}_{\gamma_g}, \boldsymbol{\Omega}_{\gamma_g})$, where

   $\boldsymbol{\Omega}_{\gamma_g}^{-1} = \boldsymbol{\Lambda}_{\gamma_g}^{-1} + \sum_{i=1}^n (\mathbf{1}_{V_g}^T \otimes \mathbf{X}_{ig}^T)\mathbf{V}_g^{-1}(\mathbf{1}_{V_g} \otimes \mathbf{X}_{ig})$, $\mathbf{T}_{\gamma_g} = \sum_{i=1}^n (\mathbf{1}_{V_g}^T \otimes \mathbf{X}_{ig}^T)\mathbf{V}_g^{-1}\mathbf{u}_{ig}$,
   $\mathbf{u}_{ig} = \mathbf{Y}_{ig} - \boldsymbol{\beta}_g - \boldsymbol{\phi}_g - \mathbf{1}_{V_g} \otimes \boldsymbol{\alpha}_{ig}$ and
   $\boldsymbol{\gamma}_g|\Lambda_{\gamma_g} \sim \text{N}\big(\mathbf{0}, \underbrace{\text{diag}(\tau_{g11}^2, \tau_{g21}^2, \ldots, \tau_{g1Q}^2, \tau_{g2Q}^2)}_{\boldsymbol{\Lambda}_{\gamma_g}}\big)$.

10. $\tau_{gjq}^2 \sim \text{Gamma}\Big(e_0 + \frac{1}{2}, \big[\frac{1}{f_0} + \frac{1}{2}\boldsymbol{\gamma}_{gjq}^2\big]^{-1}\Big)$.

4.3.3 Prediction

A key advantage of the Bayesian modeling framework, relative to approaches employing classical statistical methods, is that we obtain samples from the joint posterior distribution for all of the model parameters. A related advantage is that we can easily estimate posterior distributions of functions of the model parameters using the posterior samples.

After obtaining the estimates of the parameters from our model, we can use the summary statistics (e.g. mean) of the posterior samples of the estimates to predict the session 2 (month 6) scans for the future patients, based on their baseline scans. The prediction proceeds as follows (following the procedure described in Gelman et al. (2004), Chapter 4). We first draw the covariance matrix $\mathbf{\Gamma}$ from its full conditional distribution. We can then draw $\boldsymbol{\alpha}_i$'s, given $\mathbf{\Gamma}$, from the prior distribution for $\boldsymbol{\alpha}$. If $\mathbf{Y}_{i^*1}$ denotes the baseline scan and $\mathbf{Y}_{i^*2}$ the month 6 scan for a future subject, we now have all the necessary information (parameter estimates for $\boldsymbol{\beta}$, $\boldsymbol{\phi}$, $\mathbf{\Gamma}$, $\boldsymbol{\gamma}$, $\mathbf{Z}$, $\boldsymbol{\alpha}$ and the covariate matrix $\mathbf{X}$) to estimate $\mathbf{Y}_{i^*2}|\mathbf{Y}_{i^*1}$.

At the region level, we can re-organize the data and the parameters as follows. We can write $\mathbf{Y} = (\mathbf{Y}_1^T, \mathbf{Y}_2^T)^T$, where $\mathbf{Y}_1$ contains all the session 1 voxel values, and $\mathbf{Y}_2$ all the session 2 voxel values. For region g, then $\mathbf{Y}_g = (\mathbf{Y}_{g,1}^T, \mathbf{Y}_{g,2}^T)^T \sim \mathrm{N}\big((\boldsymbol{\mu}_{g,1}^T, \boldsymbol{\mu}_{g,2}^T)^T, \mathbf{\Sigma}_g\big)$, where $\mathbf{\Sigma}_g = \mathbf{Z}_g \otimes \mathbf{I}_{V_g}$.

From model (4.3), it follows that $\mathbf{Y}_{i^*,2}|\mathbf{Y}_{i^*,1} \sim \mathrm{N}(\mathbf{b}_{i^*g}, \mathbf{A}_{i^*g})$. For region g, $\mathbf{b}_{i^*g} = \boldsymbol{\mu}_{i^*g,2} + \mathbf{\Sigma}_{12}^T \mathbf{\Sigma}_{11}^{-1} (\mathbf{Y}_{i^*g,1} - \boldsymbol{\mu}_{i^*g,1})$, and $\boldsymbol{\mu}_{i^*g} = \boldsymbol{\beta}_g + \boldsymbol{\phi}_g + \mathbf{1}_{V_g} \otimes \boldsymbol{\alpha}_{i^*g} + \mathbf{1}_{V_g} \otimes \mathbf{X}_{i^*gv} \boldsymbol{\gamma}_{gv}$. $\boldsymbol{\mu}_{i^*g,2}$ is a vector consisting of even elements of $\boldsymbol{\mu}_{i^*g}$. If we write $\mathbf{Z}_g = \begin{bmatrix} a & c \\ c & b \end{bmatrix}$, then it is easy to see that $\mathbf{A}_{i^*g} = b\mathbf{I}_{V_g} - c\mathbf{I}_{V_g} \cdot \frac{1}{a}\mathbf{I}_{V_g} \cdot c\mathbf{I}_{V_g} = (b - \frac{c^2}{a})\mathbf{I}_{V_g}$.

By inputting the posterior mean of the parameters obtained from the MCMC estimation, we obtain the estimated conditional mean $\hat{\mathbf{b}}_{i^*g}$ and covariance matrix $\hat{\mathbf{A}}_{i^*g}$. The follow-up regional glucose uptake $\mathbf{Y}_{g,2}$ are predicted using the mean of the estimated conditional distribution, i.e. $\hat{\mathbf{b}}_{i^*g}$. Also, a 100(1-α)% prediction interval for $\mathbf{Y}_{g,2}$ can be constructed based on the estimated conditional variance.

4.3.4 Model Validation: Estimation of the prediction error.

To estimate the accuracy of our proposed prediction model, applied to the ADNI data, we estimate the prediction error for the test subjects' data. For each of the two groups (AD and HC), we applied our algorithm to predict the follow-up regional glucose uptake for 33 new (test) subjects. We used the estimated parameters from the training data.

We estimate the prediction error by comparing the observed and predicted follow-up (6 month) brain activity. To evaluate the predicted error, typically the squared error and absolute error functions are used to define the loss functions. These functions are based on the absolute difference between the observed and predicted values, and hence depend on the magnitude of the brain activity. The mean squared error of prediction of a model, denoted PMSE($\hat{p}$) is defined as $E[(y - f(X, \hat{p}))^2 | \hat{p}]$, where y is the quantity to be predicted and $f(X, \hat{p})$ is the prediction given by the model f. In neuroimaging, prediction of the brain activity is performed on each voxel, and the prediction error is evaluated across all voxels. The square or absolute error functions are inappropriate in functional imaging, however, since the brain activity measurement (such as regional glucose uptake, or BOLD signal) have different baseline values across the brain. For that reason we propose a use of a scale-free loss function, so that the prediction error is comparable across all brain voxels. The proposed function is the ratio of the square root of the PMSE and the average effects at voxel v. This function measures the magnitude of the PMSE relative to the average brain activity at each voxel. For our data, this function (we will refer to it as the "standardized square root of the prediction mean squared error – stPMSE") is defined as

$$\mathrm{stPMSE}\big(\{Y_i^{(2)}(v)\}, \{\hat{Y}_i^{(2)}(v)\}\big) = \frac{\sqrt{(1/N \sum_{i=1}^{N} [\hat{Y}_i^{(2)}(v) - Y_i^{(2)}(v)]^2}}{1/N \sum_{i=1}^{N} Y_i^{(2)}(v)} \tag{4.4}$$

This measure of prediction error is similar to one of the two measures proposed in Guo et al. (2008), but we divide by an average of the observed, not the predicted values at voxel v. We think this is more reasonable since otherwise predicted values can have undue influence on the measure of performance. For example, if the predicted values are overestimated, this could results in misleadingly small stPMSE.

4.4 Results

We apply our Bayesian spatial hierarchical model to PET data from the study of Alzheimers disease (ADNI). We considered 39 regions relevant in AD in our analysis. Those were selected from a set of 46 regions of interest considered in the Alzheimer's disease study in Bowman et al. (2008). We excluded some AAL (Automated Anatomical Labeling, Tzourio-Mazoyer et al. (2002)) regions which we think are less relevant in AD, an included several additional areas (e.g. left and right posterior cingulum). The complete list of the regions with region sizes is given in Table 5.2.3, Appendix 5.2.3. Areas of the temporal and limbic lobes are of particular interest, having been indicated as having either increased or decreased encoding activity between at-risk subjects and controls in the first wave of the Alzheimers data or being thought to be involved with the (verbal memory) paradigm (Bassett et al., 2006). We do not perform the whole brain (e.g. including all 116 AAL regions) analysis for two reasons. One reason is that the estimation time increases significantly with the number of regions (we comment on the computation cost later in this section). The other reason is that the number of regions in the analysis needs to be smaller than the number of subjects in the smallest group (in our case, 40) in order to have a stability of estimation of the between-region covariance parameters.

Our model captures the short-range spatial correlation between voxels within a defined anatomical region as well as the (potentially) long-range inter-regional connectivity, stemming from the covariance matrix $\mathbf{\Gamma}_j$, of the random effect parameters $\boldsymbol{\alpha}_{ij}$. We specify a Wishart prior for $\boldsymbol{\alpha}_{ij}^{-1}$.

For the inverse-Wishart prior, the degrees of freedom must satisfy $h_0 > G$ to yield a proper prior distribution. This prior becomes more diffuse as h_0 gets smaller (West and Harrison, 1989); hence, we set $h_0 = G$ to reflect the most diffuse proper prior that our data can support. A seemingly natural choice for H_{0j} is a point estimate of $\mathbf{\Gamma}_j$. We use the sample covariance matrix to obtain H_{0j}, calculated from the subject-specific mean activity levels in each of the anatomical regions. In addition, we examine the sensitivity of our results to the sample covariance matrix by artificially reducing the correlations (covariances) using

$$H^*_{0j} = (1 - w)H_{0j} + w\{diag(H_{0j})\}.$$

We also use inverse Wishart priors for Σ and $\mathbf{Z}_g$, i.e. $\Sigma^{-1} \sim \text{Wishart}\big((c_2\mathbf{\Omega}_2)^{-1}, c_2\big)$, $\mathbf{Z}_g^{-1} \sim \text{Wishart}\big((c_1\mathbf{\Omega}_1)^{-1}, c_1\big)$, where $\mathbf{\Omega}_1$ and $\mathbf{\Omega}_2$ are 2×2 matrices and c_1 and c_2 are shape parameters. Since we have no prior knowledge regarding the nature and extent of dependence, we choose $\mathbf{\Omega}_1$ and $\mathbf{\Omega}_2$ to be diagonal. The data will inform about dependence a posteriori. We set $c_1 = c_2 = 4$ to provide low precision for these priors.

To complete our Bayesian hierarchical model, we set $a_1 = a_2 = 0.1$, $b1 = 0.005$, $b_2 = 0.001$, and $h_0 = G$, resulting in vague or weakly informative priors, to ensure that the information in the data primarily governs the results. However, more informative priors may be employed when fairly precise information is available.

To estimate the model parameters, we performed 3,000 iterations with burn-in of 2,000 iterations and thinning of 5 iterations (for storage and computation time). The programming was done in Matlab. The estimation was performed on a Linux cluster. The test/experiment environment consisted of an 8-core system with 16GB of RAM. Each MATLAB job was given a dedicated CPU and ran for approximately 26 hours. Future optimizations include taking advantage of the multi-threaded nature of several MATLAB image analysis functions or using the MATLAB Distributed Computing Environment to use multiple cores per job for speed improvements.

Due to a large number of parameters in our model that need to be estimated, it is impractical (if not impossible) to monitor the trace plots for each of the parameters. Also, to determine the stopping time, running parallel chains is recommended to monitor how the chains mix. Again, for the computational reasons it was not practical in our case. We did, however, obtained trace plots for a number of randomly selected voxels in several regions, for each of the voxel-level parameters. (Several plots and histograms are given in Appendix 5.2.3). The chains obtained by Gibbs sampler (trace plots) are satisfactory, i.e., the generated Markov chains seem to converge, after the burn-in period, to our distributions of interest. We also monitored the trace plots for other (region-specific and scalar) parameters and found them to be satisfactory, considering relatively small number of iterations we performed. For more details on convergence, see Appendix 5.2.3.

4.4.1 Prediction of brain activity for PET data from a study of Alzheimer's disease (ADNI)

Using the estimates from the model, we apply the proposed prediction algorithm to forecast the follow-up regional glucose uptake for each of the subjects in the test data set, which consists of 33 subjects per each group. Our proposed prediction algorithm provides individualized predictions of the regional glucose uptake, based on the unique information in each subject's baseline scan, and the relevant personal characteristics (e.g., ADAS-cog score). For the prediction, we estimate the covariates at the second scanning sessions by the mean of the second session in the training data. We note that, for both groups (AD and HC), the posterior mean of the spatial effect ρ was 0.99, for both scanning sessions data, indicating very strong spatial dependence. Using two separate spatial effect parameters could be considered in future analyses to see whether these differ for different sessions. Also, estimated covariance matrices between the baseline and follow-up regional glucose uptake at a region level, $\mathbf{Z}_g$, yield high between-session correlations (> 0.7 for almost all regions) at a region level.

4.4.1.1 Individualized prediction maps

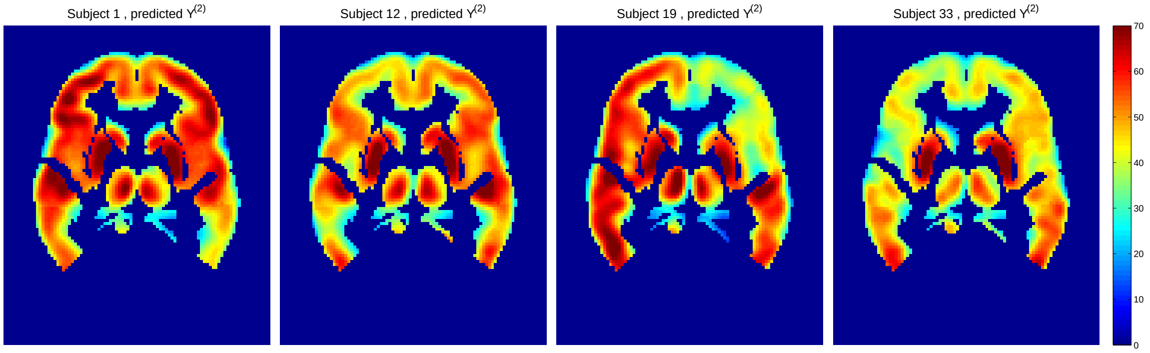
Figure 4.1 shows the individual prediction maps for of the regional glucose uptake at 6 month follow-up for AD patients. Here we only show 4 subjects' maps, for brevity. The prediction maps for several other subjects are given in Appendix 5.2.3. We can see that there are notable differences between subject in the predicted follow-up activity, indicating that possibly different stages of the disease are present in those individuals. We compare the individual predicted maps (Figure 4.1 (a))to the observed maps (Figure 4.1 (b)) and notice satisfactory agreement between the observed and predicted brain activity for these 4 subjects. Similar correspondence is observed for most of the subjects. The individual prediction maps, such as those in Figure 4.1 (a), highlight possible individual differences in the progression of the disease.

We also predict the follow-up brain activity for the healthy controls test data set, which consisted of 33 subjects as well. The results for 4 selected subject are given in Figure 4.2. Similar satisfactory agreement is observed between the predicted and observed maps. We also notice that this groups exhibits smaller between-subject differences, compared to the AD group.

4.4.2 Comparisons with competing prediction models

We compare our prediction results with the results obtained using two proposed competing methods: the General Linear Model (GLM) based method, and the method based on the Bayesian spatial hierarchical model for activation and connectivity analysis (BSMac) proposed in Bowman et al. (2008). We applied both of those methods to the same experimental data set from a study of Alzheimer's disease. First we give some details on how the prediction is performed based on those two methods, and then we compare the obtained result with the results obtained applying our method. To evaluate prediction accuracy, we computed the standardized root mean square er-

(a) predicted month 6 maps



(b) observed month 6 maps

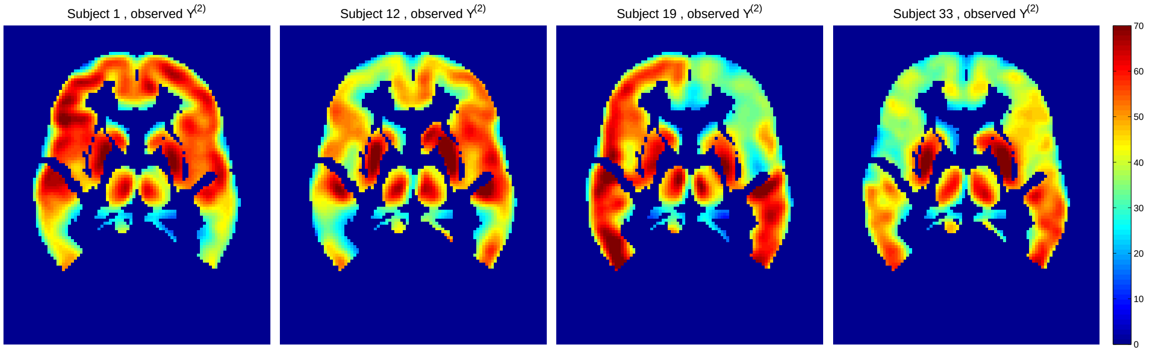


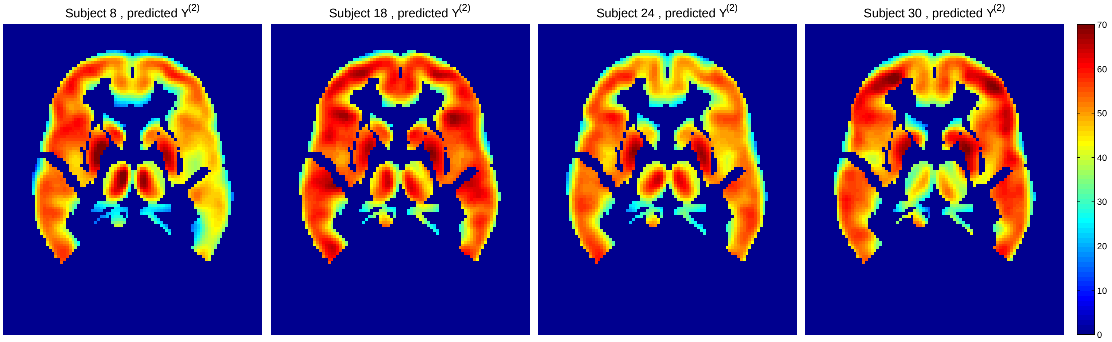
Figure 4.1: Individualized predicted and observed 6 month follow-up regional glucose uptake measurements for 4 AD patients from the test data set. Axial slice 40 is shown in radiological view. There is a satisfactory agreement between the observed and predicted post-treatment regional glucose uptake.

ror (described above) between the predicted regional glucose uptake and its observed value for 33 new subjects in each group (AD and HC).

Comparison with predictions based on the GLM. A similar prediction algorithm idea to the one we propose can be applied to develop a prediction algorithm based on a GLM. The GLM models the brain activity for all subjects using common population parameters. Independence and sphericity between scans at baseline and at follow-up scans are assumed.

Estimates from the GLM are obtained using OLS. The predicted follow-up brain activity reflects only the population-level expectation and does not take into account the information from the subject's pretreatment scans. That is, $\mathbf{Y}^*_{g,2} = \mathbf{X}^*_{g,2} \cdot \hat{\boldsymbol{\beta}}_g$, where subscript 2 denotes session 2 (follow-up) brain activity at region g, superscript *

(a) predicted month 6 maps



(b) observed month 6 maps

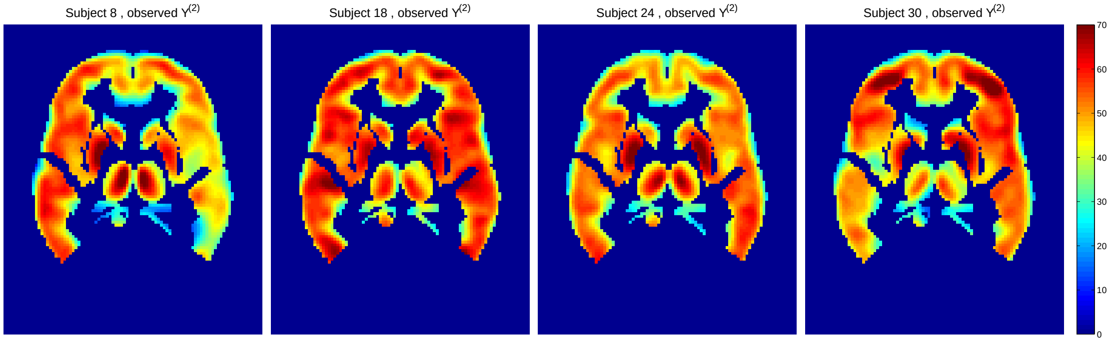


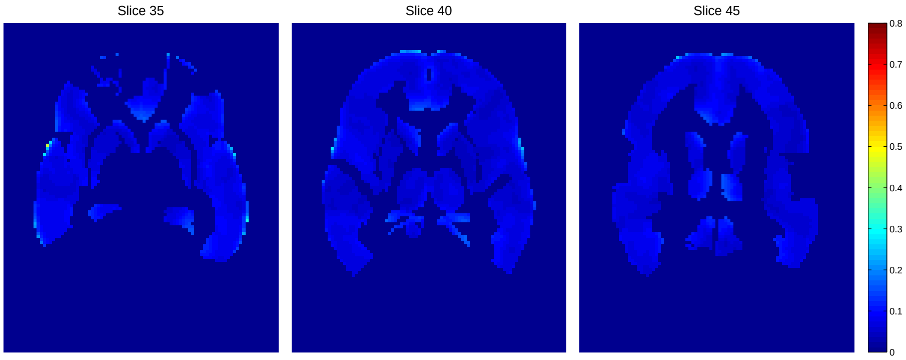
Figure 4.2: Individualized predicted and observed 6 month follow-up regional glucose uptake measurements for 4 HC subjects from the test data set; axial slice 40 is shown in radiological view. There is a satisfactory agreement between the observed and predicted post-treatment regional glucose uptake.

denotes the new subject, and $\hat{\boldsymbol{\beta}}_g$ denotes the estimate of the session 2 mean parameter obtained from the training data set.

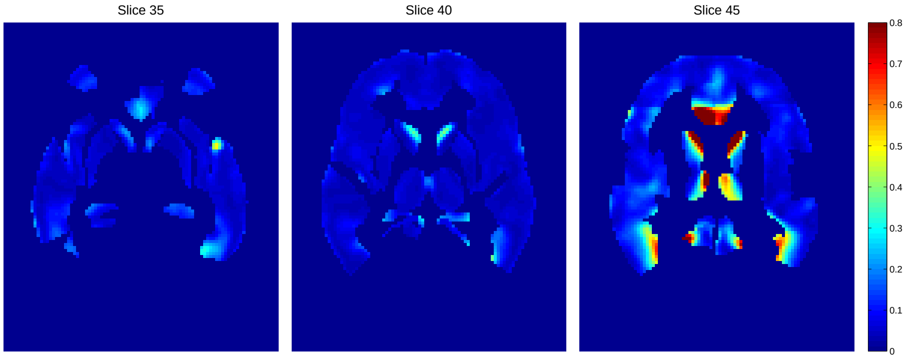
Figure 4.3(c) displays the square root of the PMSE, relative to the average brain activity, based on the GLM. A comparison between Figure 4.3(a) and Figure 4.3(c) indicates that prediction errors based on our proposed model are lower than those from the GLM, on average. The average error, for the model based on the GLM is 0.156.

Comparison with prediction based on BSMac. Prediction algorithm based on the BSMac is defined in a similar way to the algorithm based on our proposed BSPM. Since BSMac does not estimate the correlations between baseline and follow-up brain activity scans ($\mathbf{Y}_i$'s), i.e., temporal correlations, we estimate those by

(a) BSPM



(b) BSMac



(c) GLM

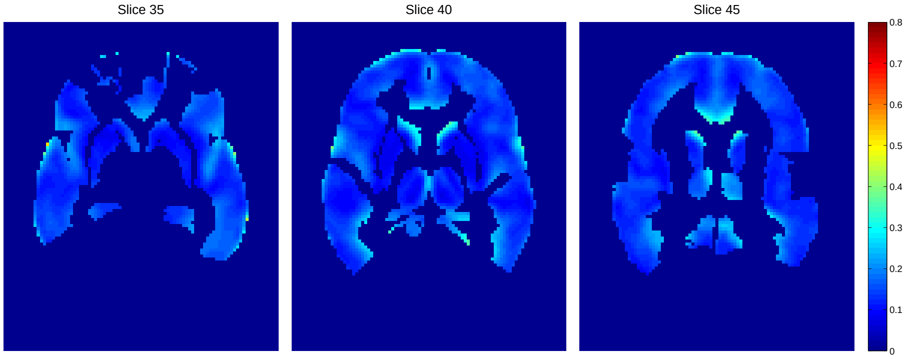
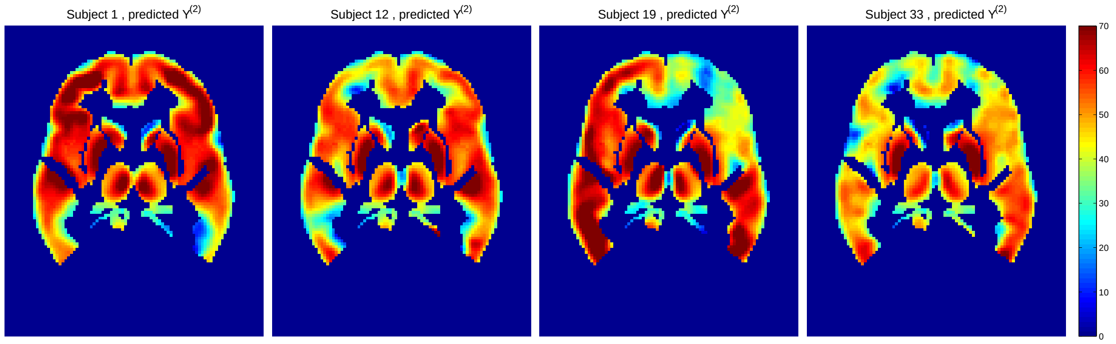


Figure 4.3: The images depict the square root of the prediction mean square error, divided by the average observed brain activity (stPMSE, (4.4)) at each voxel for prediction of the follow-up activity for 33 test subject in the AD group. Axial slices 35, 40 and 45 are shown in radiological view. (a) the stPMSE based on our proposed model; (b) the stPMSE based on BSMac; (c) the stPMSE based on the GLM. Average errors: 0.084 for the BSPM (a); 0.104 for the BSMac (b), and 0.156 for the GLM (c).

(a) predicted month 6 maps



(b) observed month 6 maps

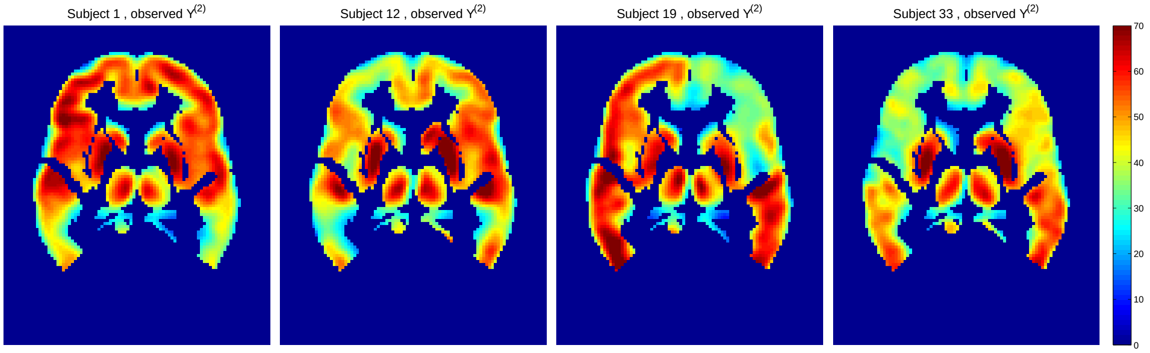


Figure 4.4: Prediction based on the BSMac. Individualized predicted and observed 6 month follow-up regional glucose uptake measurements for 4 test subjects in the AD group; axial slice 40 is shown.

the sample covariances, at a voxel level. The local spatial parameters are also not estimated, and are dropped from the estimated conditional mean. Other than those changes, the prediction algorithm is the same as the one based on the proposed BSPM.

Figure 4.4 depicts the predicted and observed regional glucose uptake maps for the same subjects as in Figure 4.2. We observe a fairly satisfactory agreement, but smaller than the agreement between the observed and predicted maps based on the BSPM.

Figure 4.3(b) displays the square root of the PMSE, relative to the average brain activity, based on the BSMac. A comparison between Figure 4.3(a) and Figure 4.3(b) indicates that the prediction errors based on our proposed model are also lower than those from the BSMac, on average. We also notice improvement over the GLM model. The average errors (total sum over all voxels, divided by the number of voxels included in the analysis) are 0.084 for the BSPM, 0.104 for the BSMac, and 0.156 for the GLM based prediction method.

4.5 Simulation

We performed a small simulation study to evaluate the accuracy of our estimation procedure. To generate data for the simulations, we selected 5 AAL regions (regions 79, 73, 15, 13, and 85) with sizes ranging from 234 to 4,655 voxels and simulated data for 15 subjects. We used our experimental data described in Section 4.2 to specify the *true* values for $\boldsymbol{\beta}_g$, spatial parameters $\boldsymbol{\phi}_g$, random effects parameters $\boldsymbol{\alpha}_i$, and the covariate parameters $\boldsymbol{\gamma}_g$. The rest of the (hyper)parameters were drawn from the prior distributions specified in the model. After all the parameters that define the mean and the covariance parameters in model (4.3) were specified, 200 simulated data sets were drawn, with the specified mean parameters and the variance-covariance matrix. We then applied our estimation procedure to each of the data sets and analyzed the posterior distributions of the parameters. Table 4.5 summarizes the results for the parameters of the covariance matrix $\mathbf{Z}_g$, for each of the 5 selected regions. We see that the biases for the covariance matrix parameters $\mathbf{Z}_g$ are small, for each of the regions. Also, spatial parameter ρ has only a small bias: we set the true value in simulated data sets to 0.9 and the mean of the estimated ρ, across 200 data sets, was found to be 0.9134 (for each data set, we first find a median of the parameter's posterior distribution, then average those medians). We believe that this bias would be even smaller if a larger number of iterations were performed. For brevity, posterior means of the estimated voxel-level parameters (for randomly selected 25 voxels in each region), as well as the posterior means of the random effects $\boldsymbol{\alpha}_{ij}$ are presented in the Appendix 5.2.4.

In general, we notice that the parameters that are most relevant in the prediction

Param.	Region									
	1		**2**		**3**		**4**		**5**	
	True	Est.	True	Est.	True	Est.	True	Est.	True	Est.
Z_g^{11}	323.26	321.25	160.61	160.05	11.46	11.46	3.44	3.44	3.57	3.57
Z_g^{22}	120.75	120.50	45.30	45.15	37.22	37.19	10.45	10.45	3.38	3.38
Z_g^{12}	-41.95	-41.40	-33.70	-33.43	16.32	16.30	4.08	4.08	2.40	2.40

Table 4.1: Summary of the simulation results for the parameters in the covariance matrix $\mathbf{Z}_g$, for each region. Z_g^{11} and Z_g^{22} denote the variance, and Z_g^{12} the covariance components.

($\boldsymbol{\beta}$'s, $\boldsymbol{\phi}$'s, $\boldsymbol{\alpha}$'s, $\mathbf{Z}_g$) have small or relatively small biases, while the parameters that contribute the prediction less directly, have larger biases. We believe that these would be reduced if longer Markov chains, i.e. larger number of iterations (at least 5,000), are generated.

Figures with trace plots for some selected voxels and parameters are given in Appendix 5.2.4.

4.6 Discussion

In this work, we describe a framework for spatiotemporal modeling of functional neuroimaging data that provides important advantages over some other methods. We propose a novel method for predicting post-baseline (or follow-up) brain scans, based on the baseline scans. The prediction algorithm is based on a novel Bayesian spatial hierarchical model. Our method is applicable in many clinical situations (e.g., for predicting the progression of a disease, or to predict after treatment brain activity, based on the baseline (pre-treatment) activity). The proposed method may be useful in clinical situations where it is too costly to acquire multiple (repeated) scans on the same subjects.

Our model builds on the predictive model by Guo et al. (2008), and a spatial Bayesian hierarchical model by Bowman et al. (2008), by incorporating a proper multivariate CAR prior for the spatial effect. The proposed model captures the short-range correlations between voxels within a defined anatomical region as well as the (potentially) long-range inter-regional correlations, which provide information about functional connectivity between the brain regions. We consider 3D neighborhood structure for estimation of the local spatial associations.

Based on the proposed model, we formulate a prediction algorithm for the follow-up brain activity, based on an individual's baseline functional neuroimaging data and relevant subject characteristics. By borrowing strength from the spatial information, we achieved improved prediction as compared to a couple of competitive prediction models (GLM and BSMac).

We apply our Bayesian spatial hierarchical model to the PET data from the study of Alzheimers disease, but the same methodology can easily be applied to a data from an fMRI study. In that case, an individual summary statistics (i.e., regression coefficients) from a typical first stage GLM based fMRI data analysis would first be obtained (in practice often obtained using software packages such as SPM or FSL), and then used as $\mathbf{Y}$'s in our model (4.3).

One of the limitations of our proposed method is that the computing time is very long (even for relatively small number of iterations of 3,000) for the estimation part of the algorithm. However, this step is done one time, and the prediction step is very fast (it is in order of seconds per one individual). Computing time can be improved by implementing some parallel computing steps (as described in Section 4.4). Another limitation is that we consider only within-region local 3D neighborhood in our model, which results in some region border effects in the prediction maps. We will consider across-region local 3D neighborhood in our future work.

Our future goal is to apply the proposed method to the MCI group and compare the results with AD and HC groups. We also plan to apply it to different data sets (e.g., from a treatment study on depression to evaluate the effect of the treatment using posterior inferences). In terms of comparisons with the competing methods, we plan to compare the prediction error and the performance of the BSPM with the spatio-temporal model proposed in Chapter 2 (Derado et al., 2010), and with the model proposed in Guo et al. (2008), in addition to GLM and BSMac.

Some recent studies De Meyer et al. (2010), Anoop et al. (2010), Buerger et al. (2002) (De Meyer et al. (2010), Buerger et al. (2002)) found that the β-amyloid protein 1-42, total tau protein, and phosphorylated $\text{tau}_{181\text{P}}$ protein concentrations, each derived from cerebral spinal fluid in the brain, may be clinically relevant biological markers for the differential diagnosis of AD. These biological measures may, therefore, serve as potentially useful covariates (predictors) in our model. Some of these proteins were collected in the ADNI study, but the rate of missing data (among subjects with FDG-PET scans) was too substantial for inclusion in our analysis.

Our ultimate goal is to develop a formalized algorithm for predicting the symptom response to treatment using baseline scans, predicted post-treatment (follow-up) activity and patient characteristics.

Chapter 5

Summary and Future Work

Brain imaging applications often produce large data sets that pose challenges for statistical analyses due, in part, to the intricate neurophysiology. To date, there is a paucity of methods that incorporate spatial considerations present in functional neuroimaging data. The main objective of our research is to uncover aspects of the complex spatial relationships present in functional neuroimaging data and to develop statistical methods that either evaluate or leverage those correlations.

Our first proposed method is related to activation studies, which attempt to localize regions of the brain activity when performing an experimental task. We propose a two-stage, spatio-temporal, autoregressive model which simultaneously accounts for spatial dependencies between voxels within the same anatomical region and for temporal dependencies between a subject's estimates from multiple sessions (Derado et al., 2010). Our approach provides a unified framework for voxel-level and region-level inferences. An important contribution of our work is that we derive efficient computational solutions to facilitate implementation. Using our proposed method, we analyze fMRI data from a study of inhibitory control in cocaine addicts. In summary, our proposed spatio-temporal model provides an appealing, computationally efficient alternative to standard GLM-based methods for analyzing fMRI data. Our model is based on assumptions that are more neurophysiologically plausible, capturing correlations between different brain locations and between estimates of neural activity at different scanning sessions. These correlations lead to interpretive advantages over the GLM, e.g., revealing information about the degree of coherence in brain activity within defined neuroanatomic regions. Our simulation studies demonstrate that our model estimates are quite accurate and that the standard errors associated with estimates of our mean model parameters are on average smaller than those from a GLM.

Our next proposed methodology is related to connectivity studies, which seek to identify what brain areas show similar patterns of activity over time, yielding distributed networks of brain function (Derado et al., 2010a). We propose a method to evaluate the level of connectivity within functionally defined neural networks. We introduce a global index to evaluate the validity of neural processing networks identified by functional clustering or by ICA. Specifically, we propose the use of Moran's I statistic (Cliff and Ord, 1973) to quantify and test the autocorrelation present within computed functional networks. We tailor our definition of the global index to neuroimaging applications through the use of network-based neighborhoods, and we also compute network-specific contributions to the overall autocorrelation. Consequently, our methods target autocorrelation exhibited by voxels within defined neural processing networks, rather than between voxels globally throughout the entire brain. We illustrate the use of our methodology with the data from two studies: a PET study of regional cerebral blood flow (rCBF) correlates of parametrically manipulated working memory among individuals with schizophrenia, and an fMRI resting-state study of depression. We also provide empirical support for the use of our hypothesis testing framework using a bootstrap analysis.

One of our future research goals is to evaluate the accuracy of the normal approximation formulas under randomization assumption (Cliff and Ord, 1973) derived for neuroimaging data. In particular, we think of the spatial correlations present in neuroimaging data. Inherent spatial correlations can be present in the data due to neurophysiology, data preprocessing, scanning acquisition etc. There are couple of different scenarios in our evaluation of the appropriateness of these approximations: to consider the data at the voxel level or to consider regional based clustering. Because the number of voxels or regions play the central role in the calculations, we may observe smaller or larger degrees of bias when we consider these two settings. In case we find that the answers vary, we hope to propose an alternative hypothesis testing framework to the one presented in Chapter 3. In our simulation study based on the PET data, the normal approximation seems to perform reasonably well. However, it is an open question whether or not the regional level analysis will need to do some adjustment for inherent spatial correlations in the data. We plan to investigate this in our future work.

Prediction studies try to use neuroimaging scans to predict some behavioral, psychological, or clinical outcomes, or to classify subgroups of individuals (e.g. for diagnostic purposes). Although functional brain imaging has become a powerful tool to investigate the physiological basis of psychiatric or neurological disorders, it has only recently been used to predict clinical outcomes in individual patients.

The third proposed method in this dissertation is a novel Bayesian hierarchical framework for predicting follow-up neural activity based on the baseline functional neuroimaging data. By borrowing strength from the spatial correlations present in the data, the prediction algorithm based on this model has improved prediction accuracy, compared to some competing models. The spatial correlations are incorporated in the model in two ways: the short-range correlations between neighboring voxels are incorporated through a multivariate conditional autoregressive (CAR) prior of the spatial parameters, while the long-range correlations between anatomical brain regions are incorporated through the covariance matrix of the random effect parameters. We applied our proposed model to the data of 40 AD patients and 40 healthy controls to estimate the parameters in our model which we then used to predict 6 month brain activity, based on the baseline activity and some relevant subject's characteristics to a set of 33 test subjects (in each group).

In the future, we plan to apply the proposed prediction algorithm to the mild cognitive impairment (MCI) group of subjects' data to predict the 6 month, 12 month, and 24 month brain activity. We think that this method had a potential clinical application is a sense that would enable clinicians to forecast future brain states, monitor the progression of dementia and to predict the onset of Alzheimer's disease. We also plan to apply the algorithm to the data from treatment studies, e.g. to forecast the effect of the treatment. Also, we will investigate how to modify our prediction model to further improve the prediction accuracy and to improve the computation cost.

Our long-term future goal is to develop a framework for predicting an individual's clinical response to treatment. To potentially improve classification (prediction) accuracy, we plan to investigate and develop methods that borrow strength from the spatial correlations present in the data. We plan to predict an individual's clinical response to treatment using spatially correlated, region-specific neural processing characteristics, such as region trends in serial brain activity measurements.

Appendices

5.1 Chapter 2 Appendices

5.1.1 Appendix A: Spatio-temporal Model in Matrix Form

Our stage-two spatio-temporal model (2.1) can be written in matrix form as follows:

$$\mathbf{B}_{igs} = \boldsymbol{\beta}_{gs} + \rho_{gs}\mathbf{W}^*(\mathbf{B}_{igs} - \boldsymbol{\beta}_{gs}) + \xi_{gs}\mathbf{M}^*(\mathbf{B}_{igs} - \boldsymbol{\beta}_{gs}) + \boldsymbol{e}_{igs}\,,$$

where $\boldsymbol{e}_{igs} \sim \mathrm{MVN}(\mathbf{0}, \phi_{gs}^2\mathbf{I})$ and where $\mathbf{W}^* = \mathbf{I}_q \otimes \mathbf{W}_g$, $\mathbf{M}^* = 1/(q-1)(\mathbf{J}_q \otimes \mathbf{I}_{V_g} - \mathbf{I}_q \otimes \mathbf{I}_{V_g})$ and $\mathbf{W}_g = 1/(V_g - 1)(\mathbf{J}_{V_g} - \mathbf{I}_{V_g})$. We can rewrite the above equation as

$$\mathbf{B}_{igs} = \boldsymbol{\beta}_{gs} + (\mathbf{I} - \rho_{gs}\mathbf{W}^* - \xi_{gs}\mathbf{M}^*)^{-1}\boldsymbol{e}_{igs} = \boldsymbol{\beta}_{gs} + \boldsymbol{\Psi}_{gs}\boldsymbol{e}_{igs} \tag{5.1}$$

Therefore, $\mathbf{B}_{igs} \sim \mathrm{MVN}(\boldsymbol{\beta}_{gs}, \phi_{gs}^2\boldsymbol{\Psi}_{gs}\boldsymbol{\Psi}_{gs})$.

5.1.2 Appendix B: Exploratory Data Analysis

B.1. *Crude estimates of the spatial and temporal correlations.*

To motivate our approach, we first performed an exploratory data analysis. We obtained crude empirical estimates of the spatial as well as temporal correlations in our data. To estimate the spatial correlations, we randomly selected a small percentage ($< 1\%$) of voxel-pairs from each anatomically defined brain region and calculated

Pearson correlation coefficients between the response variable values for these voxel pairs, across the subjects. Then we plotted the histogram of medians of these estimated correlations of each region.

To estimate the temporal correlations, we randomly selected a number of voxels from each region, calculated correlations between pre- and post treatment measurements, across the subjects, and plotted the medians of the region estimates. The estimates, shown in Figure 5.1, suggest that the voxels within the same regions have correlated brain activity and that there is correlated brain activity between pre- and post-treatment sessions.

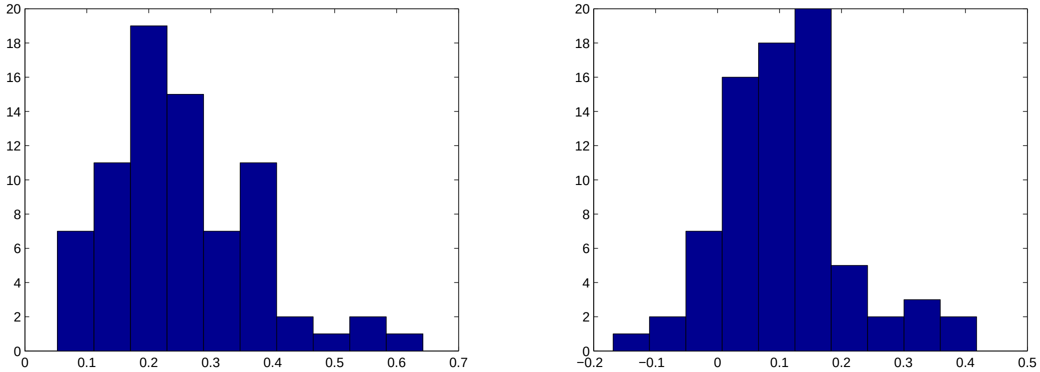


Figure 5.1: Empirical estimates of spatial correlations between randomly selected intra-regional voxel pairs for patients for the 6 mm FWHM smoothed data (left); empirical estimates of temporal (repeated measurement: pre- and post-treatment) correlations between 30% randomly selected voxels from each cluster (right).

B.2. *Exploring spatial correlations within anatomically defined regions.*

Descriptive covariance plots. To investigate the spatial correlations for our experimental data, we calculated the empirical correlation matrix separately for each of the regions in a Brodmann area-based parcellation of the brain. For every voxel in the region, the mean and the standard error of the correlations with each of the other voxels in that region were calculated (across the subjects) and plotted. The mean values were smoothed using the 'loess' curve to get a better idea of the overall spatial correlation for that region (see Figure 5.1.2). We performed this step for both patients and controls. We found that the means of the correlations did not fluctuate much within a region. Figure 5.1.2 shows such plots of several regions from the anatomical parcellation. Figure 5.1.2 provides support for the compound symmetric spatial covariance structure that we assume in our model.

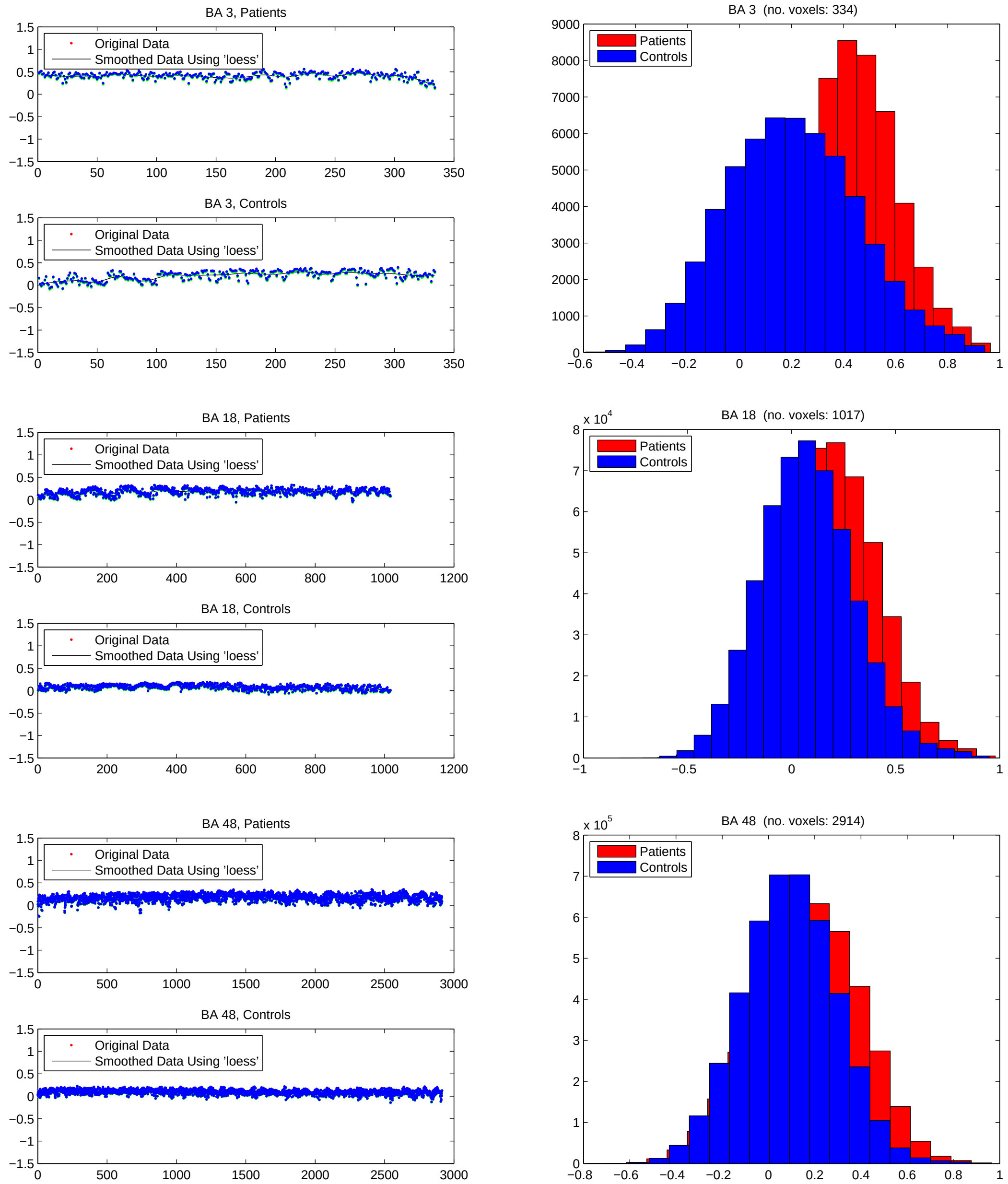


Figure 5.2: Empirical estimates of spatial correlations for 3 Brodmann regions (BA 3 (L) (part of the primary somatosensory cotrex), BA 18 (L) (V2-visual association cortex), and BA 48 (L) (retrosubicular area) with 334, 1017 and 2914 voxels respectivelly.

5.1.3 Appendix C. Computing Variance-Covariance Matrix: Coefficients of $\mathbf{\Psi}_g$.

Let $a = 1 + \frac{\rho_g}{V_g - 1}$, $b = -\frac{\rho_g}{V_g - 1}$ and $k = -\frac{\xi_g}{q-1}$. (For simplicity we drop the group index s here.) $\mathbf{\Psi}_g$ (see (2.3) in section 2.3.2) is a partitioned matrix with the *diagonal blocks* determined by

$$\begin{aligned} d &= \frac{a + (q-2)k}{a^2 + (q-2)ak - (q-1)k^2} \\ f &= \frac{-b\{a^2 + [1 + (q-1)(q-2)]k^2 + V_g[ab + (q-2)bk] + 2(q-2)ak\}}{[a^2 + (q-2)ak - (q-1)k^2]\{a^2 + (q-2)ak - (q-1)k^2 + V_g b[2a + (q-2)k + V_g b\}} \end{aligned}$$

The *off-diagonal blocks* are determined by

$$\begin{aligned} u &= \frac{-k}{a^2 + (q-2)ak - (q-1)k^2} \\ z &= \frac{bk[2a + (q-2)k + V_g b]}{[a^2 + (q-2)ak - (q-1)k^2]\{a^2 + (q-2)ak - (q-1)k^2 + V_g b[2a + (q-2)k + V_g b\}} \end{aligned}$$

This result was derived recursively on q, the number of blocks (sessions represented) in $\mathbf{\Psi}_g^{-1}$.

5.1.4 Appendix D. Score Functions.

We present the score functions, which facilitate implementation of our model. For simplicity, we drop the group notation s. Let $\mathbf{r}_{ig} = \mathbf{B}_{ig} - \mathbf{X}_i \boldsymbol{\beta}_g$ and we denote ϕ_g^2 by w_g. Also, from now on $\mathbf{I}$ will denote an identity matrix of size V_g and $\mathbf{J}$ will denote a $V_g \times V_g$ matrix of ones. In addition, $d = d(\rho_g, \xi_g, V_g, q)$, $f = (\rho_g, \xi_g, V_g, q)$, $u = (\rho_g, \xi_g, V_g, q)$, and $z = z(\rho_g, \xi_g, V_g, q)$ are the entries in matrix $\mathbf{\Psi}_g$, defined in section 2.3.2. $A = A(d, f, u, z, V_g, q)$, $B = B(d, f, u, z, V_g, q)$, $C = C(d, f, u, z, V_g, q)$, and $D = D(d, f, u, z, V_g, q)$ are the entries in $\mathbf{\Psi}_g \mathbf{\Psi}_g$, also defined in section 2.3.2.

The *score functions* are as follows:

(1)

$$S_{\boldsymbol{\beta}_g} = \frac{\partial \log L}{\partial \boldsymbol{\beta}_g} = \frac{1}{w_g} \sum_{i=1}^{K} \big\{ \mathbf{B}'_{ig} \boldsymbol{\Psi}_g^{-1} \boldsymbol{\Psi}_g^{-1} \mathbf{X}_i - \mathbf{X}'_i \boldsymbol{\Psi}_g^{-1} \boldsymbol{\Psi}_g^{-1} \mathbf{X}_i \boldsymbol{\beta}_g \big\}$$

From this we get (since in our case $\mathbf{X}_i = \mathbf{I}_{qV_g}$)):

$$\hat{\boldsymbol{\beta}}_{g_{\mathrm{OLS}}} = \frac{1}{K} \sum_{i=1}^{K} \mathbf{B}_{ig} \quad g = 1, \ldots, G.$$

(2)

$$S_{\rho_g} = \frac{\partial \log L}{\partial \rho_g} = \sum_{i=1}^{K} \mathrm{tr}\left\{ \frac{\partial \boldsymbol{\Psi}_g^{-1}}{\partial \rho_g} \cdot \boldsymbol{\Psi}_g \right\} - \frac{1}{2w_g} \sum_{i=1}^{K} \mathbf{r}'_{ig} \frac{\partial}{\partial \rho_g} \big(\boldsymbol{\Psi}_g^{-1} \boldsymbol{\Psi}_g^{-1} \big) \mathbf{r}_{ig} =$$

$$= -K \cdot q \cdot f \cdot V_g - \frac{1}{w_g} \sum_{i=1}^{K} \tfrac{1}{V_g - 1} \mathbf{r}'_{ig} \begin{bmatrix} k_1\mathbf{I} + k_2\mathbf{J} & k_3\mathbf{I} - k_3\mathbf{J} & \cdots & k_3\mathbf{I} - k_3\mathbf{J} \\ k_3\mathbf{I} - k_3\mathbf{J} & k_1\mathbf{I} + k_2\mathbf{J} & \cdots & k_3\mathbf{I} - k_3\mathbf{J} \\ \vdots & \vdots & \ddots & \vdots \\ k_3\mathbf{I} - k_3\mathbf{J} & k_3\mathbf{I} - k_3\mathbf{J} & \cdots & k_1\mathbf{I} + k_2\mathbf{J} \end{bmatrix} \mathbf{r}_{ig}$$

where

$$k_1 = 1 + \frac{\rho_g}{V_g - 1}, \quad k_2 = \rho_g - 1 - \frac{\rho_g}{V_g - 1}, \quad \text{and} \quad k_3 = -\frac{\xi_g}{q - 1}.$$

(3)

$$S_{\xi_g} = \frac{\partial \log L}{\partial \xi_g} = \sum_{i=1}^{K} \mathrm{tr}\left\{ \frac{\partial \boldsymbol{\Psi}_g^{-1}}{\partial \xi_g} \cdot \boldsymbol{\Psi}_g \right\} - \frac{1}{2w_g} \sum_{i=1}^{K} \mathbf{r}'_{ig} \frac{\partial}{\partial \xi_g} \big(\boldsymbol{\Psi}_g^{-1} \boldsymbol{\Psi}_g^{-1} \big) \mathbf{r}_{ig} =$$

$$= -K \cdot q \cdot (u + z) V_g - \frac{1}{w_g} \sum_{i=1}^{K} \mathbf{r}'_{ig} \begin{bmatrix} m_1\mathbf{I} & m_2\mathbf{I} + m_3\mathbf{J} & \cdots & m_2\mathbf{I} + m_3\mathbf{J} \\ m_2\mathbf{I} + m_3\mathbf{J} & m_1\mathbf{I} & \cdots & m_2\mathbf{I} + m_3\mathbf{J} \\ \vdots & \vdots & \ddots & \vdots \\ m_2\mathbf{I} + m_3\mathbf{J} & m_2\mathbf{I} + m_3\mathbf{J} & \cdots & m_1\mathbf{I} \end{bmatrix} \mathbf{r}_{ig}$$

where

$$m_1 = \frac{\xi_g}{q-1},\ m_2 = \frac{(q-2)\xi_g-(q-1)}{(q-1)^2} - \frac{\rho_g}{(q-1)(V_g-1)},\ \text{and}\ \ m_3 = \frac{\rho_g}{(q-1)(V_g-1)}\,.$$

(4)

$$\begin{aligned}
S_{w_g} &= \frac{\partial \log L}{\partial w_g} = \sum_{i=1}^{K}\Big\{-\frac{qV_g}{2w_g} + \frac{1}{2w_g^2}\mathbf{r}'_{ig}\boldsymbol{\Psi}_g^{-1}\boldsymbol{\Psi}_g^{-1}\mathbf{r}_{ig}\Big\} = \\
&= -\frac{KqV_g}{2w_g} + \frac{1}{2w_g^2}\sum_{i=1}^{K}\mathbf{r}'_{ig}\boldsymbol{\Psi}_g^{-1}\boldsymbol{\Psi}_g^{-1}\mathbf{r}_{ig} = \\
&= -\frac{KqV_g}{2w_g} + \frac{1}{2w_g^2}\sum_{i=1}^{K}\mathbf{r}'_{ig}\begin{bmatrix} u_1\mathbf{I}+u_2\mathbf{J} & u_3\mathbf{I}+u_4\mathbf{J} & \cdots & u_3\mathbf{I}+u_4\mathbf{J} \\ u_3\mathbf{I}+u_4\mathbf{J} & u_1\mathbf{I}+u_2\mathbf{J} & \cdots & u_3\mathbf{I}+u_4\mathbf{J} \\ \vdots & \vdots & \ddots & \vdots \\ u_3\mathbf{I}+u_4\mathbf{J} & u_3\mathbf{I}+u_4\mathbf{J} & \cdots & u_1\mathbf{I}+u_2\mathbf{J} \end{bmatrix}\mathbf{r}_{ig}
\end{aligned}$$

where

$$\begin{aligned}
u_1 &= \Big(1+\frac{\rho_g}{V_g-1}\Big)^2 + \frac{\xi_g^2}{q-1} & u_2 &= \frac{\rho_g}{V_g-1}\big(\rho_g - 2 - \frac{\rho_g}{V_g-1}\big) \\
u_3 &= -\frac{\xi_g}{q-1}\Big(2+\frac{2\rho_g}{V_g-1} - \frac{(q-2)\xi_g}{q-1}\Big) & u_4 &= \frac{2\xi_g\rho_g}{(q-1)(V_g-1)}\,.
\end{aligned}$$

In a similar fashion, one can derive second derivatives and their expected values.

5.2 Chapter 4 Appendices

5.2.1 Appendix E: Bayesian Spatial Hierarchical Model - region level

Let $\mathbf{Y}_{ig} = \left(\mathbf{Y}_{ig}^T(v_1), \ldots, \mathbf{Y}_{ig}^T(v_{V_g})\right)^T_{(2V_g \times 1)}$, and $v_1, \ldots, v_{V_G}$ denote the voxels in region g. The Kronecker product $(\mathbf{I}_{V_g} \otimes \mathbf{X}_{ig}) \cdot (\mathbf{1}_{V_g} \otimes \boldsymbol{\gamma}_g)$, i.e.

$$\begin{bmatrix} \mathbf{X}_{ig} & & & \\ & \mathbf{X}_{ig} & & \\ & & \ddots & \\ & & & \mathbf{X}_{ig} \end{bmatrix} \begin{bmatrix} \boldsymbol{\gamma}_g \\ \boldsymbol{\gamma}_g \\ \vdots \\ \boldsymbol{\gamma}_g \end{bmatrix} = (\mathbf{I}_{V_g} \mathbf{1}_{V_g}) \otimes (\mathbf{X}_{ig}\boldsymbol{\gamma}_g) = \mathbf{1}_{V_g} \otimes (\mathbf{X}_{ig}\boldsymbol{\gamma}_g).$$

Let also $\boldsymbol{\beta}_g = \left(\boldsymbol{\beta}_g^T(v_1), \ldots, \boldsymbol{\beta}_g^T(v_{V_g})\right)^T$. Since $\boldsymbol{\beta}_{gj}|\lambda_{gj}^2 \sim \mathrm{N}(\boldsymbol{\beta}_{0gj}, \lambda_{gj}^2 \mathbf{I})$ are assumed to be independent, then $\boldsymbol{\beta}_g \sim \mathrm{N}(\boldsymbol{\beta}_{0g}, \boldsymbol{\Lambda}_g)$, where

$$\boldsymbol{\beta}_{0g} = \begin{bmatrix} \boldsymbol{\beta}_{0g}^{(1)}(v_1) \\ \boldsymbol{\beta}_{0g}^{(2)}(v_1) \\ \vdots \\ \boldsymbol{\beta}_{0g}^{(1)}(v_{V_g}) \\ \boldsymbol{\beta}_{0g}^{(2)}(v_{V_g}) \end{bmatrix}, \ \boldsymbol{\Lambda}_g = \begin{bmatrix} \lambda_{g1}^2 & & & & \\ & \lambda_{g2}^2 & & & \\ & & \ddots & & \\ & & & \lambda_{g1}^2 & \\ & & & & \lambda_{g2}^2 \end{bmatrix} = \mathbf{I}_{V_g} \otimes \begin{bmatrix} \lambda_{g1}^2 & 0 \\ 0 & \lambda_{g2}^2 \end{bmatrix} = \mathbf{I}_{V_g} \otimes \boldsymbol{\Lambda}_v.$$

Finally, let $\mathbf{V}_g = \mathbf{I}_{V_g} \otimes \mathbf{Z}_g$, where $\mathbf{Z}_g = \begin{bmatrix} Z_g^{11} & Z_g^{12} \\ Z_g^{21} & Z_g^{22} \end{bmatrix}$ (same for all $v \in g$), and let $\boldsymbol{\phi}_g = [\boldsymbol{\phi}_g^T(v_1) \cdots \boldsymbol{\phi}_g^T(v_{V_g})]^T$. Then model (4.3) can be re-expressed as

Model (region level):

$$\begin{aligned}
\mathbf{Y}_{ig}|\boldsymbol{\beta}_g,\boldsymbol{\phi}_g,\boldsymbol{\alpha}_{ig},\boldsymbol{\gamma}_{gv},\mathbf{V}_g &\sim \mathrm{N}\big(\boldsymbol{\beta}_g+\boldsymbol{\phi}_g+\mathbf{1}_{V_g}\otimes\boldsymbol{\alpha}_{ij}+\mathbf{1}_{V_g}\otimes(\mathbf{X}_{ig}\boldsymbol{\gamma}_g),\mathbf{V}_g\big)\\
\boldsymbol{\beta}_{gj}|\lambda^2_{gj} &\sim \mathrm{N}(\mathbf{1}_{V_g}\boldsymbol{\beta}_{0gj},\lambda^2_{gj}\mathbf{I}_{V_g})\\
\boldsymbol{\phi}_g|\boldsymbol{\Sigma}_{\phi_g^{-1}} &\sim \mathrm{N}(\mathbf{0},\boldsymbol{\Sigma}_{\phi_g^{-1}}),\quad \boldsymbol{\Sigma}_{\phi_g^{-1}}=(D_{W_g}-\rho W_g)^{-1}\otimes\Sigma\\
\boldsymbol{\alpha}_{ij}|\boldsymbol{\Gamma}_j &\sim \mathrm{N}(\mathbf{0},\boldsymbol{\Gamma}_j),\quad \text{where } \boldsymbol{\alpha}_i=[\boldsymbol{\alpha}_{i1}^T,\ldots,\boldsymbol{\alpha}_{iG}^T]^T,\ j=1,2\\
\boldsymbol{\gamma}_{gj}|\boldsymbol{\tau}^2_{gj} &\sim \mathrm{N}(\mathbf{0},\mathrm{diag}(\tau^2_{gj1},\ldots,\tau^2_{gjQ}))\\
(\boldsymbol{\Gamma}_j)^{-1} &\sim \mathrm{Wishart}\{(h_jH_j)^{-1},h_j\},\quad j=1,2\\
\lambda^{-2}_{gj} &\sim \mathrm{Gamma}(a_j,b_j)\\
\tau^{-2}_{gjq} &\sim \mathrm{Gamma}(e_{0q},f_{0q})\\
\mathbf{Z}_g^{-1} &\sim \mathrm{Wishart}\big((c_1\boldsymbol{\Omega}_1)^{-1},c_1\big)\\
\Sigma^{-1} &\sim \mathrm{Wishart}\big((c_2\boldsymbol{\Omega}_2)^{-1},c_2\big)\\
\rho &\sim \mathrm{Uniform}(\{0,0.05,0.1,\ldots,0.8,0.81,\ldots,\\
&\qquad 0.9,0.91,\ldots,0.99\})
\end{aligned}\tag{5.2}$$

5.2.2 Appendix F: Full Conditional Distributions for the Bayesian Spatial Hierarchical Model

Let for each $g=1,\ldots,G$

$$\mathbf{Y}_{ig}=\begin{bmatrix}\mathbf{Y}_{ig}(v_1)\\ \vdots\\ \mathbf{Y}_{ig}(v_{V_g})\end{bmatrix}_{(2V_g\times 1)},\ \boldsymbol{\phi}_g=\begin{bmatrix}\boldsymbol{\phi}_g(v_1)\\ \vdots\\ \boldsymbol{\phi}_g(v_{V_g})\end{bmatrix}_{(2V_g\times 1)},\quad \boldsymbol{\alpha}_{ig}=\begin{bmatrix}\boldsymbol{\alpha}_{ig}(v_1)\\ \vdots\\ \boldsymbol{\alpha}_{ig}(v_{V_g})\end{bmatrix}_{(2V_g\times 1)}=\mathbf{1}_{V_g}\otimes\boldsymbol{\alpha}_{ig}$$

Then we can write the model (5.2) as

$$\mathbf{Y}_{ig}=\boldsymbol{\beta}_g+\boldsymbol{\phi}_g+\mathbf{1}_{V_g}\otimes\boldsymbol{\alpha}_{ig}+\mathbf{1}_{V_g}\otimes\mathbf{X}_{igv}\boldsymbol{\gamma}_{gv}+\boldsymbol{\varepsilon}_{ig}\tag{5.3}$$

where $\boldsymbol{\beta}_g = (\boldsymbol{\beta}_g^T(v_1), \ldots, \beta_g^T(v_{V_g}))$.

Let $\boldsymbol{\beta}_{0g} = (\boldsymbol{\beta}_{0g}^T(v_1), \ldots, \boldsymbol{\beta}_{0g}^T(v_{V_g}))$.

Prior for $\boldsymbol{\beta}$ (combined pre and post):

$$\boldsymbol{\beta}_g^{(1)} \sim \mathrm{N}(\boldsymbol{\beta}_{0g1}^{(1)}, \lambda_{g1}^2 \mathbf{I}_{V_g})$$
$$\boldsymbol{\beta}_g^{(2)} \sim \mathrm{N}(\boldsymbol{\beta}_{0g2}^{(2)}, \lambda_{g2}^2 \mathbf{I}_{V_g})$$

Under the independence assumption

$$\boldsymbol{\beta}_g \sim \mathrm{N}\big(\boldsymbol{\beta}_{0g}, \boldsymbol{\Lambda}_g\big)$$

where

$$\boldsymbol{\Lambda}_g = \begin{bmatrix} \lambda_{g1}^2 & 0 & & & & \\ 0 & \lambda_{g2}^2 & & & & \\ & & \ddots & & & \\ & & & \ddots & & \\ & & & & \lambda_{g1}^2 & 0 \\ & & & & 0 & \lambda_{g2}^2 \end{bmatrix} = \mathbf{I}_{V_g} \otimes \begin{bmatrix} \lambda_{g1}^2 & 0 \\ 0 & \lambda_{g2}^2 \end{bmatrix}$$

Deriving Full Conditional Distributions.

Denote by $\boldsymbol{\Lambda}(\boldsymbol{\alpha})$ the set of all parameters, excluding parameter $\boldsymbol{\alpha}$. The full conditional distributions for each of the parameters in the model are as follows.

(1)

$$
\begin{aligned}
[\boldsymbol{\beta}_g|\boldsymbol{\Lambda}(\boldsymbol{\beta}_g),\mathbf{Y}] &\propto [\boldsymbol{\beta}_g|\boldsymbol{\Lambda}_g]\times\prod_{i=1}^{n}[Y_{ig}|\boldsymbol{\beta}_g,\boldsymbol{\phi}_g,\boldsymbol{\alpha}_{ig},\boldsymbol{\gamma}_g,\mathbf{V}_g]\\
&\propto \exp\big\{-\tfrac{1}{2}(\boldsymbol{\beta}_g-\boldsymbol{\beta}_{0g})^T\boldsymbol{\Lambda}_g^{-1}(\boldsymbol{\beta}_g-\boldsymbol{\beta}_{0g})\big\}\times\\
&\quad |\mathbf{V}_g|^{-\frac{n}{2}V_g}\exp\Big\{-\tfrac{1}{2}\sum_{i=1}^{n}[\mathbf{Y}_{ig}-\boldsymbol{\beta}_g-\boldsymbol{\phi}_g-\mathbf{1}\otimes\boldsymbol{\alpha}_{ig}-\mathbf{1}\otimes\mathbf{X}_{ig}\boldsymbol{\gamma}_g]^T\mathbf{V}_g^{-1}\\
&\quad [\mathbf{Y}_{ig}-\boldsymbol{\beta}_g-\boldsymbol{\phi}_g-\mathbf{1}\otimes\boldsymbol{\alpha}_{ig}-\mathbf{1}\otimes\mathbf{X}_{ig}\boldsymbol{\gamma}_g]\Big\}\\
&\propto \exp\big\{-\tfrac{1}{2}[\boldsymbol{\beta}_g^T\boldsymbol{\Lambda}_g^{-1}\boldsymbol{\beta}_g-2\boldsymbol{\beta}_g^T\boldsymbol{\Lambda}_g^{-1}\boldsymbol{\beta}_{0g}]\big\}\times\\
&\quad \exp\Big\{-\tfrac{1}{2}\sum_{i=1}^{n}[\boldsymbol{\beta}_g^T\mathbf{V}_g^{-1}\boldsymbol{\beta}_g-2\boldsymbol{\beta}_g^T\mathbf{V}_g^{-1}\underbrace{(\mathbf{Y}_{ig}\boldsymbol{\phi}_g-\mathbf{1}\otimes\boldsymbol{\alpha}_{ig}-\mathbf{1}\otimes\mathbf{X}_{ig}\boldsymbol{\gamma}_g)}_{:=\mathbf{r}_{ig}}]\Big\}\\
&\propto \exp\big\{-\tfrac{1}{2}[\boldsymbol{\beta}_g^T[\boldsymbol{\Lambda}_g^{-1}+n\mathbf{V}_g^{-1}]\boldsymbol{\beta}_g-2\boldsymbol{\beta}_g^T[\boldsymbol{\Lambda}_g^{-1}\boldsymbol{\beta}_{0g}+\mathbf{V}_g^{-1}\sum_{i=1}^{n}\mathbf{r}_{ig}]\big\}\\
&\propto \mathrm{N}\big(\boldsymbol{\Omega}_g\mathbf{T}_g,\boldsymbol{\Omega}_g\big)
\end{aligned}
$$

where $\boldsymbol{\Omega}_g$, $\mathbf{T}_g$, $\boldsymbol{\Lambda}_g$ are as defined in Section 4.3.2

(2) The full conditional distribution for $\boldsymbol{\phi}_v$ is a likelihood adjusted version of the conditional distribution for MCAR(ρ,Σ):

$$
\begin{aligned}
&[\boldsymbol{\phi}_v|\boldsymbol{\Lambda}(\boldsymbol{\phi}_v), \mathbf{Y}] \propto \\
&\propto \quad \mathrm{N}\big(\rho \sum_{v' \neq v} \frac{w_{vv'}}{w_{v+}} \boldsymbol{\phi}_{v'}, w_{v+}^{-1}\Sigma\big) \cdot \prod_{i=1}^{n} \mathrm{N}(\boldsymbol{\mu}_{igv}, \mathbf{Z}_g) \ \ (\text{denote } b_{v'} = \tfrac{w_{vv'}}{w_{v+}}) \\
&\propto \quad \exp\Big\{-\frac{1}{2}\big(\boldsymbol{\phi}_v - \rho \sum_{v'} b_{v'}\boldsymbol{\phi}_{v'}\big)^T w_{v+}\Sigma^{-1}\big(\boldsymbol{\phi}_v - \rho \sum_{v'} b_{v'}\boldsymbol{\phi}_{v'}\big)\Big\} \cdot \\
&\qquad \exp\Big\{-\frac{1}{2}\sum_{i=1}^{n}[\mathbf{Y}_{ig}(v) - (\underbrace{\boldsymbol{\beta}_g(v) + \boldsymbol{\phi}_G(v) + \boldsymbol{\alpha}_{ig}(v) + \mathbf{X}_{ig}\boldsymbol{\gamma}_g}_{\mathbf{r}})]^T \mathbf{Z}_g^{-1}[\mathbf{Y}_{ig} - \mathbf{r}]\Big\} \\
&\propto \quad \exp\Big\{-\frac{1}{2} w_{v+}[\boldsymbol{\phi}_v^T \Sigma^{-1} \boldsymbol{\phi}_v - 2\boldsymbol{\phi}_v^T \Sigma^{-1} \rho \sum_{v'} b_{v'} \boldsymbol{\phi}_{v'}]\Big\} \cdot \\
&\qquad \exp\Big\{-\frac{1}{2}\sum_{i=1}^{n}[\boldsymbol{\phi}_v - (\underbrace{\mathbf{Y}_{ig}(v) - \boldsymbol{\beta}_g(v) - \boldsymbol{\phi}_G(v) - \boldsymbol{\alpha}_{ig} - \mathbf{X}_{ig}\boldsymbol{\gamma}_g}_{\mathbf{s}_{ig}})]^T[\boldsymbol{\phi}_v - \mathbf{s}_{ig}]\Big\} \\
&\propto \quad \exp\Big\{ \text{ same as above } \Big\} \cdot \\
&\qquad \exp\Big\{-\frac{1}{2}\Big[\sum_{i=1}^{n} \boldsymbol{\phi}_v^T \mathbf{Z}_g^{-1} \boldsymbol{\phi}_v - 2\boldsymbol{\phi}_v^T \mathbf{Z}_g^{-1} \sum_{i=1}^{n} \mathbf{s}_{ig}\Big]\Big\} \\
&\propto \quad \exp\Big\{ \text{ same as above } \Big\} \cdot \\
&\qquad \exp\Big\{n\boldsymbol{\phi}_v^T \mathbf{Z}_g^{-1} \boldsymbol{\phi}_v - 2\boldsymbol{\phi}_v^T \mathbf{Z}_g^{-1} n \cdot \big(\bar{\mathbf{Y}}_g(v) - \boldsymbol{\beta}_g(v) - \bar{\boldsymbol{\alpha}}_g - \bar{\mathbf{X}}_g \boldsymbol{\gamma}_g\big)\Big\}
\end{aligned}
$$

where

$$
\begin{aligned}
\bar{\mathbf{Y}}_g(v) &= \frac{1}{n}\sum_{i=1}^{n} \mathbf{Y}_{ig}(v) \\
\bar{\boldsymbol{\alpha}}_g &= \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{\alpha}_{ig} \\
\bar{\mathbf{X}}_g &= \frac{1}{n}\sum_{i=1}^{n} \mathbf{X}_{ig}
\end{aligned}
$$

Then

$$
\begin{aligned}
[\boldsymbol{\phi}_v|\boldsymbol{\Lambda}(\boldsymbol{\phi}_v),\mathbf{Y}] \propto \exp\Big\{-\frac{1}{2}\Big[\boldsymbol{\phi}_v^T(\underbrace{w_{v+}\Sigma^{-1}+n\mathbf{Z}_g^{-1}}_{\mathbf{H}_v^{-1}})\boldsymbol{\phi}_v\Big]\Big\} \\
-2\boldsymbol{\phi}_v^T\underbrace{\big(w_{v+}\rho\Sigma^{-1}\sum_{v'\neq v}\frac{w_{vv'}}{w_{v+}}\boldsymbol{\phi}_{v'}+n\mathbf{Z}_g^{-1}(\bar{\mathbf{Y}}_g(v)-\boldsymbol{\beta}_g(v)-\bar{\boldsymbol{\alpha}}_g-\bar{\mathbf{X}}_g\boldsymbol{\gamma}_g)\big)}_{\mathbf{P}_v}\Big]\Big\} \\
\propto \quad \mathrm{N}(\mathbf{H}_v\mathbf{P}_v,\mathbf{H}_v).
\end{aligned}
$$

(3)

$$
\begin{aligned}
&[\boldsymbol{\alpha}_i|\boldsymbol{\Lambda}(\boldsymbol{\alpha}_i),\mathbf{Y}) \propto \mathrm{N}(\mathbf{0},\boldsymbol{\Gamma})\cdot\prod_{g=1}^{G}\mathrm{N}(\boldsymbol{\beta}_g+\boldsymbol{\phi}_g+\mathbf{1}_{V_g}\otimes\boldsymbol{\alpha}_{ig}+\mathbf{1}_{V_g}\otimes\mathbf{X}_{ig}\boldsymbol{\gamma}_g,\mathbf{V}_g) \\
&\propto \quad \exp\big\{-\tfrac{1}{2}\boldsymbol{\alpha}_i^T\boldsymbol{\Gamma}^{-1}\boldsymbol{\alpha}_i\big\}\cdot \\
&\qquad \exp\big\{-\tfrac{1}{2}\sum_{g=1}^{G}[\mathbf{Y}_{ig}-\boldsymbol{\beta}_g-\boldsymbol{\phi}_g-\mathbf{1}_{V_g}\otimes\boldsymbol{\alpha}_{ig}-\mathbf{1}_{V_g}\otimes\mathbf{X}_{ig}\boldsymbol{\gamma}_g]^T\mathbf{V}_g^{-1}\cdot \\
&\qquad\quad [\mathbf{Y}_{ig}-\boldsymbol{\beta}_g-\boldsymbol{\phi}_g-\mathbf{1}_{V_g}\otimes\boldsymbol{\alpha}_{ig}-\mathbf{1}_{V_g}\otimes\mathbf{X}_{ig}\boldsymbol{\gamma}_g]\big\} \\
&\propto \quad \exp\big\{-\tfrac{1}{2}\boldsymbol{\alpha}_i^T\boldsymbol{\Gamma}^{-1}\boldsymbol{\alpha}_i\big\}\cdot \\
&\qquad \exp\big\{-\tfrac{1}{2}\sum_{g=1}^{G}[\mathbf{1}_{V_g}\otimes\boldsymbol{\alpha}_{ig}-(\underbrace{\mathbf{Y}_{ig}-\boldsymbol{\beta}_g-\boldsymbol{\phi}_g-\mathbf{1}_{V_g}\otimes\mathbf{X}_{ig}\boldsymbol{\gamma}_g}_{=:\mathbf{t}_{ig}})]^T\mathbf{V}_g^{-1}[\mathbf{1}_{V_g}\otimes\boldsymbol{\alpha}_{ig}-\mathbf{t}_{ig}]\big\} \\
&\propto \quad \exp\big\{-\tfrac{1}{2}\boldsymbol{\alpha}_i^T\boldsymbol{\Gamma}^{-1}\boldsymbol{\alpha}_i\big\}\cdot \\
&\qquad \exp\Big\{-\tfrac{1}{2}\Big[\underbrace{\sum_{g=1}^{G}(\mathbf{1}_{V_g}\otimes\boldsymbol{\alpha}_{ig})^T\mathbf{V}_g^{-1}(\mathbf{1}_{V_g}\otimes\boldsymbol{\alpha}_{ig})}_{\mathbf{I}}-2\underbrace{\sum_{g=1}^{G}(\mathbf{1}_{V_g}\otimes\boldsymbol{\alpha}_{ig})^T\mathbf{V}_g^{-1}\mathbf{t}_{ig}}_{\mathbf{II}}\Big]\Big\}
\end{aligned}
$$

$$\mathbf{I} = \sum_{g=1}^{G} (\mathbf{1}_{V_g} \otimes \boldsymbol{\alpha}_{ig})^T \mathbf{V}_g^{-1} (\mathbf{1}_{V_g} \otimes \boldsymbol{\alpha}_{ig}) = \underbrace{[\boldsymbol{\alpha}_{i1}^T, \ldots, \boldsymbol{\alpha}_{i1}^T]}_{V_1 \text{ times}} \begin{bmatrix} \mathbf{Z}_1^{-1} & & \\ & \ddots & \\ & & \mathbf{Z}_1^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha}_{i1} \\ \vdots \\ \boldsymbol{\alpha}_{i1} \end{bmatrix} + \cdots +$$

$$\underbrace{[\boldsymbol{\alpha}_{iG}^T, \ldots, \boldsymbol{\alpha}_{iG}^T]}_{V_G \text{ times}} \begin{bmatrix} \mathbf{Z}_G^{-1} & & \\ & \ddots & \\ & & \mathbf{Z}_G^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha}_{iG} \\ \vdots \\ \boldsymbol{\alpha}_{iG} \end{bmatrix} =$$

$$= V_1 \boldsymbol{\alpha}_{i1}^T \mathbf{Z}_1^{-1} \boldsymbol{\alpha}_{i1} + \ldots + V_G \boldsymbol{\alpha}_{iG}^T \mathbf{Z}_G^{-1} \boldsymbol{\alpha}_{iG} \quad (V_g \text{ is the number of voxels in region } g)$$

$$= \underbrace{[\boldsymbol{\alpha}_{i1}^T, \ldots, \boldsymbol{\alpha}_{iG}^T]}_{\boldsymbol{\alpha}_i^T} \underbrace{\begin{bmatrix} V_1 \mathbf{Z}_1^{-1} & & \\ & \ddots & \\ & & V_G \mathbf{Z}_G^{-1} \end{bmatrix}}_{\boldsymbol{\Omega}_\alpha^{-1}} \begin{bmatrix} \boldsymbol{\alpha}_{i1} \\ \vdots \\ \boldsymbol{\alpha}_{iG} \end{bmatrix}$$

$$\mathbf{II} = \sum_{g=1}^{G} (\mathbf{1}_{V_g} \otimes \boldsymbol{\alpha}_{ig})^T \mathbf{V}_g^{-1} \mathbf{t}_{ig} = [\boldsymbol{\alpha}_{i1}^T, \ldots, \boldsymbol{\alpha}_{i1}^T] \begin{bmatrix} \mathbf{Z}_1^{-1} & & \\ & \ddots & \\ & & \mathbf{Z}_1^{-1} \end{bmatrix} \mathbf{t}_{i1} + \cdots +$$

$$[\boldsymbol{\alpha}_{iG}^T, \ldots, \boldsymbol{\alpha}_{iG}^T] \begin{bmatrix} \mathbf{Z}_G^{-1} & & \\ & \ddots & \\ & & \mathbf{Z}_G^{-1} \end{bmatrix} \mathbf{t}_{iG} =$$

$$= \underbrace{[\boldsymbol{\alpha}_{i1}^T, \ldots, \boldsymbol{\alpha}_{iG}^T]}_{\boldsymbol{\alpha}_i^T} \begin{bmatrix} \underbrace{\mathbf{Z}_1^{-1} \ldots \mathbf{Z}_1^{-1}}_{V_1 \text{ times}} & & \\ & \ddots & \\ & & \underbrace{\mathbf{Z}_G^{-1} \ldots \mathbf{Z}_G^{-1}}_{V_G \text{ times}} \end{bmatrix} \begin{bmatrix} \mathbf{t}_{i1} \\ \vdots \\ \mathbf{t}_{iG} \end{bmatrix}$$

$$= \underbrace{\boldsymbol{\alpha}_i^T}_{(2 \times 2G)} \underbrace{\begin{bmatrix} \mathbf{1}_{V_1} \otimes \mathbf{Z}_1^{-1} & & \\ & \ddots & \\ & & \mathbf{1}_{V_G} \otimes \mathbf{Z}_G^{-1} \end{bmatrix}}_{\mathbf{P}} \begin{bmatrix} \mathbf{t}_{i1} \\ \vdots \\ \mathbf{t}_{iG} \end{bmatrix}.$$

Hence, it follows that

$$\begin{aligned}
[\boldsymbol{\alpha}_i|\boldsymbol{\Lambda}(\boldsymbol{\alpha}_i),\mathbf{Y}) &\propto \exp\bigl\{-\tfrac{1}{2}\boldsymbol{\alpha}_i^T\boldsymbol{\Gamma}^{-1}\boldsymbol{\alpha}_i\bigr\}\times\exp\bigl\{-\tfrac{1}{2}\boldsymbol{\alpha}_i^T\boldsymbol{\Omega}_\alpha^{-1}\boldsymbol{\alpha}_i-2\boldsymbol{\alpha}_i^T\mathbf{P}\bigr\}\\
&\propto \exp\bigl\{-\tfrac{1}{2}\boldsymbol{\alpha}_i^T\underbrace{(\boldsymbol{\Gamma}^{-1}+\boldsymbol{\Omega}_\alpha^{-1})}_{\boldsymbol{\Psi}^{-1}}\boldsymbol{\alpha}_i-2\boldsymbol{\alpha}_i^T\mathbf{P}\bigr\}\\
&\propto \mathrm{N}(\boldsymbol{\Psi}\mathbf{P},\boldsymbol{\Psi}),
\end{aligned}$$

where

$\boldsymbol{\Psi}=\bigl(\boldsymbol{\Gamma}^{-1}+\boldsymbol{\Omega}_\alpha^{-1}\bigr)^{-1}$,

$$\mathbf{P}=\begin{bmatrix}1_{V_1}\otimes\mathbf{Z}_1^{-1} & \dots & \dots\\ \vdots & \ddots & \vdots\\ \dots & \dots & 1_{V_G}\otimes\mathbf{Z}_G^{-1}\end{bmatrix}\cdot\begin{bmatrix}\mathbf{Y}_{i1}-\boldsymbol{\beta}_1-\boldsymbol{\phi}_1-1_{V_1}\otimes\mathbf{X}_{i1}\boldsymbol{\gamma}_1\\ \vdots\\ \mathbf{Y}_{iG}-\boldsymbol{\beta}_G-\boldsymbol{\phi}_G-1_{V_G}\otimes\mathbf{X}_{iG}\boldsymbol{\gamma}_G\end{bmatrix}, \text{ and}$$

$$\boldsymbol{\Gamma}^{-1}=\boldsymbol{\Gamma}_1^{-1}\otimes\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}+\boldsymbol{\Gamma}_2^{-1}\otimes\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}, \text{ while } \boldsymbol{\Omega}_\alpha^{-1}=\begin{bmatrix}V_1\mathbf{Z}_1^{-1} & \dots & \dots\\ \vdots & \ddots & \vdots\\ \dots & \dots & V_G\mathbf{Z}_G^{-1}\end{bmatrix}.$$

(4)

$$\begin{aligned}
[\boldsymbol{\Gamma}_j^{-1}|\boldsymbol{\Lambda}(\boldsymbol{\Gamma}_j^{-1}),\mathbf{Y}] &\propto [\boldsymbol{\Gamma}_j^{-1}]\cdot\prod_{i=1}^{n}[\boldsymbol{\alpha}_i^{(j)}|\boldsymbol{\Gamma}_j^{-1}]\\
&\propto \mathrm{Wishart}\bigl((h_jH_j)^{-1},h_j\bigr)\cdot\prod_{i=1}^{n}\mathrm{N}(\mathbf{0},\boldsymbol{\Gamma}_j)\\
&\propto |\boldsymbol{\Gamma}_j|^{(h_j-G-1)/2}\exp\bigl\{-\frac{h_j}{2}\mathrm{tr}(H_j\boldsymbol{\Gamma}_j^{-1})\bigr\}\cdot\\
&\qquad |\boldsymbol{\Gamma}_j|^{n/2}\exp\bigl\{-\frac{1}{2}\sum_{i=1}^{n}\boldsymbol{\alpha}_{ij}^T\boldsymbol{\Gamma}_j^{-1}\boldsymbol{\alpha}_{ij}\bigr\}\\
&\propto |\boldsymbol{\Gamma}_j|^{(h_j+n-G-1)/2}\exp\bigl\{-\frac{1}{2}\mathrm{tr}[(h_jH_j+\sum_{i=1}^{n}\boldsymbol{\alpha}_{ij}\boldsymbol{\alpha}_{ij}^T)\boldsymbol{\Gamma}_j^{-1}]\bigr\}\\
&\propto \mathrm{Wishart}\bigl((h_jH_j+\sum_{i=1}^{n}\boldsymbol{\alpha}_{ij}\boldsymbol{\alpha}_{ij}^T)^{-1},h_j+n\bigr)
\end{aligned}$$

(5)

$$
\begin{aligned}
[\mathbf{Z}_g^{-1}|\boldsymbol{\Lambda}(\mathbf{Z}_g^{-1}),\mathbf{Y}] &\propto [\mathbf{Z}_g^{-1}]\cdot\prod_{i=1}^{n}\prod_{v\in g}\mathrm{N}(\mathbf{Y}_{ig}(v)|\boldsymbol{\beta}_g(v),\boldsymbol{\phi}_g(v),\boldsymbol{\alpha}_{ig},\mathbf{X}_{ig}\boldsymbol{\gamma}_g,\mathbf{Z}_g)\\
&\quad(\text{ Let } \mathbf{u}_{igv}=\mathbf{Y}_{ig}(v)-\boldsymbol{\beta}_g(v)-\boldsymbol{\phi}_g(v)-\boldsymbol{\alpha}_{ig}-\mathbf{X}_{ig}\boldsymbol{\gamma}_g)\\
&\propto\ |\mathbf{Z}_g^{-1}|^{(c_1-2-1)/2}\exp\big\{-\tfrac{1}{2}\mathrm{tr}(c_1\Omega_1\mathbf{Z}_g^{-1})\big\}\cdot\\
&\quad\ |\mathbf{Z}_g^{-1}|^{nV_g/2}\exp\big\{-\tfrac{1}{2}\sum_{i=1}^{n}\sum_{v\in g}\mathbf{u}_{igv}^T\mathbf{Z}_g^{-1}\mathbf{u}_{igv}\big\}\\
&\propto\ |\mathbf{Z}_g^{-1}|^{(c_1+nV_g-2-1)/2}\exp\big\{-\tfrac{1}{2}\mathrm{tr}(c_1\Omega_1\mathbf{Z}_g^{-1})-\tfrac{1}{2}\underbrace{\sum_{i=1}^{n}\mathrm{tr}\big((\sum_{v\in g}\mathbf{u}_{igv}\mathbf{u}_{igv}^T)\mathbf{Z}_g^{-1}\big)}_{=\mathrm{tr}(\sum_{i=1}^{n}(\sum_{v\in g}u_{igv}u_{igv}^T)\mathbf{Z}_g^{-1})}\big\}\\
&\propto\ |\mathbf{Z}_g^{-1}|^{(c_1+nV_g-2-1)/2}\exp\big\{-\tfrac{1}{2}\mathrm{tr}[(c_1\Omega_1+\sum_{i=1}^{n}\sum_{v\in g}\mathbf{u}_{igv}\mathbf{u}_{igv}^T)\mathbf{Z}_g^{-1}]\big\}\\
&\propto\ \mathrm{Wishart}\big((c_1\Omega_1+\sum_{i=1}^{n}\sum_{v\in g}\mathbf{u}_{igv}\mathbf{u}_{igv}^T)^{-1},c_1+nV_g\big)
\end{aligned}
$$

(6)

$$
\begin{aligned}
[\boldsymbol{\Sigma}^{-1}|\boldsymbol{\Lambda}(\boldsymbol{\Sigma}^{-1}),\mathbf{Y}] &\propto [\boldsymbol{\Sigma}^{-1}]\cdot\prod_{v=1}^{V}\mathrm{N}(\rho\sum_{v'\neq v}\frac{w_{vv'}}{w_{v+}}\boldsymbol{\phi}_{v'},\frac{1}{w_{v+}}\boldsymbol{\Sigma})\ \ (\text{ Denote } b_{v'}:=\frac{w_{vv'}}{w_{v+}})\\
&\propto\ |\boldsymbol{\Sigma}^{-1}|^{(c_2-2-1)/2}\exp\big\{-\tfrac{1}{2}\mathrm{tr}(c_2\Omega_2\boldsymbol{\Sigma}^{-1})\big\}\cdot\\
&\quad\ \prod_{v=1}^{V}w_v|\boldsymbol{\Sigma}^{-1}|^{V/2}\exp\Big\{\tfrac{1}{2}\sum_{v=1}^{V}[\boldsymbol{\phi}_v-\rho\sum_{v'\neq v}b_{v'}\boldsymbol{\phi}_{v'}]^Tw_{v+}\boldsymbol{\Sigma}^{-1}[\boldsymbol{\phi}_v-\rho\sum_{v'\neq v}b_{v'}\boldsymbol{\phi}_{v'}]\Big\}\\
&\propto\ |\boldsymbol{\Sigma}^{-1}|^{(c_2+V-2-1)/2}\exp\big\{-\tfrac{1}{2}\mathrm{tr}\big[c_2\Omega_2+\\
&\quad\ \sum_{v=1^V}[w_{v+}^{1/2}(\boldsymbol{\phi}_v-\rho\sum_{v'\neq v}b_{v'}\boldsymbol{\phi}_{v'})][w_{v+}^{1/2}(\boldsymbol{\phi}_v-\rho\sum_{v'\neq v}b_{v'}\boldsymbol{\phi}_{v'})]^T\big]\boldsymbol{\Sigma}^{-1})\big\}\\
&\propto\ \mathrm{Wishart}\big((c_2\Omega_2+\sum_{v=1}^{V}w_{v+}[\boldsymbol{\phi}_v-\rho\sum_{v'\neq v}b_{v'}\boldsymbol{\phi}_{v'}][\boldsymbol{\phi}_v-\rho\sum_{v'\neq v}b_{v'}\boldsymbol{\phi}_{v'}]^T),c_2+V\big),
\end{aligned}
$$

where V is the total number of voxels included in the analysis (which have at least one within-region neighbor).

(7) Recall the notation: $\boldsymbol{\beta}_{gj}=\boldsymbol{\beta}_g^{(j)}$, where $j=1,2$ denote the baseline and follow-up

sessions. The independence assumption implies

$$\boldsymbol{\beta}_{gj} \sim \text{MVN}(\boldsymbol{\beta}_{0gj}, \lambda^2_{gj}),\ \boldsymbol{\beta}_g \sim \text{MVN}\Big(\boldsymbol{\beta}_{0g}, \underbrace{\mathbf{I}_{V_g} \otimes \begin{bmatrix} \lambda^2_{g1} & 0 \\ 0 & \lambda^2_{g2} \end{bmatrix}}_{\boldsymbol{\Lambda}_g}\Big).$$

Then

$$\begin{aligned}
&[\lambda^{-2}_{gj}|\boldsymbol{\Lambda}(\lambda^{-2}_{gj}), \mathbf{Y}] \propto [\lambda^{-2}_{gj}] \cdot [\boldsymbol{\beta}_{gj}|\lambda^{-2}_{gj}] \\
&\quad\propto \;\; \text{Gamma}(a_j, b_j) \cdot \text{Normal}(1_{V_g}\boldsymbol{\beta}_{ogj}, \lambda^2_{gj}\mathbf{I}_{V_g}) \\
&\quad\propto \;\; (\lambda^{-2}_{gj})^{a_j-1} \exp\Big(\frac{\lambda^{-2}_{gj}}{b_j}\Big)(\lambda^2_{gj})^{-V_g/2} \exp\Big\{-\frac{1}{2\lambda^2_{gj}}(\boldsymbol{\beta}_{gj} - \mathbf{1}_{V_g}\boldsymbol{\beta}_{0gj})^T(\boldsymbol{\beta}_{gj} - \mathbf{1}_{V_g}\boldsymbol{\beta}_{0gj})\Big\} \\
&\quad\propto \;\; (\lambda^{-2}_{gj})^{a_j+V_g-1} \exp\Big\{-\lambda^2_{gj}\Big[\frac{1}{b_j} + \frac{(\boldsymbol{\beta}_{gj} - \mathbf{1}_{V_g}\boldsymbol{\beta}_{0gj})^T(\boldsymbol{\beta}_{gj} - \mathbf{1}_{V_g}\boldsymbol{\beta}_{0gj})}{2}\Big]^{-1}\Big\} \\
&\quad\propto \;\; \text{Gamma}\Big(a_j + V_g/4, \Big[\frac{1}{b_j} + \frac{(\boldsymbol{\beta}_{gj} - \mathbf{1}_{V_g}\boldsymbol{\beta}_{0gj})^T(\boldsymbol{\beta}_{gj} - \mathbf{1}_{V_g}\boldsymbol{\beta}_{0gj})}{2}\Big]^{-1}\Big)
\end{aligned}$$

(8)

Let $\{\rho_1, \ldots, \rho_n\}$ be an array of values on which we specify a discrete uniform prior for ρ. ($n = 36$ in our analysis.) Let $L(\rho) = \prod_{v=1}^{V} \text{N}(\rho \sum_{v' \neq v} b_{v'}\boldsymbol{\phi}_{v'}, w^{-1}_{v+}\boldsymbol{\Sigma})$. Then

$$\begin{aligned}
L(\rho) &\propto \exp\Big\{-\tfrac{1}{2}\sum_{v=1}^{V}[\boldsymbol{\phi}_v - \rho\sum_{v'\neq v}\frac{w_{vv'}}{w_{v+}}\boldsymbol{\phi}_{v'}]^T w_{v+}\boldsymbol{\Sigma}^{-1}[\boldsymbol{\phi}_v - \rho\sum_{v'\neq v}\frac{w_{vv'}}{w_{v+}}\boldsymbol{\phi}_{v'}]\Big\} \\
&\propto \exp\Big\{-\tfrac{1}{2}\sum_{v=1}^{V}[\rho\underbrace{\sum_{v'\neq v}\frac{w_{vv'}}{w_{v+}}\boldsymbol{\phi}_{v'}}_{A} - \boldsymbol{\phi}_v]^T w_{v+}\boldsymbol{\Sigma}^{-1}[\rho\sum_{v'\neq v}\frac{w_{vv'}}{w_{v+}}\boldsymbol{\phi}_{v'} - \boldsymbol{\phi}_v]\Big\} \\
&\propto \exp\Big\{-\tfrac{1}{2}\sum_{v=1}^{V}(\rho A^T - \boldsymbol{\phi}_v^T)w_{v+}\boldsymbol{\Sigma}^{-1}(\rho A - \boldsymbol{\phi}_v)\Big\} \\
&\propto \exp\Big\{-\tfrac{1}{2}\sum_{v=1}^{V}(\rho A^T w_{v+}\boldsymbol{\Sigma}^{-1}\rho A - \boldsymbol{\phi}_v^T w_{v+}\boldsymbol{\Sigma}_{-1}\rho A - \rho A^T w_{v+}\boldsymbol{\Sigma}^{-1}\boldsymbol{\phi}_v\Big\} \\
&\propto \exp\Big\{-\tfrac{1}{2}\sum_{v=1}^{V}\rho^2 A^T w_{v+}\boldsymbol{\Sigma}^{-1}A - 2\rho\sum_{v=1}^{V}\boldsymbol{\phi}_v^T w_{v+}\boldsymbol{\Sigma}^{-1}A\Big\} \\
&\propto \exp\Big\{-\tfrac{1}{2}\rho^2\underbrace{\Big(\sum_{v=1}^{V}A^T w_{v+}\boldsymbol{\Sigma}^{-1}A\Big)}_{\eta^{-1}} - 2\rho\underbrace{\Big(\sum_{v=1}^{V}\boldsymbol{\phi}_v^T w_{v+}\boldsymbol{\Sigma}^{-1}A\Big)}_{\xi}\Big\}
\end{aligned}$$

where

$$\begin{aligned}
\eta &= \Big(\sum_{v=1}^{V}\big(\sum_{v'\neq v}\frac{w_{vv'}}{w_{v+}}\boldsymbol{\phi}_{v'}^T\cdot w_{v+}\boldsymbol{\Sigma}^{-1}\cdot\sum_{v'\neq v}\frac{w_{vv'}}{w_{v+}}\boldsymbol{\phi}_{v'}\Big)^{-1} \\
&= \Big(\sum_{v=1}^{V}\frac{1}{w_{v+}}M^T\boldsymbol{\Sigma}^{-1}M\Big)^{-1}, \quad M = \sum_{v'\neq v}w_{vv'}\boldsymbol{\phi}_{v'} \\
\xi &= \sum_{v=1}^{V}\boldsymbol{\phi}_v^T\boldsymbol{\Sigma}^{-1}M.
\end{aligned}$$

The corresponding set of normalized likelihood weights is then given by $\{p1, \ldots, p_n\}$, where

$$p_k = \frac{L(\rho_k)}{L(\rho_1) + \ldots + L(\rho_n)} = \frac{1}{\frac{L(\rho_1)}{L(\rho_k)} + \ldots + \frac{L(\rho_n)}{L(\rho_k)}},\ k = 1, \ldots, n.$$

For example, for $n = 1$

$$\begin{aligned}
\frac{L(\rho_1)}{L(\rho_k)} &= \frac{\exp\{-\frac{1}{2}(\rho_1^2\eta^{-1} - 2\rho_1\xi)\}}{\exp\{-\frac{1}{2}(\rho_k^2\eta^{-1} - 2\rho_k\xi)\}} \\
&= \exp\big\{-\tfrac{1}{2}(\rho_1^2\eta^{-1} - 2\rho_1\xi - \rho_k^2\eta^{-1} + 2\rho_k\xi)\big\} \\
&= \exp\big\{-\tfrac{1}{2}[(\rho_1^2 - \rho_k^2)\eta^{-1} - 2\xi(\rho_1 - \rho_k)]\big\} \\
&= \exp\big\{-\tfrac{1}{2}[(\rho_1 - \rho_k)(\rho_1 + \rho_k)\eta^{-1} - 2\xi(\rho_1 - \rho_k)]\big\} \\
&= \exp\big\{-\tfrac{1}{2}\big((\rho_1 - \rho_k)[(\rho_1 + \rho_k)\eta^{-1} - 2\xi]\big)\big\}.
\end{aligned}$$

(9) $[\boldsymbol{\gamma}_g|\boldsymbol{\Lambda}(\boldsymbol{\gamma}_g), \mathbf{Y}] =?$ Recall the notation: $\boldsymbol{\gamma}_g = [\gamma_{g1}^{(1)}, \gamma_{g1}^{(2)}, \ldots, \gamma_{gQ}^{(1)}, \gamma_{gQ}^{(2)}]^T$, where Q denotes the number of covariates in the model and the superscripts (1) and (2) denote the baseline and follow-up sessions. We assume independence between the covariates, i.e.,

$$\boldsymbol{\gamma}_g|\Lambda_{\gamma_g} \sim \text{MVN}\big(0, \text{diag}(\tau_{g11}^2, \tau_{g21}^2, \ldots, \tau_{g1Q}^2, \tau_{g2Q}^2)\big)$$

$$\begin{aligned}
&[\boldsymbol{\gamma}_g|\boldsymbol{\Lambda}(\gamma_g), \mathbf{Y}] \propto [\boldsymbol{\gamma}_g|\boldsymbol{\Lambda}(\boldsymbol{\Gamma}_g)] \cdot \prod_{i=1}^{n} \mathbf{Y}_{ig}|\boldsymbol{\beta}_g, \boldsymbol{\phi}_g, \boldsymbol{\alpha}_{ig}, \boldsymbol{\gamma}_g, \mathbf{V}_g \\
&\propto \quad \exp\big\{-\tfrac{1}{2}\boldsymbol{\gamma}_g^T\boldsymbol{\Lambda}_{\gamma_g}^{-1}\boldsymbol{\gamma}_g\big\} \cdot \\
&\qquad \exp\big\{-\tfrac{1}{2}\sum_{i=1}^{n}(\mathbf{Y}_{ig} - \boldsymbol{\beta}_g - \boldsymbol{\phi}_g - \mathbf{1}_{V_g} \otimes \boldsymbol{\alpha}_{ig} - \mathbf{1}_{V_g} \otimes \mathbf{X}_{ig}\boldsymbol{\gamma}_g)^T\mathbf{V}_g^{-1} \cdot \\
&\qquad (\mathbf{Y}_{ig} - \boldsymbol{\beta}_g - \boldsymbol{\phi}_g - \mathbf{1}_{V_g} \otimes \boldsymbol{\alpha}_{ig} - \mathbf{1}_{V_g} \otimes \mathbf{X}_{ig}\boldsymbol{\gamma}_g)\big\}.
\end{aligned}$$

The sum in the second exponential term can be written as

$$
\begin{aligned}
&\sum_{i=1}^{n}[\mathbf{1}_{V_g}\otimes \mathbf{X}_{ig}\boldsymbol{\gamma}_g - (\underbrace{\mathbf{Y}_{ig}-\boldsymbol{\beta}_g-\boldsymbol{\phi}_g-\mathbf{1}_{V_g}\otimes\boldsymbol{\alpha}_{ig})}_{:=\mathbf{u}_{ig}})]^T\mathbf{V}_g^{-1}[\mathbf{1}_{V_g}\otimes\mathbf{X}_{ig}\boldsymbol{\gamma}_g-\mathbf{u}_{ig}] \\
&= \sum_{i=1}^{n}\big[(\mathbf{1}_{V_g}\otimes\mathbf{X}_{ig}\boldsymbol{\gamma}_g)^T-\mathbf{u}_{ig}^T\big]\mathbf{V}_g^{-1}\big[(\mathbf{1}_{V_g}\otimes\mathbf{X}_{ig}\boldsymbol{\gamma}_g)-\mathbf{u}_{ig}\big] = \\
&\quad \Big[(\mathbf{1}_{V_g}\otimes\mathbf{X}_{ig}\boldsymbol{\gamma}_g)^T=\mathbf{1}_{V_g}^T\otimes\boldsymbol{\gamma}_g^T\mathbf{X}_{ig}^T=[\boldsymbol{\gamma}_g^T\mathbf{X}_{ig}^T\ldots\boldsymbol{\gamma}_g^T\mathbf{X}_{ig}^T]=\boldsymbol{\gamma}_g^T[\mathbf{X}_{ig}^t\ldots\mathbf{X}_{ig}^T]= \\
&\qquad \boldsymbol{\gamma}_g^T(\mathbf{1}_{V_g}^T\otimes\mathbf{X}_{ig}^T)\Big] \\
&= \sum_{i=1}^{n}\big[\boldsymbol{\gamma}_g^T(\mathbf{1}_{V_g}^T\otimes\mathbf{X}_{ig}^T)-\mathbf{u}_{ig}\big]\mathbf{V}_g^{-1}\big[\mathbf{1}_{V_g}\otimes\mathbf{X}_{ig}\boldsymbol{\gamma}_g-\mathbf{u}_{ig}\big] \\
&= \sum_{i=1}^{n}\big[\boldsymbol{\gamma}_g^T(\mathbf{1}_{V_g}^T\otimes\mathbf{X}_{ig})\mathbf{V}_g^{-1}(\mathbf{1}_{V_g}\otimes\mathbf{X}_{ig})\boldsymbol{\gamma}_g-2\boldsymbol{\gamma}_g^T(\mathbf{1}_{V_g}^T\otimes\mathbf{X}_{ig})\mathbf{V}_g^{-1}\mathbf{u}_{ig}\big] \\
&= \boldsymbol{\gamma}_g^T\Big(\sum_{i=1}^{n}(\mathbf{1}_{V_g}^T\otimes\mathbf{X}_{ig})\mathbf{V}_g^{-1}(\mathbf{1}_{V_g}\otimes\mathbf{X}_{ig})\Big)\boldsymbol{\gamma}_g-2\boldsymbol{\gamma}_g^T\sum_{i=1}^{n}(\mathbf{1}_{V_g}^T\otimes\mathbf{X}_{ig})\mathbf{V}_g^{-1}\mathbf{u}_{ig}
\end{aligned}
$$

Hence,

$$
\begin{aligned}
[\boldsymbol{\gamma}_g|\boldsymbol{\Lambda}(\gamma_g),\mathbf{Y}] \propto \exp\Big\{&-\tfrac{1}{2}\boldsymbol{\gamma}_g^T\underbrace{\big[\boldsymbol{\Lambda}_{\gamma_g}^{-1}+\sum_{i=1}^{n}(\mathbf{1}_{V_g}^T\otimes\mathbf{X}_{ig}^T)\mathbf{V}_g^{-1}(\mathbf{1}_{V_g}\otimes\mathbf{X}_{ig})\big]}_{\boldsymbol{\Omega}_{\gamma_g}^{-1}}\boldsymbol{\gamma}_g \\
&-2\boldsymbol{\gamma}_g^T\underbrace{\sum_{i=1}^{n}(\mathbf{1}_{V_g}^T\otimes\mathbf{X}_{ig}^T)\mathbf{V}_g^{-1}\mathbf{u}_{ig}}_{\mathbf{T}_{\gamma_g}}\Big\} \\
&\propto \mathrm{N}(\boldsymbol{\Omega}_{\gamma_g}\mathbf{T}_{\gamma_g},\boldsymbol{\Omega}_{\gamma_g})
\end{aligned}
$$

where $\boldsymbol{\Omega}_{\gamma_g}^{-1}=\boldsymbol{\Lambda}_{\gamma_g}^{-1}+\sum_{i=1}^{n}(\mathbf{1}_{V_g}^T\otimes\mathbf{X}_{ig}^T)\mathbf{V}_g^{-1}(\mathbf{1}_{V_g}\otimes\mathbf{X}_{ig})$, $\mathbf{T}_{\gamma_g}=\sum_{i=1}^{n}(\mathbf{1}_{V_g}^T\otimes\mathbf{X}_{ig}^T)\mathbf{V}_g^{-1}\mathbf{u}_{ig}$ and $\mathbf{u}_{ig}=\mathbf{Y}_{ig}-\boldsymbol{\beta}_g-\boldsymbol{\phi}_g-\mathbf{1}_{V_g}\otimes\boldsymbol{\alpha}_{ig}$.

Computational Details: $\boldsymbol{\Omega}_{\gamma_g}$ and $\mathbf{T}_{\gamma_g}$ can further be simplified for computational reasons.

$$\begin{aligned}\boldsymbol{\Omega}_{\gamma_g}^{-1} &= \boldsymbol{\Lambda}_{\gamma_g}^{-1} + \sum_{i=1}^{n}(\mathbf{1}_{V_g}^{T} \otimes \mathbf{X}_{ig}^{T})(\mathbf{I}_{V_g} \otimes \mathbf{Z}_g^{-1})(\mathbf{1}_{V_g} \otimes \mathbf{X}_{ig}) = \\ &= \boldsymbol{\Lambda}_{\gamma_g}^{-1} + V_g \sum_{i=1}^{n} \mathbf{X}_{ig}^{T}\mathbf{Z}_g^{-1}\mathbf{X}_{ig}\end{aligned}$$

and

$$\begin{aligned}\mathbf{T}_{\gamma_g} &= \sum_{i=1}^{n}[\mathbf{X}_{ig} \cdots \mathbf{X}_{ig}] \begin{bmatrix} \mathbf{Z}_g^{-1} & & \\ & \ddots & \\ & & \mathbf{Z}_g^{-1} \end{bmatrix} \mathbf{u}_{ig} = \\ &= \sum_{i=1}^{n}(\mathbf{1}_{V_g} \otimes \mathbf{X}_{ig}^{T}\mathbf{Z}_g^{-1})\mathbf{u}_{ig}\end{aligned}$$

(10)

$$\begin{aligned}[\tau_{gjq}^2 | \boldsymbol{\Lambda}(\tau_{gjq}^2), \mathbf{Y}] &\propto [\tau_{gjq}^2][\boldsymbol{\gamma}_{gjq} | \tau_{gjq}^2] \\ &\propto \mathrm{Gamma}(e_0, f_0) \cdot \mathrm{N}(0, \tau_{gjq}^2) \\ &\propto (\tau_{gjq}^{-2})^{e_0 - 1} \exp\Big(-\frac{\tau_{gjq}^{-2}}{f_0}\Big)(\tau_{gjq}^2)^{-1/2} \exp\Big\{-\tfrac{1}{2\tau_{gjq}^2}\boldsymbol{\gamma}_{gjq}^2\Big\} \\ &\propto (\tau_{gjq}^{-2})^{e_0 + 1/2 - 1} \exp\Big\{-\tau_{gjq}^{-2}\big[\tfrac{1}{f_0} + \tfrac{1}{2}\boldsymbol{\gamma}_{gjq}^2\big]\Big\} \\ &\propto \mathrm{Gamma}\Big(e_0 + \tfrac{1}{2}, \big[\tfrac{1}{f_0} + \tfrac{1}{2}\boldsymbol{\gamma}_{gjq}^2\big]^{-1}\Big)\end{aligned}$$

5.2.3 Appendix G: Additional results from the analysis of ADNI data

Regions included in analysis of ADNI data

We include 39 regions from the AAL map (Tzourio-Mazoyer et al., 2002) in our analysis of the PET data from the study Alzheimer's disease study (ADNI). The

No.	Region Name	V_g	No.	Region Name	V_g
1	Caudate L	877	21	Insula L	1898
2	Caudate R	950	22	Insula R	1760
3	Cingulum Ant L	1352	23	ParaHippocampal R	982
4	Cingulum Ant R	1352	24	Putamen L	976
5	Cingulum Mid L	1845	25	Putamen R	1058
6	Cingulum Mid R	2098	26	Rolandic Oper L	963
7	Frontal Inf Oper L	970	27	Rolandic Oper R	1278
8	Frontal Inf Oper R	1321	28	Temporal Mid L	4655
9	Frontal Inf Orb L	1537	29	Temporal Mid R	4337
10	Frontal Inf Orb R	1580	30	Temporal Pole Sup L	1189
11	Frontal Inf Tri L	2376	31	Temporal Pole Sup R	1171
12	Frontal Inf Tri R	1891	32	Temporal Sup L	2278
13	Frontal Sup L	3208	33	Temporal Sup R	3112
14	Frontal Sup R	3864	34	Thalamus L	1076
15	Frontal Mid L	4618	35	Thalamus R	1055
16	Frontal Mid R	4896	36	Cingulum Post L	466
17	Frontal Sup Medial L	2763	37	Cingulum Post R	281
18	Frontal Sup Medial R	2059	38	Precuneus L	3475
19	Hippocampus L	929	39	Precuneus R	2832
20	Hippocampus R	941			

Table 5.1: Selected 39 AAL regions of interest considered in the Alzheimer's disease study. V_g denotes the regions size, in voxels. Notation is as follows: R=right, L=left, Inf=inferior, Med=medial, Mid=middle, Orb=orbital, Oper=operculum, Post=posterior, Sup=superior, Temp=temporal, Tri=triangularis.

regions are listed in Table 5.2.3.

Figures 5.3 and 5.4 give the trace plot and the corresponding histogram plot of the posterior distribution for parameter $\boldsymbol{\beta}_g$. The figures show the trace plot and histogram for one selected voxel in region g.

Individualized prediction maps for subjects from ADNI data set. Figure 5.2.3 shows the individualized prediction maps for four additional AD patients from the test data set. Figure 5.2.3 shows the individualized prediction maps for two additional HC subjects from the test data set.

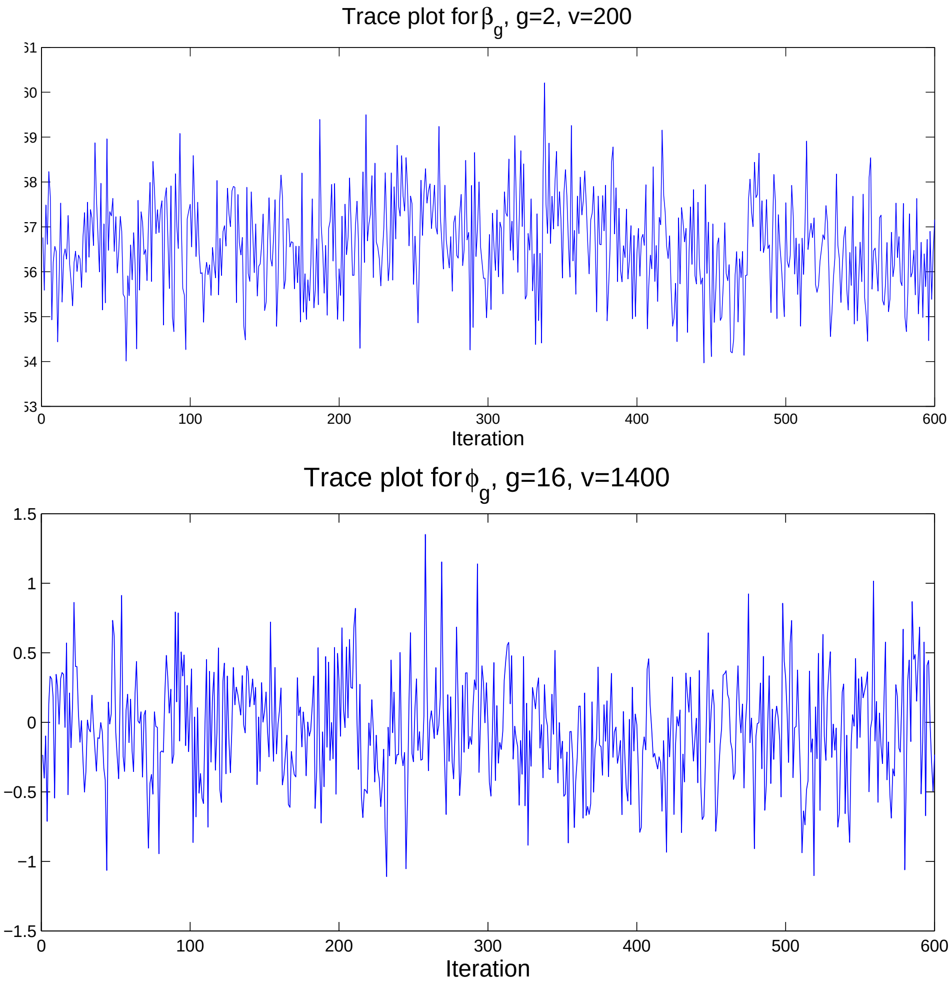


Figure 5.3: Trace plot for $\boldsymbol{\beta}_g$ parameter for a randomly selected voxel in region 2 (top) and for $\boldsymbol{\phi}_g$ parameter for a randomly selected voxel in region 16 (bottom).

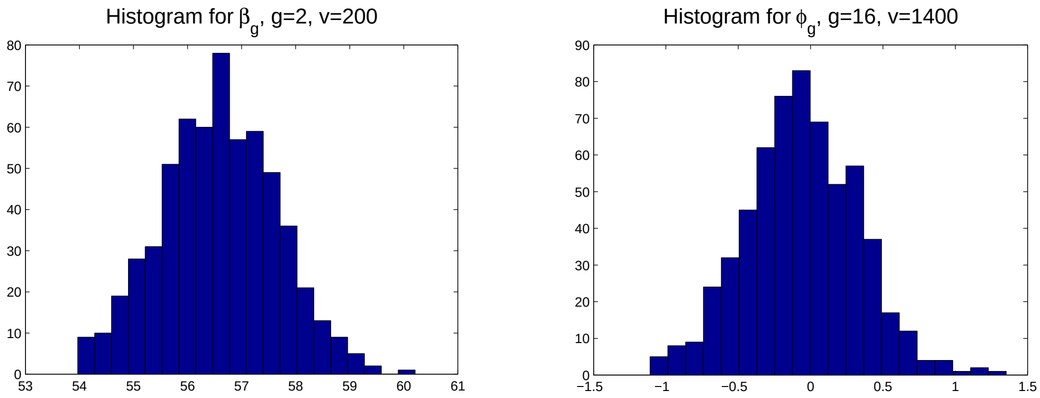
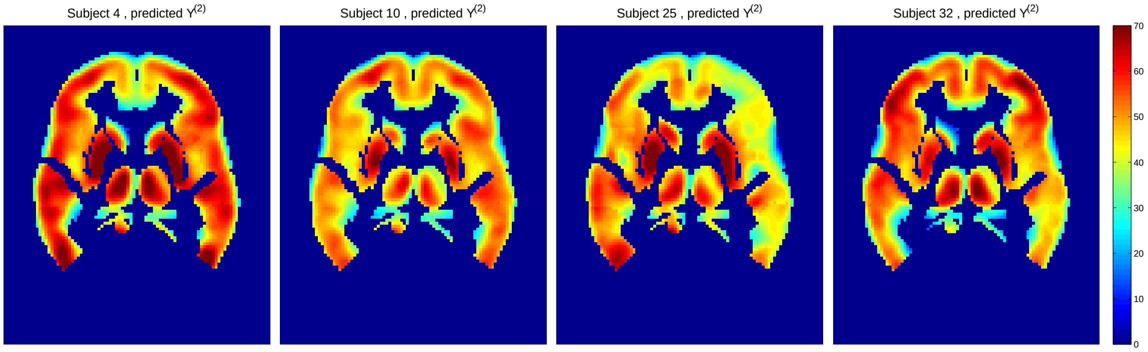


Figure 5.4: Histogram plot for $\boldsymbol{\beta}_g$, for a randomly selected voxel in region 2 (left) and for $\boldsymbol{\phi}_g$, for a randomly selected voxel in region 16 (right).

(a) predicted month 6 maps



(b) observed month 6 maps

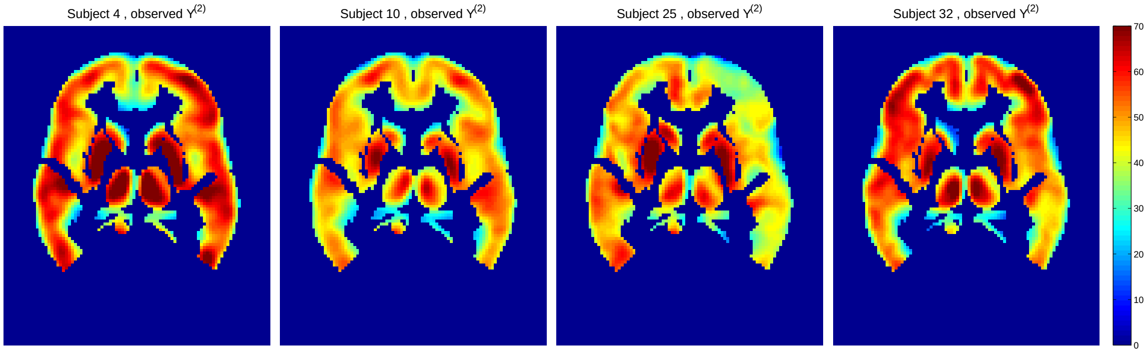
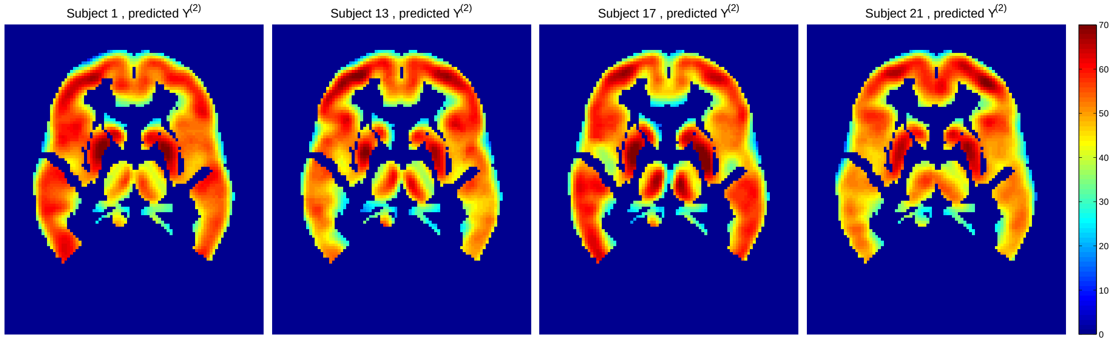


Figure 5.5: Individualized predicted and observed 6 month follow-up glucose uptake measurements for 4 AD patients from the test data set. Axial slice 40 is shown in radiological view.

(a) predicted month 6 maps



(b) observed month 6 maps

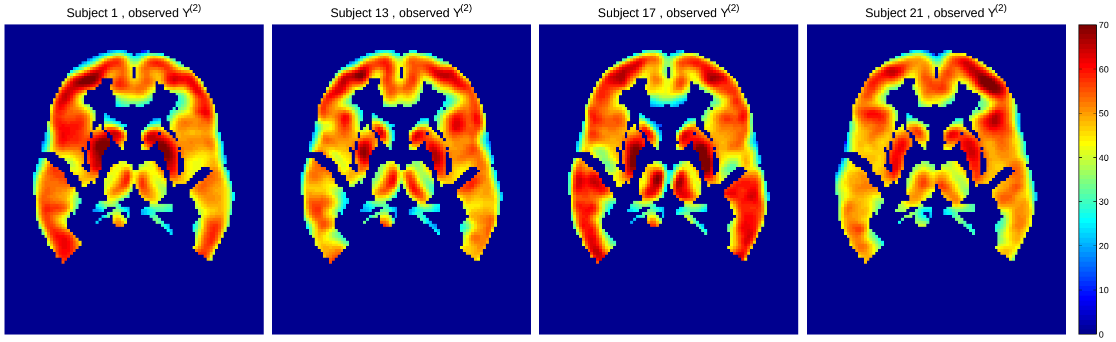


Figure 5.6: Individualized predicted and observed 6 month follow-up glucose uptake measurements for 4 HC subjects from the test data set. Axial slice 40 is shown in radiological view.

5.2.4 Appendix H: Simulation Results.

Tables 5.2 through 5.4 give results from the simulation study described in Section 4.5. Tables 5.2 and 5.3 list the estimated and the true values, as well as the bias, for parameter $\boldsymbol{\beta}$ from the mean of the top level of our model (BSPM). The results are shown for all 5 regions included in the simulation study and for 25 randomly selected voxels from each region. Table 5.2 gives the simulation results for the baseline session, and Table 5.3 for the follow-up session. Table 5.4 gives the simulation results for the random effect parameters $\boldsymbol{\alpha}_i$, for baseline session (results for the follow-up session are similar).

Voxel	**Region**														
	1			**2**			**3**			**4**			**5**		
	Est.	True	Bias	Est.	True	Bias	Est.	True	Bias	Est.	True	Bias	Est.	True	Bias
1	14.42	12.95	1.47	20.50	20.93	-0.43	37.26	34.18	3.08	62.55	60.16	2.39	63.848	63.02	0.83
2	11.53	10.35	1.18	15.77	15.32	0.45	31.06	28.66	2.40	63.44	61.75	1.69	62.2453	61.95	0.29
3	14.07	12.20	1.87	15.63	17.00	-1.37	22.97	22.11	0.86	63.86	62.62	1.24	60.8592	59.62	1.24
4	14.00	13.61	0.39	22.71	22.48	0.23	24.48	22.93	1.55	64.32	62.90	1.42	56.8206	55.89	0.93
5	16.20	14.53	1.67	26.29	27.40	-1.11	31.97	29.74	2.23	65.45	62.90	2.55	51.2923	50.55	0.75
6	16.53	14.95	1.57	30.89	31.70	-0.81	37.20	35.21	1.99	64.03	62.63	1.40	43.5581	43.90	-0.34
7	16.36	14.89	1.47	32.75	35.23	-2.48	41.85	39.72	2.13	63.24	61.78	1.46	38.7694	36.57	2.20
8	13.88	12.01	1.87	36.43	37.82	-1.38	45.41	43.36	2.05	61.25	59.88	1.37	28.7784	29.39	-0.61
9	15.14	13.85	1.29	37.20	39.40	-2.20	47.60	46.22	1.38	57.60	56.55	1.05	23.6408	22.79	0.85
10	16.76	15.28	1.48	39.33	40.21	-0.88	49.55	48.27	1.28	53.41	51.47	1.95	24.191	23.43	0.76
11	17.50	16.21	1.29	40.13	40.49	-0.35	51.18	49.35	1.83	45.85	44.67	1.17	31.0402	30.27	0.77
12	18.51	16.61	1.90	38.96	40.43	-1.48	52.09	49.29	2.80	38.29	36.73	1.55	37.2314	37.14	0.10
13	17.16	16.52	0.65	38.40	39.99	-1.59	49.76	48.15	1.62	29.10	27.93	1.17	43.9114	43.91	0.00
14	16.33	16.00	0.33	38.52	39.09	-0.57	47.91	46.23	1.68	18.60	17.88	0.71	49.7671	49.87	-0.10
15	13.03	11.05	1.98	35.81	37.73	-1.93	46.02	43.64	2.38	15.45	15.00	0.46	54.8177	54.54	0.28
16	14.54	13.37	1.17	35.99	36.42	-0.43	42.07	40.23	1.84	26.76	24.97	1.78	57.9509	57.83	0.13
17	17.48	15.28	2.21	33.89	35.43	-1.54	38.37	36.00	2.37	35.16	34.31	0.85	59.7718	59.77	0.00
18	18.33	16.75	1.58	33.51	34.70	-1.19	33.39	31.07	2.32	44.32	43.18	1.14	60.6768	60.53	0.15
19	19.36	17.63	1.73	28.43	28.92	-0.49	43.93	41.93	2.00	51.92	51.13	0.79	61.4144	60.45	0.96
20	20.34	17.95	2.38	28.12	28.75	-0.63	48.47	45.31	3.16	58.40	57.33	1.07	58.9549	59.75	-0.79
21	18.98	17.82	1.15	30.45	31.15	-0.70	48.67	47.71	0.96	62.89	61.56	1.33	59.6355	58.62	1.01
22	18.32	17.39	0.93	33.73	33.19	0.54	50.91	49.18	1.73	65.36	64.22	1.14	58.7166	57.38	1.33
23	17.97	16.74	1.22	33.61	34.58	-0.98	51.78	49.88	1.90	67.10	65.87	1.23	58.1904	56.33	1.86
24	12.08	9.82	2.26	32.62	35.45	-2.83	52.50	50.08	2.42	68.38	66.77	1.61	57.539	55.69	1.84
25	11.00	8.82	2.17	34.55	36.15	-1.60	52.01	49.65	2.36	69.19	67.04	2.15	56.4307	55.49	0.94

Table 5.2: Estimated and true values of $\boldsymbol{\beta}$ for selected 25 voxels from each region in the Simulation study, baseline session.

Voxel	**Region**														
	1			**2**			**3**			**4**			**5**		
	est.	true	bias	est.	true	bias	est.	true	bias	est.	true	bias	est.	true	bias
1	27.03	27.95	-0.92	25.75	27.44	-1.69	28.45	28.89	-0.44	53.78	52.62	1.16	64.42	62.24	2.18
2	24.21	24.78	-0.57	19.41	20.68	-1.27	25.25	25.52	-0.26	56.05	54.74	1.31	62.84	60.55	2.29
3	26.88	27.54	-0.66	19.00	20.78	-1.78	21.33	21.58	-0.25	56.73	56.09	0.63	60.00	57.42	2.58
4	27.70	29.37	-1.67	25.02	26.09	-1.07	19.45	20.48	-1.03	58.06	56.92	1.14	55.33	52.88	2.45
5	28.82	30.51	-1.69	29.41	31.40	-1.99	25.04	25.45	-0.41	59.35	57.22	2.13	49.50	47.13	2.38
6	31.20	30.97	0.23	34.53	36.33	-1.80	29.70	29.76	-0.05	57.62	56.93	0.69	42.53	40.36	2.17
7	30.84	31.01	-0.17	39.39	40.68	-1.28	32.90	33.76	-0.86	57.02	56.00	1.02	35.73	33.08	2.64
8	26.63	26.94	-0.31	43.23	44.08	-0.85	36.46	37.22	-0.75	55.11	54.23	0.88	27.71	26.12	1.59
9	28.94	30.01	-1.07	44.24	46.48	-2.23	38.98	39.91	-0.93	52.32	51.35	0.97	22.87	20.04	2.83
10	31.13	31.91	-0.79	45.54	47.89	-2.35	40.45	41.49	-1.04	48.55	47.10	1.46	24.68	22.83	1.85
11	32.47	33.12	-0.66	46.52	48.36	-1.84	41.00	42.09	-1.09	40.99	41.35	-0.36	30.90	28.54	2.36
12	33.36	33.84	-0.48	46.82	48.13	-1.31	41.22	41.70	-0.48	35.52	34.33	1.19	37.87	35.14	2.72
13	33.04	34.01	-0.97	46.10	47.56	-1.46	39.51	40.36	-0.85	27.85	26.79	1.05	44.23	41.94	2.29
14	32.24	33.64	-1.40	45.65	46.94	-1.30	37.69	38.13	-0.45	20.79	19.45	1.34	50.70	48.00	2.70
15	23.91	24.66	-0.75	43.79	46.22	-2.43	34.78	35.12	-0.34	16.91	15.42	1.48	54.86	52.81	2.04
16	27.62	28.20	-0.58	44.54	45.51	-0.96	31.34	31.54	-0.20	24.42	23.04	1.38	58.25	56.41	1.84
17	31.38	31.56	-0.18	42.73	45.14	-2.41	27.01	27.45	-0.44	31.57	30.88	0.69	60.90	58.96	1.93
18	33.26	34.03	-0.76	43.25	45.18	-1.93	22.57	23.41	-0.84	40.01	38.65	1.36	62.62	60.68	1.94
19	34.78	35.66	-0.88	40.01	42.14	-2.13	32.96	33.31	-0.34	47.47	45.91	1.56	64.05	61.67	2.38
20	35.72	36.59	-0.87	39.34	40.67	-1.33	36.64	36.62	0.01	52.68	51.95	0.73	64.14	62.06	2.08
21	35.51	36.88	-1.37	41.05	42.81	-1.75	38.69	39.48	-0.80	57.65	56.27	1.38	64.32	61.84	2.49
22	35.06	36.65	-1.59	43.62	44.40	-0.79	40.24	41.56	-1.31	60.34	59.15	1.19	63.64	61.29	2.34
23	34.44	35.82	-1.38	43.70	45.22	-1.52	42.07	42.68	-0.61	62.10	60.80	1.30	63.28	60.74	2.54
24	33.23	34.33	-1.09	43.23	45.72	-2.49	42.22	42.76	-0.53	62.76	61.66	1.10	61.83	60.39	1.45
25	29.40	30.23	-0.83	44.92	46.33	-1.41	41.82	41.88	-0.07	63.18	62.09	1.09	62.93	60.26	2.67

Table 5.3: Estimated and true values of $\boldsymbol{\beta}$ for selected 25 voxels from each region in the Simulation study, follow-up session.

Voxel	Region														
	1			**2**			**3**			**4**			**5**		
	est.	true	bias	est.	true	bias	est.	true	bias	est.	true	bias	est.	true	bias
1	5.63	6.65	-1.02	9.94	8.46	1.48	-1.18	1.30	-2.48	-1.68	0.00	-1.68	-1.9278	-1.20	-0.72
2	-21.20	-23.68	2.48	1.39	0.14	1.25	5.91	-4.67	10.58	-0.47	1.75	-2.22	-2.7562	-2.38	-0.37
3	-29.16	-19.74	-9.43	-8.01	-10.56	2.55	9.95	-7.40	17.34	2.54	3.21	-0.67	-0.4261	1.47	-1.90
4	14.32	23.95	-9.63	5.23	2.75	2.48	-8.01	10.48	-18.49	-2.46	-2.00	-0.46	1.332	3.17	-1.84
5	-0.57	-19.71	19.14	-5.89	-4.52	-1.37	-0.28	-1.09	0.81	-1.07	2.22	-3.29	-1.3073	-3.92	2.61
6	-8.26	-8.63	0.38	10.00	8.93	1.08	2.95	-1.88	4.83	1.52	2.96	-1.44	0.2527	0.57	-0.32
7	19.37	29.75	-10.38	4.19	1.57	2.62	-1.42	4.04	-5.46	-1.54	-0.94	-0.60	-2.9279	-0.97	-1.95
8	1.20	-5.15	6.34	4.25	4.72	-0.47	1.52	-1.98	3.50	0.59	1.61	-1.02	0.6604	-0.60	1.26
9	2.38	3.00	-0.63	3.98	3.47	0.51	-1.96	2.47	-4.44	-0.74	-0.47	-0.26	-0.9299	-1.14	0.21
10	-17.23	-16.06	-1.16	-16.26	-18.02	1.76	-1.32	3.08	-4.40	0.81	2.55	-1.74	1.7195	2.54	-0.82
11	14.58	17.68	-3.09	-4.06	-5.02	0.96	-0.35	1.31	-1.66	0.99	1.29	-0.30	0.8902	1.05	-0.16
12	5.86	0.07	5.79	2.11	2.16	-0.05	-1.14	1.09	-2.22	1.15	2.67	-1.52	2.4952	1.66	0.84
13	17.38	19.58	-2.19	-4.90	-6.78	1.88	-7.64	9.52	-17.16	-0.75	1.10	-1.85	3.0931	4.14	-1.05
14	-17.23	-7.53	-9.70	-8.57	-10.88	2.31	3.06	-0.75	3.80	0.28	0.35	-0.07	-0.2636	1.35	-1.62
15	12.86	17.16	-4.29	6.13	3.49	2.63	0.44	2.20	-1.76	1.32	3.83	-2.51	0.974	2.85	-1.88

Table 5.4: Estimated and true values of α_i (random effects), baseline session.

Bibliography

Anoop, A., Singh, P. S., Jacob, R. S., and Maji, S. K. (2010). CSF biomarkers for Alzheimer's disease diagnosis. *International Journal of Alzheimer's Disease* doi:10.4061/2010/606802.

Aron, A. R., Behrens, T. E., Smith, S., Frank, M. J., and Poldrack, R. A. (2007). Triangulating a cognitive control network using diffusion-weighted magnetic resonance imaging (MRI) and functional MRI. *The Journal of Neuroscience* **14,** 3743–3752.

Aron, A. R., Durston, S., Eagle, D. M., Logan, G. D., Stinear, C. M., and Stuphorn, V. (2007a). Converging evidence for a fronto-basal-ganglia network for inhibitory control of action and cognition. *The Journal of Neuroscience* **27,** 11860–11864.

Aron, A. R. and Poldrack, R. A. (2006). Cortical and subcortical contributions to stop signal response inhibition: role of the subthalamic nucleus. *Journal of Neuroscience* **26,** 2424–2433.

Balslev, D., Nielsen, F., Frutiger, S., Sidtis, J., Christiansen, T., Svarer, C., Strother, S., Rottenberg, D., Hansen, L., Paulson, O., and Law, I. (2002). Cluster analysis of activity-time series in motor learning. *Human Brain Mapping* **15,** 135–145.

Banerjee, S., Gelfand, A., and Carlin, B. (2004). *Hierarchical Modeling and Analysis for Spatial Data.* Chapman & Hall/CRC.

Bartholomew, D. (1987). *Latent Variable Models and Factor Analysis.* Charles Griffin & Co. Ltd, London.

Bassett, S., Yousem, D., Christinzio, C., Kusevic, I., Yassa, M., Caffo, B., and Zeger, S. (2006). Familiar risk for Alzheimers disease alters fMRI activation patterns. *Brain* **129,** 1229 –1239.

Baumgartner, R., Ryner, L., Richter, W., Summers, R., Jarmasz, M., and Somorjai, R. (2000). Comparison of two exploratory data analysis methods for fMRI: fuzzy clustering vs. principal component analysis. *Magnetic Resonance Imaging* **18,** 89–94.

Beckmann, C. and Smith, S. (2004). Probabilistic independent component analysis for functional magnetic resonance imaging. *IEEE Transactions on Medical Imaging* **23,** 137–152.

Beckmann, C. and Smith, S. (2005). Tensorial extensions of independent component analysis for multisubject FMRI analysis. *NeuroImage* **25,** 294–311.

Bell, A. and Sejnowski, T. (1995). An information maximisation approach to blind separation and blind deconvolution. *Neural Computation* **7,** 1129–1159.

Benali, H., Buvat, I., Anton, J. L., Pélégrini, M., Di Paola, M., Bittoun, J., Burnod, Y., and Di Paola, R. (1997). Space-time statistical model for functional MRI image sequences. In Duncan, J. and Gindi, G., editors, *Proceedings of the 15th International Conference on Information Processing in Medical Image*, volume 1230 of *Lecture Notes in Computer Science*, pages 285–298. Springer-Verlag London, UK.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)* **57,** 289–300.

Besag, J. (1974). Spatial interaction and the spatial analysis of lattice systems (with discussion). *J. R. Statist. Soc. B* **36,** 192–236.

Besag, J., Green, P., Higdon, D., and Mengersen, K. (1995). Bayesian computation and stochastic systems. *Statistical Science* **10,** 3–66.

Besag, J., York, J., and Molie, A. (1991). Bayesian image restoration with two applications in spatial statistics (with discussion). *Annals of the Institute of Statistical Mathematics* **43,** 1–59.

Bolla, K. I., Eldreth, D. A., London, E. D., Kiehl, K. A., Mouratidis, M., Contoreggi, C., Matochik, J. A., Kurian, V., Kimes, A. S., and M., E. (2003). Orbitofrontal cortex dysfunction in abstinent cocaine abusers performing a decision-making task. *NeuroImage* **19,** 1085–1094.

Bowman, F. (2005). Spatio-temporal modeling of localized brain activity. *Biostatistics* **6,** 558–575.

Bowman, F., Caffo, B., Bassett, S., and Kilts, C. (2008). A Bayesian hierarchical framework for spatial modeling of fMRI data. *NeuroImage* **39,** 146–156.

Bowman, F. and Kilts, C. (2003). Modeling intra-subject correlation among repeated scans in positron emission tomography (PET) neuroimaging data. *Human Brain Mapping* **20,** 59–70.

Bowman, F. and Patel, R. (2004). Identifying spatial relationships in neural processing using a multiple classification approach. *NeuroImage* **23,** 260–268.

Bowman, F., Patel, R., and Lu, C. (2004). Methods for detecting functional classifications in neuroimaging data. *Human Brain Mapping* **23,** 109–119.

Brodmann, K. (1909a). *Beschreibung der einzelnen Hirnkarten, IV. Kapitel in Vergleichende Lokalisationslehre der Grosshirnrinde.* (Leipzig: Verlag von Johann Ambrosias Barth).

Brodmann, K. (1994[original publication date 1909]b). *"Vergleichende Lokalisationlehre der Grosshirnrinde in ihren Prinzipien Dargestellt auf Grund des Zellenbaues," translated by Garey, L.J. as Brodmann's 'Localisation in the Cerebral Cortex'.* Smith-Gordon, London.

Buerger, K., Zinkowski, R., Teipel, S. J., Tapiola, T., Arai, H., Blennow, K., Andreasen, N., Hofmann-Kiefer, K., DeBernardis, J., Kerkman, D., McCulloch, C., Kohnken, R., Padberg, F., Pirttilä, T., Schapiro, M. B., Rapoport, S. I., Möller, H., Davies, P., and Hampel, H. (2002). Differential diagnosis of Alzheimer disease with cerebrospinal fluid levels of tau protein phosphorylated at threonine 231. *Archives of Neurology* **59,** 1267–1272.

Bullmore, E., Brammer, M., Williams, S., Rabe-Hesketh, S., Janot, N., David, A., Mellers, J., Howard, R., and Sham, P. (1996). Statistical methods of estimation and inference for functional MR image analysis. *Magnetic Resonance in Medicine* **35,** 261–277.

Calhoun, V., Adali, T., Hansen, L. K., Larsen, J., and Pekar, J. (2003). ICA of functional MRI data: an overview. In *Fourth International Symposium on Independent Component Analysis and Blind Source Separation, Nara, Japan.*, pages 281–288.

Calhoun, V., Adali, T., Pearlson, G., and Pekar, J. (2001). Spatial and temporal independent component analysis of functional MRI data containing a pair of task-related waveforms. *Human Brain Mapping* **13,** 43–53.

Calhoun, V. and Pekar, J. (2000). *When and where are components independent? On the applicability of spatial- and temporal- ICA to functional MRI data*, volume 11.

Sixth International Conference on Functional Mapping of the Human Brain., Neuroimage, Academic Press.

Carlin, B. and Banerjee, S. (2003). Hierarchical multivariate CAR models for spatially correlated survival data. In *Bayesian Statistics 7*, pages 45–64. Oxford: Oxford University Press.

Cattaneo, Z., Silvanto, J., Pascual-Leone, A., and Battelli, L. (2009). The role of the angular gyrus in the modulation of visuospatial attention by the mental number line. *NeuroImage* **44,** 563–568.

Chiu, K.-C., Liu, Z.-Y., and Xu, L. (2003). A statistical approach to testing mutual independence of ica recovered sources. In *Proceedings of 4th International Symposium on Independent Component Analysis and Blind Signal Separation (ICA2003), Japan.*

Clayton, D. and Kaldor, J. (1987). Empirical bayes estimates of age-standardized relative risks for use in disease mapping. *Biometrics* **43,** 671–681.

Cliff, A. and Ord, J. (1973). *Spatial Autocorrelation.* Pion Limited, London.

Cliff, A. and Ord, J. (1981). *Spatial Processes. Models and Applications.* London: Pion.

Common, P. (1994). Independent component analysis, a new concept? *Signal Processing* **36,** 287–314.

Cordes, D., Haughton, V., Carew, J., Arfanakis, K., and Maravilla, K. (2002). Hierarchical clustering to measure connectivity in fMRI resting-state data. *Magnetic Resonance Imaging* **20,** 305–317.

Cox, D. and Savoy, R. (2003). Functional magnetic resonance imaging (fMRI) "brain reading": Detecting and classifying distributed patterns of fMRI activity in human visual cortex. *NeuroImage* **19,** 261–270.

Cressie, N. A. C. (1993). *Statistics for Spatial Data.* New York: Wiley.

De Meyer, G., Shapiro, F., Vanderstichele, H., Vanmechelen, E., Engelborghs, S., De Deyn, P., Coart, E., Hansson, O., Minthon, L., Zetterberg, H., Blennow, K., Shaw, L., and Trojanowski, J. Q. (2010). Diagnosis-independent Alzheimer disease biomarker signature in cognitively normal elderly people. *Archives of Neurology* **67,** 949–956.

Derado, G., Bowman, F., and Kilts, C. (2010). Modeling the spatial and temporal dependence in fMRI data. *Biometrics* **66,** 949 –957.

Derado, G., Bowman, F. D., Ely, T., and Kilts, C. D. (2010a). Evaluating functional autocorrelation within spatially distributed neural processing networks. *Statistics and Its Interface* **3,** 45–48.

D'Esposito, M. (2001). *Handbook of Functional Neuroimaging of Cognition*, chapter Functional Neuroimaging of Working Memory, pages 293–327. The MIT Press, Cambridge, MA.

D'Esposito, M., Aguirre, G., Zarahn, E., and Ballard, D. (1998). Functional MRI studies of spatial and non-spatial working memory. *Cognitive Brain Research* **7,** 1–13.

Efron, B. and Tibshirani, R. J. (1998). *An Introduction to the Bootstrap.* Chapman & Hall.

Esposito, F., Formisano, E., Seifritz, E., Goebel, R., Morrone, R., Tedeschi, G., and Salle, F. (2002). Spatial independent component analysis of functional MRI time-series: to what extent do results depend on the algorithm used? *Human Brain Mapping* **16,** 146–157.

Evans, A., Collins, D., Mills, S., Brown, E., Kelly, R., and Peters, T. (1993). 3d statistical neuroanatomical models from 305 mri volumes. In *Proc. IEEE Nucl. Sci. Symp. Med. Imaging Conf.*, pages 1813–1817.

Evans, K., Dougherty, D., Pollack, M., and Rauch, S. (2006). Using neuroimaging to predict treatment response in mood and anxiety disorders. *Ann Clin Psychiatry* **18,** 33–42.

Fadili, M., Ruan, S., Bloyet, D., and Mazoyer, B. (2001). On the number of clusters and the fuzziness index for unsupervised FCA application to BOLD fMRI time series. *Medical Image Analysis* **5,** 55–67.

Fischl, B., Salat, D., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., van der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., and Dale, A. M. (2002). Whole brain segmentation. Automated labeling of neuroanatomical structures in the human brain. *Neuron* **33,** 341–355.

Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Ségonne, F., Salat, D., Busa, E., Seidman, L., Goldstein, J., Kennedy, D., Caviness, V., Makris, N., Rosen, B., and Dale, A. M. (2004). Automatically parcellating the human cerebral cortex. *Cerebral Cortex* **14,** 11–22.

Fisher, R. (1936). The use of multiple measurements in taxonomic problems. *Annals of Eugenics* **7,** 179–188.

Friston, K., Frith, C., Liddle, P., and Frackowiak, R. (1993). Functional connectivity: the principal component analysis of large data sets. *Journal of Cerebral Blood Flow & Metabolism* **13,** 5–14.

Friston, K., Glaser, D. E., Henson, R. N., Kiebel, S., Phillips, C., and Ashburner, J. (2002). Classical and Bayesian inference in neuroimaging: applications. *Neuroimage* **16,** 484–512.

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J.-P., Frith, C. D., and Frackowiak, R. S. J. (1995). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping* **2,** 189–210.

Garavan, H. and Hester, R. (2007). The role of cognitive control in cocaine dependence. *Neuropsychology Review* **17,** 337–345.

Gelfand, A. E. and Vounatsou, P. (2003). Proper multivariate conditional autoregressive models for spatial data analysis. *Biostatistics* **4,** 11–25.

Gelman, A., Carlin, J. B., Stern, S., H., and Rubin, D. B. (2004). *Bayesian Data Analysis.* Chapman and Hall/CRC, 2nd edition.

Giessing, C., Fink, G., Rösler, F., and Thiel, C. (2007). fMRI Data predict individual differences of behavioral effects of nicotine: A partial least square analysis. *Journal of Cognitive Neuroscience* **19,** 658–670.

Goldapple, K., Segal, Z., Garson, C., Lau, M., Bieling, P., Kennedy, S., and Mayberg, H. (2004). Modulation of cortical-limbic pathways in major depression: treatment specific effects of cbt. *Archives of General Psychiatry* **61,** 34–41.

Goldstein, R. Z. and Volkow, N. D. (2002). Drug addiction and its underlying basis: Neuroimaging evidence for the involvement of the frontal cortex. *American Journal of Psychiatry* **159,** 1642–1652.

Goodchild, M. (1986). *Spatial Autocorrelation.* CATMOG.

Gössl, C., Auer, D., and Fahrmeir, L. (2001). Bayesian spatiotemporal inference in functional magnetic resonance imaging. *Biometrics* **57,** 554–562.

Gray, H. (1918). *Anatomy of the Human Body.* Philadelphia: Lea & Febiger. Bartleby.com, 2000, `www.bartleby.com`.

Guo, Y., Bowman, F., and Kilts, C. (2008). Predicting the brain response to treatment using a Bayesian Hierarchical model with application to a study of schizophrenia. *Human Brain Mapping* **29,** 1092–1109.

Guo, Y. and Pagnoni, G. (2008). A unified framework for group independent component analysis for multi-subject fMRI data. *NeuroImag* **42,** 1078–1093.

Harshman, R. (1970). Foundations of the PARAFAC procedure: models and conditions for an "exploratory" multimodal factor analysis. *UCLA Working Papers in Phonetics* **16,** 1–84.

Harshman, R. and Lundy, M. (1994). Parafac: parallel factor analysis. *Computational Statistics and Data Analysis* **18,** 39–72.

Hartigan, J. (1977). *Classification and Clustering*, chapter Distribution Problems in Clustering. New York: Academic Press, Inc.

Hartigan, J. (1978). Asymptotic distributions for clustering criteria. *Annals of Statistic* **6,** 117–131.

Hartvig, N. V. (2002). A Stochastic Geometry Model for Functional Magnetic Resonance Images. *Scandinavian Journal of Statistics* **29,** 333–353.

Herault, J. and Jutten, C. (1986). Space or time adaptive signal processing by neural network model. In Denker, J. S., editor, *Neural Networks for Computing. Proceedings of AIP Conference*, pages 206–211. American Institute of Physics, New York.

Hertz, J., Krogh, A., and Palmer, R. (1991). *Introduction to the Theory of Neural Computation.* Redwood City: Addison-Wesley.

Higgins, J. (1996). *Brain imaging in psychiatry*, chapter Normal brain anatomy imaged by CT and MRI, pages 63–107. Oxford: Blackwell Science.

Himberg, J. and Hyvärinen, A. (2003). Icasso: software for investigating the reliability of ICA estimates by clustering and visualization. In *IEEE Workshop on Neural Networks for Signal Processing (NNSP2003), Toulouse, France*, pages 259–268.

Himberg, J., Hyvärinen, A., and Esposito, F. (2004). Validating the independent components of neuroimaging time series via clustering and visualization. *NeuroImage* **22,** 1214–1222.

Hoeft, F., Meyler, A., Glover, G., Kobayashi, N., Mazaika, P., and et.al (2007). Prediction of childrens reading skills using behavioral, functional, and structural neuroimaging measures. *Behavioral Neuroscience* **121,** 602–613.

Horn, N. R., Dolan, M., Elliot, R., Deakin, J. F. W., and Woodruff, P. W. R. (2003). Response inhibition and impulsivity: an fMRI study. *Neuropsychologia* **41,** 1959–1966.

Hyvärinen, A. (1999). Fast and robust fixed-point algorithms for independent component analysis. *IEEE Transactions on Neural Networks* **10,** 626–634.

Jin, X., B.P., C., and Banerjee, S. (2005). Generalized hierarchical multivariate CAR models for areal data. *Biometrics* **61,** 950–961.

Kaufman, J. N., Ross, T. J., Stein, E. A., and Garavan, H. (2003). Cingulate hypoactivity in cocaine users during a GO-NOGO task as revealed by event-related functional magnetic resonance imaging. *The Journal of Neuroscience* **23,** 7839–7843.

LaConte, S., Strother, S., Cherkassky, V., Anderson, J., and Hu, X. (2005). Support vector machines for temporal classification of block design fMRI data. *NeuroImage* **26,** 317–329.

Laird, N. M. and Ware, J. H. (1998). Random-effects models for longitudinal data. *Biometrics* **38,** 963–974.

Mardia, K. V. (1988). Multi-dimensional Multivariate Gaussian Random Markov Fields with Application to Image Processing. *Journal of Multivariate Analysis* **24,** 265–284.

Marquand, A., M., H., Brammer, M., Chu, C., Coen, S., and Mourão Miranda, J. (2010). Quantitative prediction of subjective pain intensity from whole-brain fMRI data using Gaussian processes. *NeuroImage* **49,** 2178–89.

Marshall, R. J. (1991). A review of methods for the statistical analysis of spatial patterns of disease. *J. R. Statist. Soc. A* **154,** 421–441.

Mayberg, H. (1997). Limbic-cortical dysregulation: A proposed model of depression. *Journal of Neuropsychiatry and Clinical Neurosciences* **9,** 471–481.

McCullagh, P. and Nelder, J. (1989). *Generalized Linear Models.* Boca Raton: Chapman and Hall/CRC, 2nd edition.

McIntosh, A., Bookstein, F., Haxby, J., and et al. (1996). Spatial pattern analysis of functional brain images using partial least squares. *NeuroImage* **3,** 143–157.

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

McKeown, M., Makeig, S., Brown, G., Jung, T., Kindermann, S., Bell, A., and Sejnowski, T. (1997). Analysis of fMRI data by decomposition into independent components. *Neurology* **48,** 6056.

McKeown, M., Makeig, S., Brown, G., Jung, T., Kindermann, S., Bell, A., and Sejnowski, T. (1998). Analysis of fMRI data by blind separation into independent spatial components. *Human Brain Mapping* **6,** 160–188.

Mitchell, T., Hutchinson, R., Niculescu, R., Pereiera, F., and Wang, X. (2004). Learning to decode cognitive states from brain images. *Machine Learning* **57,** 145–175.

Moran, P. (1948). Some theorems on time series. II. The Significance of the serial correlation coefficient. *Journal of the Royal Statistical Society, Series B* **35,** 255–260.

Moran, P. (1950). Notes on continuous stochastic phenomena. *Biometrika* **37,** 17–23.

Mueller, S. G., Weiner, M. W., Thal, L., Petersen, R. C., Jack, C., Jagust, W., Trojanowski, J. Q., Toga, A. W., and Beckett, L. (2005). The Alzheimers Disease Neuroimaging Initiative. *Neuroimaging Clinics of North America* **15,**.

Murata, N. (2001). Properties of the empirical characteristic function and its application to testing for independence. In *Proceedings of Third International Conference on Independent Component Analysis and Blind Signal Separation*, pages 19–24.

Nachev, P., Wydell, H., O'Neil, K., Husain, M., and Kennard, C. (2007). The role of the pre-supplementary motor area in the control of action. *NeuroImage* **36,** T155–T163.

Ogawa, S., Lee, T. M., Kay, A., and Tank, D. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciencies, Biophysics* **87,** 9868–9872.

Perlstein, W., Dixit, N., Carter, C., Noll, D., and Cohen, J. (2003). Prefrontal cortex dysfunction mediates deficits in working memory and prepotent responding in schizophrenia. *Biological Psychiatry* **53,**.

Pessoa, L., , and Padmala, S. (2007). Decoding near-threshold perception of fear from distributed single-trial brain activation. *Cerebral Cortex* **17,** 691–701.

Picard, N. and Strick, P. (1996). Motor areas of the medial wall: A review of their location and functional activation. *Cerebral Cortex* **6,** 342–35.

Pugh, K., Shaywitz, B., Shaywitz, S., Shankweiler, D., Katz, L., Fletcher, J. M., Skudlarski, P., Fulbright, R., Constable, R., Bronen, R., Lacadie, C., and Gore, J. (1996). Predicting reading performance from neuroimaging profiles: The cerebral basis of phonological effects in printed word identification. *NeuroImage* **3,** S455.

Purdon, P., Solo, V., Weisskoff, R., and Brown, E. (2001). Locally regularized spatiotemporal modeling and model comparison for functional MRI. *NeuroImage* **14,** 912–923.

Rencher, A. (2002). *Methods of Multivariate Analysis, 2nd ed.* John Wiley & Sons, Inc., New York., 2nd edition.

Searle, W. (1983). Cubic Clustering Criterion, SAS Technical Report A-108. Technical report, SAS Technical Report A-108.

Shattuck, D., Mirza, M., Adisetiyo, V., Hojatkashani, C., Salamon, G., Narr, K., Poldrack, R., Bilder, R., and Toga, A. W. (2008). Construction of a 3D probabilistic atlas of human cortical structures. *NeuroImage* **39,** 1064–1080.

Sherman, S. M. and Guillery, R. W. (2002). The role of the thalamus in the flow of information to the cortex. *Philosophical Transactions of the Royal Society London: B* **357,** 1695–1708.

Shimizu, S. and Kano, Y. (2001). Examination of independence in independent component analysis. In *Proceedingsof 2001 International Meeting of the Psychometric Society*.

Simon, O., Kherif, F., Flandin, G., Poline, J.-B., Rivière, D., Mangin, J.-F., Le Bihan, D., and Dehaene, S. (2004). Automatized clustering and functional geometry of human parietofrontal networks for language, space, and number. *NeuroImage* **23,** 1192–1202.

Stogbauer, H., Andrzejak, R., Kraskov, A., and Grassberger, P. (2004). Reliability of ICA estimates with mutual information. In *Lecture Notes in Computer Science: Independent Component Analysis and Blind Signal Separation*, volume 3195, pages 209–216.

Stonnington, C., Chu, C., Klöppel, S., Jack, C. J., Ashburner, J., Frackowiak, R., and Initiative., A. D. N. (2010). Predicting clinical scores from magnetic resonance scans in Alzheimer's disease. *NeuroImage* **51,** 1405–13.

Talairach, J. and Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain.* Thieme Medical Publishers, Inc., New York.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., and Joliot, M. (2002). Automated anatomical labelling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single subject brain. *NeuroImage* **15,** 273–289.

Vapnik, V. (1995). *The Nature of Statistical Learning Theory.* Springer Verlag, New York.

Volkow, N. D., Wang, G.-J., Fowler, J. S., Wong, C., Ding, Y.-S., Hitzemann, R., Swanson, J. M., and Kalivas, P. (2005). Activation of orbital and medial prefrontal cortex by methylphenidate in cocaine-addicted subjects but not in controls: relevance to addiction. *The Journal of Neuroscience* **25,** 3932–3939.

Waller, L. A., Carlin, B. P., Xia, H., and Gelfand, A. E. (1997). Hierarchical spatio-temporal mapping of disease rates. *Journal of Americal Statistal Association* **92,** 607–617.

Wold, H. (1966). *Multivariate Analysis.*, chapter Estimation of principal components and related models by iterative least squares., pages 391–420. New York: Academic Press.

Woods, R., Grafton, S., Holmes, C., Cherry, S., and Mazziotta, J. (1998). Automated image registration: I. general methods and intrasubject, intramodality validation. *Journal of Computer Assisted Tomography* **22,** 139–152.

Woods, R., Grafton, S., Watson, J., N.L., S., and J.C., M. (1998). Automated image registration: II. Intersubject validation of linear and nonlinear models. *Journal of Computer Assisted Tomography* **22,** 153–165.

Woolrich, M. W., Jenkinson, M., Brady, J. M., and Smith, S. M. (2004). Fully Bayesian spatio-temporal modeling of fMRI data. *IEEE Transactions on Medical Imaging* **23,** 213–231.

Worsley, K. and Friston, K. (1995). Analysis of fMRI time-series revisited-again. *Neuroimage* **2,** 173–181.

Worsley, K., Liao, C. H., Aston, J., Petre, V., Duncan, G. H., Morales, F., and Evans, A. C. (2002). A general statistical analysis for fMRI data. *NeuroImage* **15,** 1–15.

Wu, E., Yu, P., and Li, W. (2009). A smoothed bootstrap test for independence based on mutual information source. *Computational Statistics & Data Analysis Archive* **53,** 2524–253.