Identifying Heterogeneous Disorders Through Normative Modeling of the Resting State BOLD Signal

by

Ariel Oldag

A thesis submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Master of Science

Auburn, Alabama
May 2, 2020

Keywords: Normative Modeling, Bayesian Inference, fMRI, Resting State, Statistical Learning

Approved by

Dr. Mark Carpenter, Chair, Professor, Department of Mathematics and Statistics
Dr. Gopikrishna Deshpande, Professor of Electrical and Computer Engineering
Dr. Bertram Zinner, Associate Professor Emeritus of Mathematics and Statistics

Abstract

One of the recent developments of interest in neuroscience is the detailed study into resting state data. This data allows us to examine the way our brains behave in an awake state with no tasks presented. When coupled with non-imaging data, resting state data can greatly assist in the classification of disorders and speed patients along the road to recovery. The method of classification examined in this thesis is normative modeling through Scalable Multi-Task Gaussian Process Regression (S-MTGPR). Two distinct and unrelated datasets are used - a large dataset (n=172) and a small dataset (n=27) - to prove the effectiveness of S-MTGPR is unrelated to sample size. In this thesis we examine how normative models are built and how they classify subjects based on neurological activity.

Acknowledgments

I would like to thank Dr. Deshpande, Dr. Carpenter, and Dr. Zinner for their time, helpfulness, and patience. They have pushed and challenged me. I sincerely believe I have become a better researcher and student throughout this process, thanks to them. This thesis could not have been completed without their assistance, and I am incredibly grateful to them for this opportunity.

My path to success has been possible because of the unconditional love and support shown to me from my friends and family. The Ramseys, the Oldags, the Ogles, the Harrisons, Michael Schur, and the Artalonas. I'm thankful for all of you and couldn't have made it this far without you.

A special thank you to my husband, Jordan. I love you.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 MRI & BOLD Signal

Magnetic Resonance Imaging (MRI) is a method of producing three dimensional anatomical images. It is non-invasive and does not use ionizing radiation. It creates incredibly clear images of non-bony parts and soft tissue, making it more ideal than computed tomography (CT) or x-rays when it comes to imaging the brain, spinal cord, nerves, muscles, ligaments, and tendons [23].

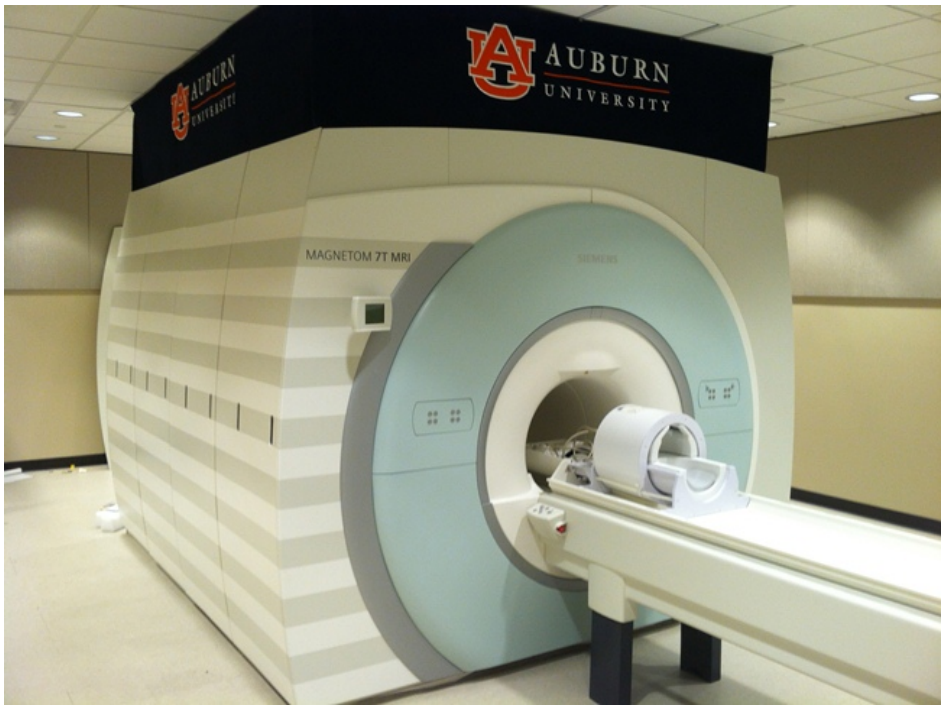

Figure 1.1: Photo of the 7T MRI Scanner used by Auburn University

The brain is composed of two different types of matter - grey and white. Grey matter contains cell bodies, dendrites, and axon terminals of neurons, meaning it ultimately contains all the synapses in the brain. White matter is composed of axons which connect the grey matter in the brain. MRI can differentiate between grey and white matter, making it a highly viable option for medical diagnosis or therapy. It's only downside is the cost and length of time required for a scan session.

Figure 1.2: Grey & White Matter. This image is reproduced from Neuroscience News & Research

MRIs work by using powerful magnets to create a magnetic field that forces protons in the body to align with the created field. A radiofrequency current is emitted and pushed through the patient, which stimulates the protons and forces them to spin out of equilibrium [23, 5]. This strains against the magnetic field and sensors within the MRI are able to detect energy released from the protons as they struggle to realign with the magnetic field. The length of time it takes for a proton to realign is dependent on its environment and chemical nature, allowing physicians to distinguish between types of tissues in the body [23].

There are several variables of interest in MRI scans, but the variable this thesis focuses on is the blood oxygen-level dependent (BOLD) signal. The BOLD signal is strongly correlated

with neural activity contributing to the local field potential (LFP) [15]. This demonstrates the concept that as activity in the brain increases, so does the consumption of oxygen. For the purposes of this thesis, the BOLD signal is used to identify neural activity.

There are several ways to manually structure brain architecture in order to examine where the extracted BOLD signal originates. This thesis will focus on the use of Regions of Interest (ROI) to determine connectivity. Specifically, we will examine how these regions interact with one another. This is known as ROI-to-ROI connectivity. The segmentation of ROIs is dependent on the atlas used. The analysis in this thesis uses the AAL atlas, which segments the brain into 120 ROIs [28].

One of the primary issues with MRIs is the necessity for patients to remain still during scans. Any movement during a scan can blur images, create artifacts, and distort or warp the extracted time series. This can especially be a problem in children or patients that are afflicted with an ailment that prevents them from keeping still for an extended period of time [5]. There is current research into creating machines designed for these patients.

A method of combating these abnormalities in scans is preprocessing. There are multiple steps involved in preprocessing, usually involving mathematical transformations of the data [2]. The end result is a collection of scans that are equally formatted and plotted. Preprocessing is necessary in statistical analysis because it standardizes the data and provides meaning for coordinates and groupings.

1.2 Geneva-Combat & ABIDE

There are two datasets used in this thesis. The first is the Geneva-Combat data. All patients within this dataset have been diagnosed with some form of mild traumatic brain injury (mTBI) caused by combat related injuries. However, four of the patients have also been diagnosed with post traumatic stress disorder (PTSD) in addition to their mTBI. This results in a control group of n=23 and treatment of n=4. The non-imaging response focuses on the King-Devick (KD) score recorded for each patient.

KD scores are obtained through a 2 minute visual test in which subjects are required to state numbers as they observe them in mildly complex patterns that are written on cards or projected on an interactive screen [18]. The times required to complete each card are recorded in seconds using a stopwatch. The sum of the three test card time scores constitutes the summary score for the entire test, the KD time score. While these scores can be standardized, they were left as raw scores for this project.

The Hippocampus (HIP) and the Amygdala (AMYG) are the two ROIs of most importance in this analysis. HIP is the ROI most commonly associated with learning and memory, and AMYG is commonly associated with fear and aggression. When either of these is damaged through a traumatic event, their interaction grows [9]. This is especially true if the traumatic event was emotional [17, 9]. The Prefrontal Cortex is commonly associated with the processing and retaining of information. When HIP and AMYG are highly active, the Prefrontal Cortex is likely to increase in activity as well [17, 9, **?**]. These three regions are the areas of focus when examining PTSD.

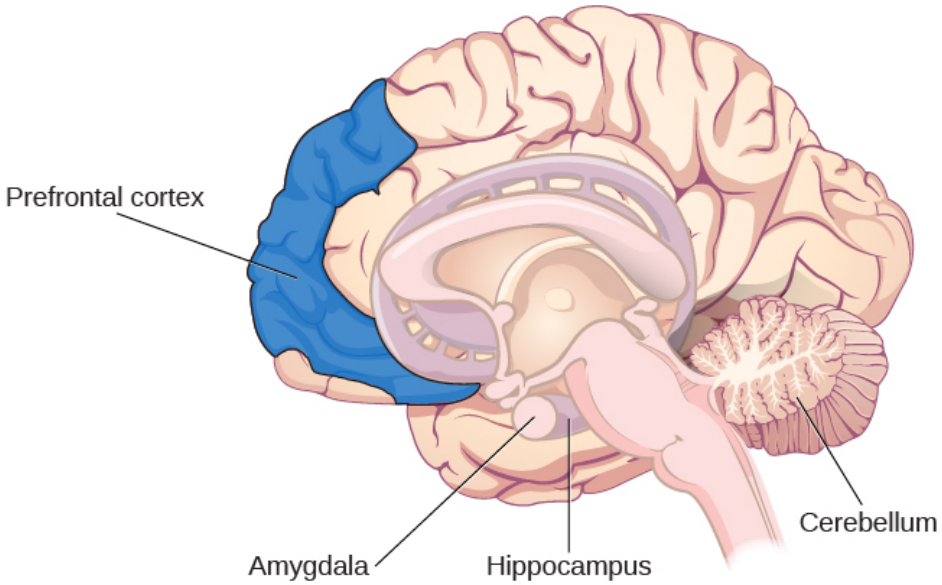

Figure 1.3: HIP, AMYG, and Prefrontal Cortex. This image was reproduced from from the Lumen Learning Group

All patients were scanned with a 7T at the MRI Facility in Auburn University. They were instructed to relax, keep their eyes open, and keep their head as still as possible throughout the duration of the scan. The subjects all provided informed consent and the scanning procedure was performed in accordance with the guidelines and the approval of the Institutional Review Board at Auburn University.

There are two methodologies underlying subject scans - resting state and task based. In resting state scans, subjects are to remain awake and alert, but are given nothing specific to focus on, recall, or do for the duration of the scan. These scans examine the activity of the brain in its natural state [34]. In task based scans, subjects are given a list of tasks to complete during the scans (i.e. tapping fingers, recalling events, examining images, etc) and the scans examine the activity of the brain in response to these tasks. The scans used in the Geneva-Combat dataset are considered to be resting state scans.

With a sample size of n=27, the Geneva-Combat dataset brings a risk of its results being anecdotal. To verify the methodology used and the results obtained, a second dataset is introduced that has no relation to Geneva-Combat - the Autism Brain Imaging Database Exchange (ABIDE).

The ABIDE dataset contains 1112 subjects and is measured across 17 international sites. For this thesis, 1 imaging site is used. All data comes from the NYU Langone Medical Center and contains 172 subjects. Of these subjects, 89 individuals have been diagnosed with Autism Spectrum Disorder (ASD) and there are 83 control patients. The subjects range in age from 7 to 64 years. This dataset was publicly released in 2011 after being successfully anonymized. The ABIDE dataset is part of the 1000 Functional Connectomes Project [13], and further information can be found on the NITRC website. There is a large collection of non-imaging data for this dataset, but this thesis will only focus on the binary response of control (0) vs. ASD (1). Further, this thesis uses the pre-processed data provided by The Preprocessed Connectomes Project [12]. The results from this dataset will be compared to

the subject's Autism Diagnostic Observation Schedule (ADOS) total score to examine the link between severity of outliers and associated severity of the ADOS score.

The ADOS is a standardized diagnostic test that measures a subject's interaction with certain objects or scenarios. It accounts for the developmental level and age of the subject, and can be modified to observe a large range of participants [4]. The ADOS is composed of four modules. A subject is scored at each module from zero to three - zero indicating non-abnormal behavior and three indication abnormal behavior.

This dataset is focused on the Temporal Lobe, which is a collection of 10 ROIs: HIP, the Parahippocampus (PHIP), AMYG, Fusiform Gyrus (FUSI), Heschl Gyrus (HES), Superior Temporal Gyrus (T1), Superior Temporal Pole (T1P), Middle Temporal Gyrus (T2), Middle Temporal Pole (T2P), and Inferior Temporal Gyrus (T3). The Temporal Lobe is commonly associated with ASD because it is responsible for the bulk of social cognition, empathy, and receptive language. The deficits of which are the typical symptoms used to diagnose ASD in subjects [4, 13, 30, 32, 33, 26].

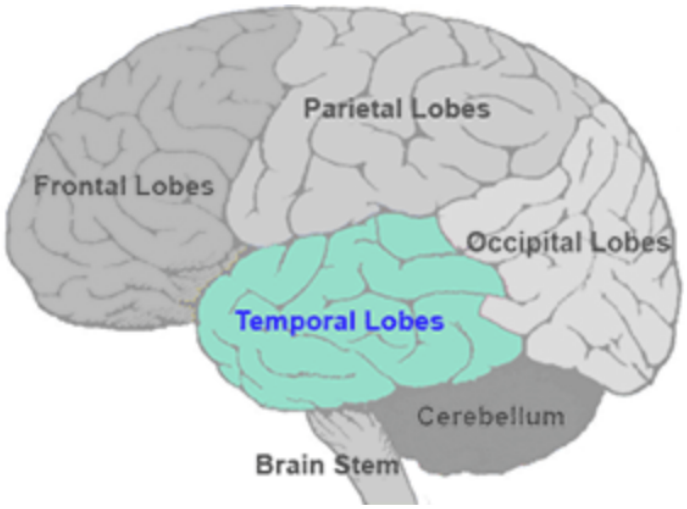

Figure 1.4: Frontal Lobe. This image was reproduced from the Queensland Health Organization

Subjects are scanned in resting state, meaning they are fully awake when the scan is performed and instructed to lie still and think of nothing. They are given no tasks and should not be attempting to remember events.

1.3 Preprocessing

After duplicating the data and saving it as a secondary dataset so as to avoid alteration of the raw data, the data is put through a pre-processing pipeline. For both functional and structural data, there are five collective steps taken in preprocessing. The Geneva-Combat dataset and the ABIDE dataset went through identical pipelines, as further described.

The first step is realignment and unwarping. This is also known as subject motion estimation and correction. Using SPM12, all scans are co-registered and re-sampled to a reference image using b-spline interpolation. B-spline interpolation is created through linear combinations of the original non-parametric function produced in the scan. The parameter B is in reference to the total number of parameters needed to accurately estimate the original function, and varies by subject [2, 29].

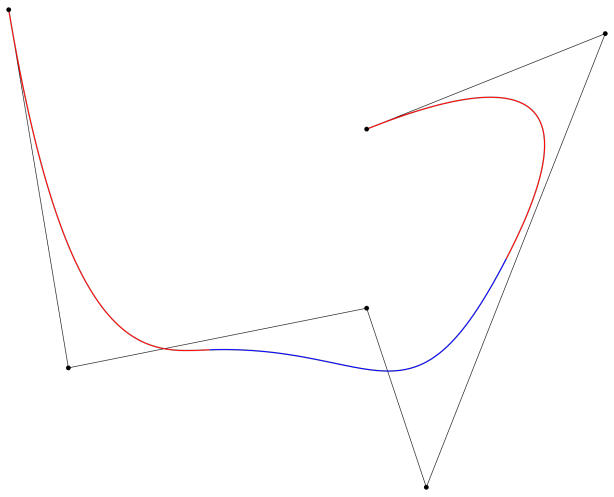

Figure 1.5: Graphical Representation of B-Spline. This image was reproduced from Wojciech Mula

The step of realignment and unwarping is performed to an attempt to correct potential distortion that is created through movement of the subject. The interaction between the distortion and the head movement is estimated through the derivatives of the deformation field [2]. Additionally, functional data is re-sampled along the phase-encoded direction in order to correct deformations caused by field distortions within the scanner. The realignment step also involves centering the data to (0,0,0) coordinates through translation. This is an important step because all connectivity measures will reference the same area of the brain within each patient.

The second step is slice-timing correction. fMRI data is sequential in nature, and creates a slight temporal misalignment between each slice. In this step, functional data is time-shifted and re-sampled using sinc-interpolation to match the time in the middle of each scan [2]. Sinc-interpolation is a method of obtaining a sequence of numbers from the fMRI scans, and converting them into a continuous-time bandlimited function [27]. This is a method of interpolating the signals between scans, thus allowing the scans to be viewed in a continuous fashion.

The third step is outlier identification. The majority of BOLD signals of interest occur within a range of 5 standard deviations of the mean BOLD signal. Anything outside of that range is typically denoted as an outlier due to head adjustment or signal noise [2]. A bounding box of 140x180x115 mm is created, and any timepoint registering outside of the bounding box is identified as an outlier and subsequently removed.

The fourth step is segmentation and normalization. Both functional and structural data are put into standard Montreal Neurological Institute (MNI) space. MNI space is the standard template used to map brains and was created through averaging 152 normal MRI scans and matching 9 parameters [28]. After the data is normalized, the scans are segmented into grey matter, white matter, and CSF tissue classes. This is done through a procedure in SPM12 that iteratively classifies tissue and estimates the posterior tissue probability map (TPM) from intensity values of the reference image [2].

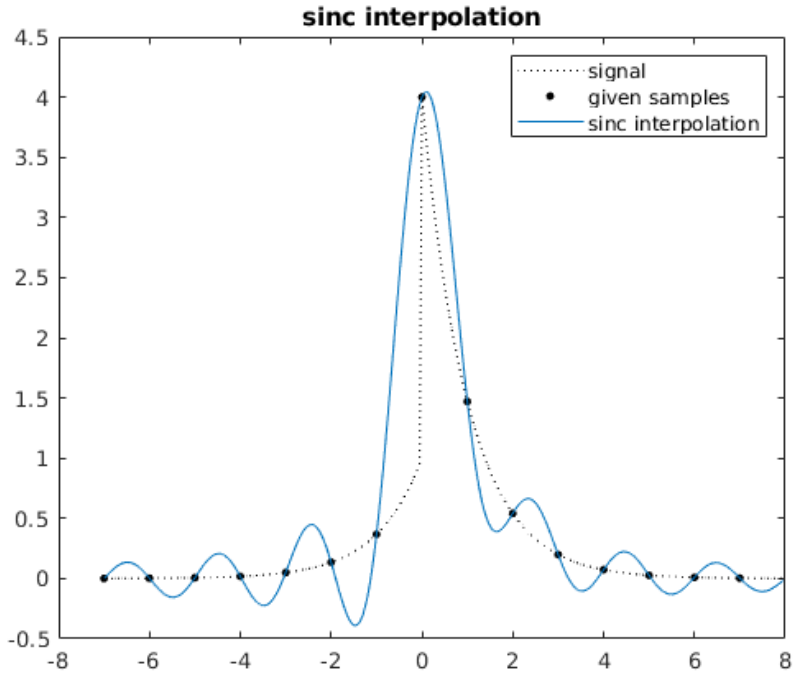

Figure 1.6: Graphical Represenation of Sinc-Interpolation. This image was reproduced from Julius O. Smith III

There is a fifth additional step taken for the functional data, in which the BOLD time-series is smoothed using spatial convolution with a gaussian kernel of 8mm. This is done to reduce variance across subjects, and increase the ratio of BOLD signal-to-noise [2].

All five of these steps are performed to both datasets, resulting in cleaned data that is ready for analysis. These are necessary because it assures that data meets the assumptions necessary for analysis - the time course comes from a single location, the data is uniformly spaced in time, and the data is spatially smooth.

1.4 Normative Modeling

After the datasets have been cleaned they are used to create a normative model. Normative modeling is ideal here, because fMRI data is inherently heterogeneous [33]. Normative models allow for the identification of deviations from a typical pattern, which in turns allows for a deeper understanding of neuroanatomical development at an individual level, rather

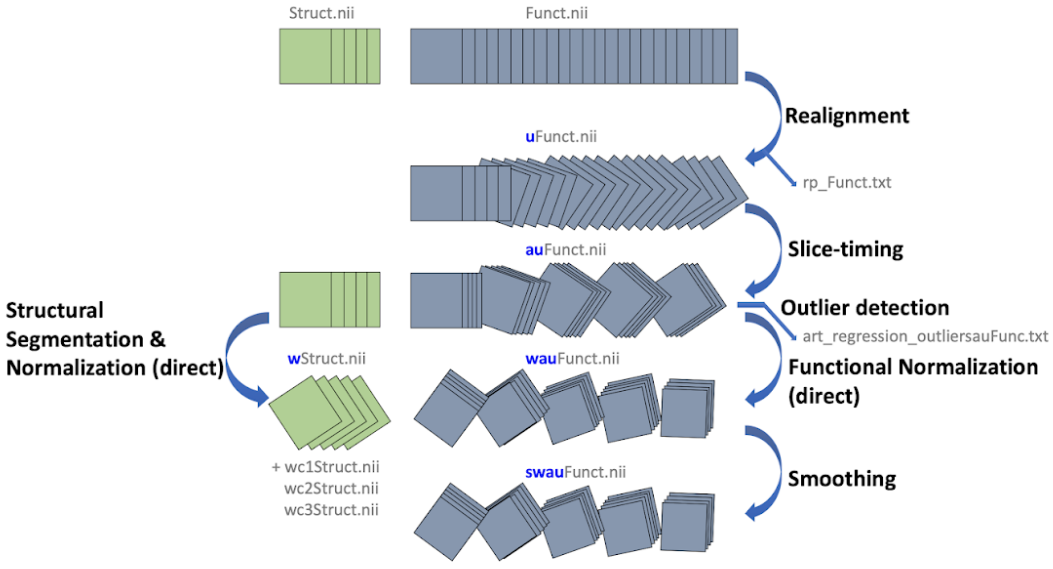

Figure 1.7: Preprocessing Steps. This image was reproduced from Alfonso Nieto-Castanon

than a group level. While normative modeling is similar in approach to the statistical learning technique of clustering, it is inherently different because it works to identify outliers - not clusters - which in turn identifies abnormalities in scans.

There are several ways to create a normative model, but the one used in this thesis is Scalable Multi-Task Gaussian Process Regression (S-MTGPR). S-MTGPR is used as opposed to other methods (ie Multi-task Kronecker Gaussian Process Regression) because of the reduced computational complexity. fMRI data typically has high dimensionality, and steps must be taken to reduce this dimensionality or the analysis will run into several problems.

1.5 Thesis Organization

A prevalent issue facing diagnosis of disorders and abnormalities in fMRI scans is the heterogeneity of samples. To advance the field of neuroscience, doctors and psychologists need to be able to scan a subject, compare the results to a large sample size of similar scans from other subjects, and quickly determine outliers in order to properly diagnose subjects.

Because of growing concern for computational complexity and length of time dedicated to diagnosis, the concept of identification of disorders within heterogeneous groups is the primary topic of this thesis. The first chapter is dedicated to generically introducing MRI and fMRI concepts, and briefly discussing the need for normative modeling. The datasets used in this thesis are also described in this chapter. In the second chapter, normative modeling is discussed in detail through the use of theorems and a discussion on how normative modeling is applied. In this chapter, both datasets undergo normative modeling and the outputs are examined in detail. Chapter three focuses on interpretation of the outputs and a discussion of possible concerns.

Chapter 2

Normative Modeling

2.1 Introduction to Normative Modeling

One of the primary reasons normative modeling is gaining traction in the field of neuroscience is because of its ability to produce reliable and reproducible results. Current methods, mainly clustering, are highly unreliable because they do not address the issue of heterogeneity and high dimensionality in the data [33]. In the medical field, conditions and diseases are typically diagnosed pathologically or clinically, meaning subjects are diagnosed based on their observed symptoms. However, many conditions have overlapping symptoms. When differentiating between two similar conditions, relying on symptoms to determine diagnosis is unwise. Both ASD and PTSD are disorders that share symptoms with many other diagnoses, which commonly creates a highly heterogeneous sample group.

S-MTGPR does not require separable clusters, but rather examines the outliers of the created normal model to identify possible disorders within subjects. This addresses the issue of heterogeneity in fMRI scans.

Normative modeling employs Gaussian process regression (GPR) to classify neuroimaging data on the basis of clinical and behavioral covariates (i.e independent variables). GPR provides measures of uncertainty and measures of variance through its creation of predictive confidence. This allows GPR models to capture multiple behaviors of subjects through fairly routine parameterization. It also allows for Bayesian inference. GPR previously could only produce one output. If more than one output was needed, it would require the technique of mulit-kriging. Boyle *et al.* proposed the idea of single-task GPR (STGPR) in 2004 by proving that outputs could be expressed through a convolutional process between a smoothing kernel and a latent function [8]. The kernel can be selected by the researcher, and the latent

function - in the case of neuroscience - can be modeled as the covariance function using the independent variables and the BOLD signal as inputs.

In 2008, Bonilla *et al.* introduced the idea of extending STGPR to multi-task GPR (MTGPR) through coupling together sets of latent functions with a shared Gaussian process prior distribution to produce a correlation between outputs. They proposed a method of taking the cross-covariance matrix into the Kronecker product of the sample and task covariance matrices, thus making it possible to model across-sample and across-task variance [7].

This was a large step forward for the field of neuroscience analysis, but the computational complexity of MTGPR still posed as a problem. When using N samples and T tasks, MTGPR has a time and space complexity of $O(N^3T^3)$ and $O(N^2T^2)$ respectively. While this may seem manageable under most datasets, it quickly can overwhelm computational power when dealing with fMRI data - typically N is larger than 10^4 while T is larger than 10^5 [16, 21]. Several efforts have been made to reduce computation time throughout this GPR. Their experiments range from the use of principal component analysis (PCA) to reduce the sample and task covariance matrix, to variational inducing kernals (VIK) which reduces time complexity to $O(NTM^2)$ through the use of inducing functions rather than inducing outputs. However, all of these methods examined made very few improvements on the order of T.

Marquand *et al.* proposed a solution to this by using a combination of low-rank approximations of the task covariance matrix through algebraic properties of the Kronecker product. This is known as the scalable MTGPR (S-MTGPR). Using a publicly available fMRI dataset, they proved that S-MTGPR has a much lower computational time than STGPR and also has higher sensitivity than STGPR [16].

2.2 Methodology

2.2.1 Notation - RRC

Boldface capital **R** is the BOLD timeseries within each ROI. Lowercase r is the matrix of correlation coefficients. Boldface captial **Z** is the ROI-to-ROI Connectivity matrix (RRC) of Fisher-transformed correlation coefficients.

2.2.2 ROI-to-ROI Connectivity

The measurement of connectivity obtained from the subjects is on an ROI-to-ROI level, meaning it uses all voxels located in a region to summarize the activity of said region. This is used because we are more interested in looking at the entire network of connections within the brain. The primary measurement used to assess connectivity is the Fisher-Transformed Bivariate Correlation Coefficient between pairs of ROI BOLD timeseries signals. Every combination of pairs is measured, and thus produces the RRC. Each element in the matrix is calculated:

Theorem 2.1

$$r(i,j) = \frac{\int R_i(t)R_j(t)dt}{(\int R_i^2(t)dt \int R_j^2(t)dt)^{\frac{1}{2}}}$$

2.2.3 Notation - S-MTGPR

Boldface capital letters, **A** are used to denote matrices. Bold face italic capital letters, A, are used to denote scalar numbers. The vertical vector that results from collapsing columns of a matrix $\mathbf{A} \in R^{N\times T}$ with $vec(\mathbf{A}) \in R^{NT}$. The Kronecker and element-wise matrix products are denoted as $\otimes$ and $\odot$, respectively.

2.2.4 Scalable Multi-Task Gaussian Process Regression

Let $\mathbf{X} \in R^{N \times F}$ be the input matrix with N samples and F covariates. Let $\mathbf{Y} \in R^{N \times T}$ represent the response variable matrix with N samples and T voxels. The multi-task Kronecker Gaussian process model (MT-Kronprod) is defined as:

Theorem 2.2

$$p(\boldsymbol{Y} \mid \boldsymbol{0}, \boldsymbol{D} \otimes \boldsymbol{R} + \sigma^2 \boldsymbol{I})$$

where $\mathbf{D} \in R^{T \times T}$ and $\mathbf{R} \in R^{N \times N}$ are respectively the voxel and sample covariance matrices. Because of the high computational complexity of matrix diagonalization operations, a low-rank approximation of $\mathbf{D}$ is used.

Let $\phi : \mathbf{Y} \rightarrow \mathbf{Z}$ be an orthogonal linear transformation that transforms $\mathbf{Y}$ to a reduced latent space $\mathbf{Z} \in R^{N \times P}$, where $P < T$, and $\mathbf{Z} = \phi(\mathbf{Y}) = \mathbf{YB}$. Here, columns of $\mathbf{B} \in R^{T \times P}$ represent a set of P orthogonal basis functions. Assuming a zero-mean matrix normal distribution for $\mathbf{Z}$, by factorizing its rows and columns:

Theorem 2.3

$$p(\boldsymbol{Z} \mid \boldsymbol{C,R}) = MN(\boldsymbol{0}, \boldsymbol{C} \otimes \boldsymbol{R}) = \frac{exp(-\frac{1}{2} Tr[\boldsymbol{C}^{-1} \boldsymbol{B}^T \boldsymbol{Y}^T \boldsymbol{R}^{-1} \boldsymbol{YB}])}{\sqrt{(2\pi)^{NP} \mid \boldsymbol{C} \mid^P \mid \boldsymbol{R} \mid^N}}$$

where $\mathbf{C} \in r^{P \times P}$ and $\mathbf{R} \in r^{N \times N}$ are columns and row covariance matrices of $\mathbf{Z}$. Using the trace invariance property under cyclic permutations, the noise-free multivariate normal distribution of $\mathbf{Y}$ can be approximated:

Theorem 2.4

$$p(\boldsymbol{Y} \mid \boldsymbol{D,R}) \approx p(\boldsymbol{Y} \mid \boldsymbol{C,B,R}) = \frac{exp(-\frac{1}{2} Tr[\boldsymbol{BC}^{-1} \boldsymbol{B}^T \boldsymbol{Y}^T \boldsymbol{R}^{-1} \boldsymbol{Y}])}{\sqrt{(2\pi)^{NT} \mid \boldsymbol{BCB}^T \mid^T \mid \boldsymbol{R} \mid^N}}$$

where $\mathbf{D}$ is approximated by $\mathbf{BCB}^T$. The S-MTGPR is then derived by marginalizing over noisy samples:

Theorem 2.5

$$p(\boldsymbol{Y} \mid \boldsymbol{D,R,}\sigma^2) \approx p(\boldsymbol{Y} \mid \boldsymbol{C,B,R,}\sigma^2) = N(\boldsymbol{Y} \mid \boldsymbol{0,}\ \boldsymbol{BCB}^T \otimes \boldsymbol{R} + \sigma^2 \boldsymbol{I})$$

Further, to obtain the mean and variance of the predictive distribution of the population, standard GPR framework is used:

Theorem 2.6

$$vec(\boldsymbol{M}^*) = (\boldsymbol{R}^* \boldsymbol{U_R} \boldsymbol{Y} \boldsymbol{U_C^T} \boldsymbol{CB^T})$$

and

Theorem 2.7

$$\boldsymbol{V}^* = (\boldsymbol{D} \otimes \boldsymbol{R}^{**}) - (\boldsymbol{BCU}_C \otimes \boldsymbol{R}^* \boldsymbol{U}_R)\boldsymbol{K}^{-1}(\boldsymbol{U}_C^T \boldsymbol{CB}^T \otimes \boldsymbol{U_R^T} \boldsymbol{R}^{*T})$$

in which $\mathbf{C} = \mathbf{U_C S_C U_C^T}$ and $\mathbf{R} = \mathbf{U_R S_R U_R^T}$ are the eigenvalue decompositions of the covariance matrices. Because $\mathbf{B}$ is assumed to be orthogonal, it is also assumed that $\mathbf{B}$ is invertible and positive definite.

To produce predictive variance, it is necessary to calculate the marginal log likelihood. This is seen as:

Theorem 2.8

$$\mathcal{L} = -\frac{N \times T}{2} ln(2\pi) - \frac{1}{2} ln \mid \boldsymbol{K} \mid -\frac{1}{2} vec(\boldsymbol{U}_R^T \boldsymbol{YBU_C})^T vec(\boldsymbol{Y})$$

There are three sets of parameters of interest in S-MTGPR. These are $\Theta_{\mathbf{C}}, \Theta_{\mathbf{R}}$, and Θ_{σ^2}, which are respectively, the reduced voxel covariance matrix, the input covariance matrix, and the noise variance. The three main parameters, $\Theta_{\mathbf{C}}, \Theta_{\mathbf{R}}$, and Θ_{σ^2} are calculated through the optimization of $\mathcal{L}$. When calculated using the reduced matrices of the inputs, these three parameters will be single vectors.

2.3 Normative Modeling & PTSD

There are a total of 27 subjects in the Geneva-Combat dataset. All of which have been diagnosed with an mTBI, but 4 have additionally been diagnosed with PTSD. All of the mTBIs of the subjects originated from combat, and were somehow related to their military training. At the time of the diagnosis of mTBI, a KD sideline score was recorded for each subject. The KD score of each subject is used as the primary covariate in the SMTGPR model. These are listed below in accordance to each subject.

Subject	KD Score	Subject	KD Score
1	43.00	15	32.72
2	42.75	16	52.44
3	56.66	17	44.85
4	41.97	18	51.40
5	36.56	19	40.63
6	58.20	20	41.93
7	64.86	21	41.93
8	52.37	22	41.93
9	52.06	23	41.93
10	44.69	24	41.93
11	44.49	25	41.93
12	40.56	26	38.91
13	37.91	27	38.72
14	51.91	-	-

Table 2.1: Associated KD Scores

After pre-preprocessing the data, the subject's scans are concatenated into a single nifti file using SPM12 3D to 4D conversion. This single file is then put into Marquand's

normalization code, which builds the S-MTGPR model. The associated KD scores are additionally entered into the model as covariates.

Typically, a standard anatomical volume, with an isotropic voxel resolution of 1mm contains almost 17 million voxels, which are arranged in a 3D matrix of 256 x 256 x 256 voxels [22]. After the S-MTGPR model is built, a volume represents a summary of each subject used in the input. An analysis of the volumes is used to identify outliers.

After experimentation with the model outputs, 4 k-folds had the lowest root mean-squared error (RMSE) at 0.05954 and was subsequently selected. RMSE was used in this calculation because it represents the absolute fit of the data and is simple to interpret.

The normative model is trained on the healthy subjects (n=23). These subjects create the average neuronal activity and estimated standard deviations that is used in comparison to the subjects with PTSD. The healthy subjects can be identified through their low ranking KD scores. The subjects that have PTSD (n=3) can easily be identified by their high KD scores. It is important to note that this is not always true but is rather a coincidence in our data. They are patients 3, 6, 7, and 16. The areas of focus in the brain include the AMGY, the HIPP, and the frontal lobe. These three areas have been proven to be heavily affected by PTSD in previous research [9]. If any deviating behavior is expected to be found in these subjects, it will likely be in these regions.

The S-MTGPR model creates several outputs, including a generalized mask of the normal functional neuronal activity for the sample. This activity is shown through the Fisher-transformed bivariate correlation coefficient, and represents the average correlation within the brain that can be found throughout the healthy subjects. This measure is pictured through taking a single voxel, recording its correlation with every other voxel, then taking the average. If this voxel has high neuronal activity, it is likely to have a high average. In a resting state, there are typically multiple areas of activity. The areas highlighted in blue show regions of the brain that are slightly more active than others. The majority of the brain in this layout hovers around a correlation value of 0, with a few significant spikes that can

be seen in blue. These spikes are not located in either the AMGY, HIPP or Frontal Lobe, but rather are spread throughout the default activity network that is commonly referred to as the standard for activity networks in resting state scans [10]. The numerical values that make up this scan can be found in the appendix.

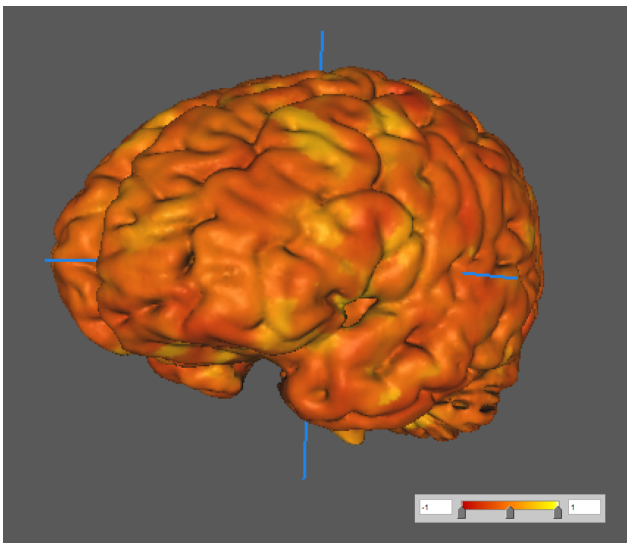

Figure 2.1: Estimated connectivity of a Healthy Subject (Geneva-Combat)

After the S-MTGPR model is created, one of the outputs is the Predicted Yhat Layout. This parses through each input and estimates which subjects contain outliers and identifies where the outliers are located. Each pane of the image shows a slice of the brain in axial view, cutting through the middle of the brain. Each slice corresponds to a different subject, showing what can be considered a summary image of each subject. Most subjects look relatively similar, with three exceptions. These exceptions appear to have more activity in their AMGY, HIPP, and Frontal Lobe. When compared to the mask of the normal functional neuronal activity of the sample, it is clear that these three subjects vary significantly from the created model, and give cause for further investigation.

Average correlation values are calculated for each volume to be more precise in analysis, in addition to their standard deviations. As seen in the table below, the averages of subjects

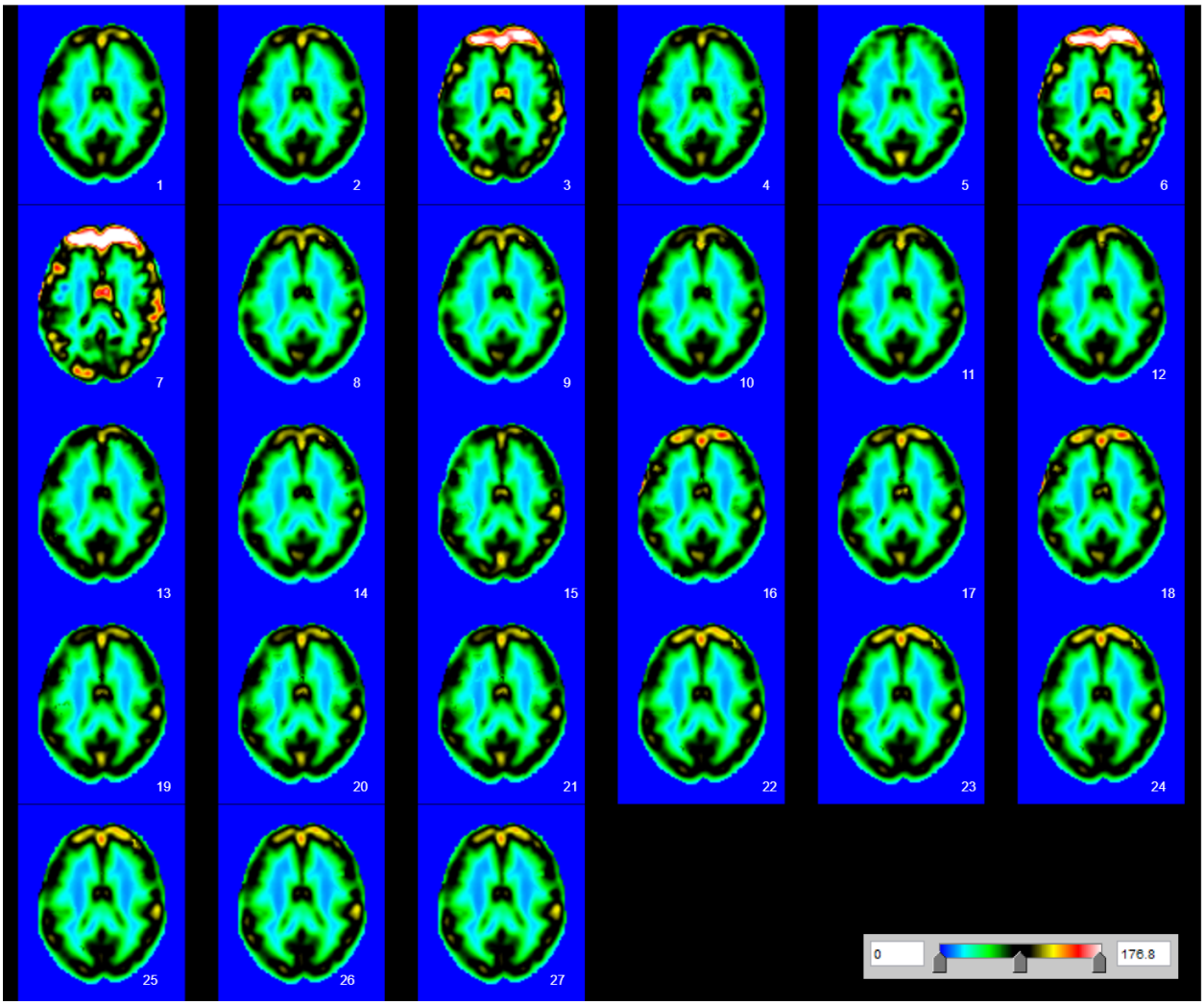

Figure 2.2: Predicted Yhat Layout (Geneva-Combat)

3, 6, and 7 are significantly higher than the others and have slightly higher standard deviations as well. Subject 16 struggles to stand out from the controls.

The hyperparameters, $\Theta_{\mathbf{C}}, \Theta_{\mathbf{R}}$, and Θ_{σ^2}, are shown below. As a reminder, $\Theta_{\mathbf{C}}$ is the reduced voxel covariance matrix, $\Theta_{\mathbf{R}}$ is the input covariance matrix, and Θ_{σ^2} is the noise variance. Each hyperparameter represents the eigenvalue decomposition of their respective matrix, and can be interpreted as the magnitude and direction of the data.

Volume	Mean	SD	Volume	Mean	SD
1	0.1365	0.2547	15	0.1384	0.2603
2	0.1362	0.2539	16	0.1341	0.2525
3	0.1550	0.2954	17	0.1365	0.2557
4	0.1353	0.2552	18	0.1348	0.2533
5	0.1263	0.2382	19	0.1362	0.2549
6	0.1573	0.3009	20	0.1358	0.2541
7	0.1667	0.3257	21	0.1359	0.2541
8	0.1296	0.2442	22	0.1368	0.2573
9	0.1297	0.2443	23	0.1368	0.2573
10	0.1330	0.2483	24	0.1368	0.2573
11	0.1330	0.2484	25	0.1368	0.2573
12	0.1343	0.2505	26	0.1368	0.2577
13	0.1354	0.2528	27	0.1368	0.2578
14	0.1298	0.2443	-	-	-

Table 2.2: Activity Volume Means & Standard Deviations

Hyperparameter	Estimation
$\Theta_{\mathbf{C}}$	-5.38E-04
$\Theta_{\mathbf{R}}$	0.1718
Θ_{σ^2}	-1.9152

Table 2.3: Hyperparameter Estimates (Geneva-Combat)

The application of the normative model in identifying heterogeneous disorders can be seen through examining the estimated mean connectivity value of the sample in comparison to individual values. The S-MTGPR model calculated a mean of 0.138 and an estimated standard deviation of 0.124. Under the assumption of a normal distribution, the empirical rule states that 68% of data should fall within one standard deviation, and 95% of data should fall within two standard deviations. Under this assumption, approximately 82% of data should fall within 1.5 standard deviations of the calculated mean. Using the empirical rule as a basis, outliers will be flagged as such if they fall outside 1.5 standard deviations in either direction. Healthy subjects should fall within the range (0.125, 0.152). The connectivity average is plotted against the calculated change in centroid location, as this location

should remain fairly constant among subjects and focuses attention on the difference in correlation values.

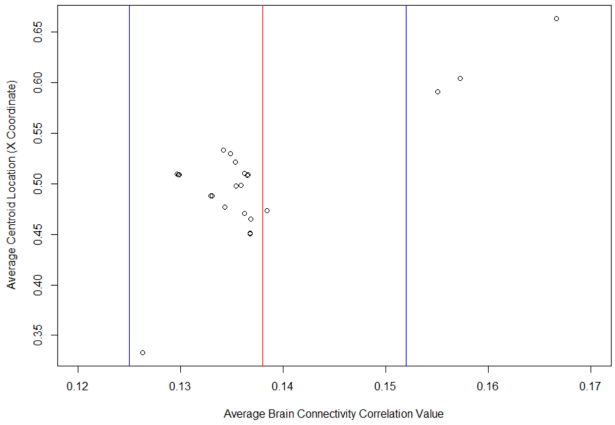

Figure 2.3: Associated Scatterplot of Volume Activity by Centroid

There are three subjects that register beyond the 1.5 standard deviation mark of 0.152. When comparing this plot to the numerical values, it is clear that the three values marked as outliers are subjects 3, 6, and 7. These subjects have been diagnosed with PTSD, and additionally have high KD scores. However, subject 16 - who has a diagnosis of PTSD - is not flagged by this marker. This results in a sensitivity rating of 75% and a specificity rating of 95.7%. Both the sensitivity rating and the specificity rating are fairly high. This methodology shows fairly positive results and gives hope that this methodology is reliable. However, this dataset is very small and unbalanced. As such, the ABIDE dataset is examined.

2.4 Normative Modeling & ASD

The dataset used in this portion is a subset of the ABIDE dataset. It only consists of scans originating from NYU. This set contains 172 subjects. Of these 172, 89 subjects have been diagnosed with ASD and 83 are considered control subjects. The process to create the SMTGPR model follows the same methodology as previously discussed. All subjects were preprocessed using the same pipeline, then concatenated into one nifti file. This file was then

used as an input in Marquand's normalization code, which performs the transformations and calculations previous discussed. The new hyperparameters have been estimated as follows:

Hyperparameter	Estimation
$\Theta_{\mathbf{C}}$	-8.29E-04
$\Theta_{\mathbf{R}}$	-0.8003
Θ_{σ^2}	-9.1990

Table 2.4: Hyperparameter Estimates (ABIDE)

The estimated correlation throughout the brain is calculated through the S-MTGPR model, and projected on an anatomical MNI T1 image.

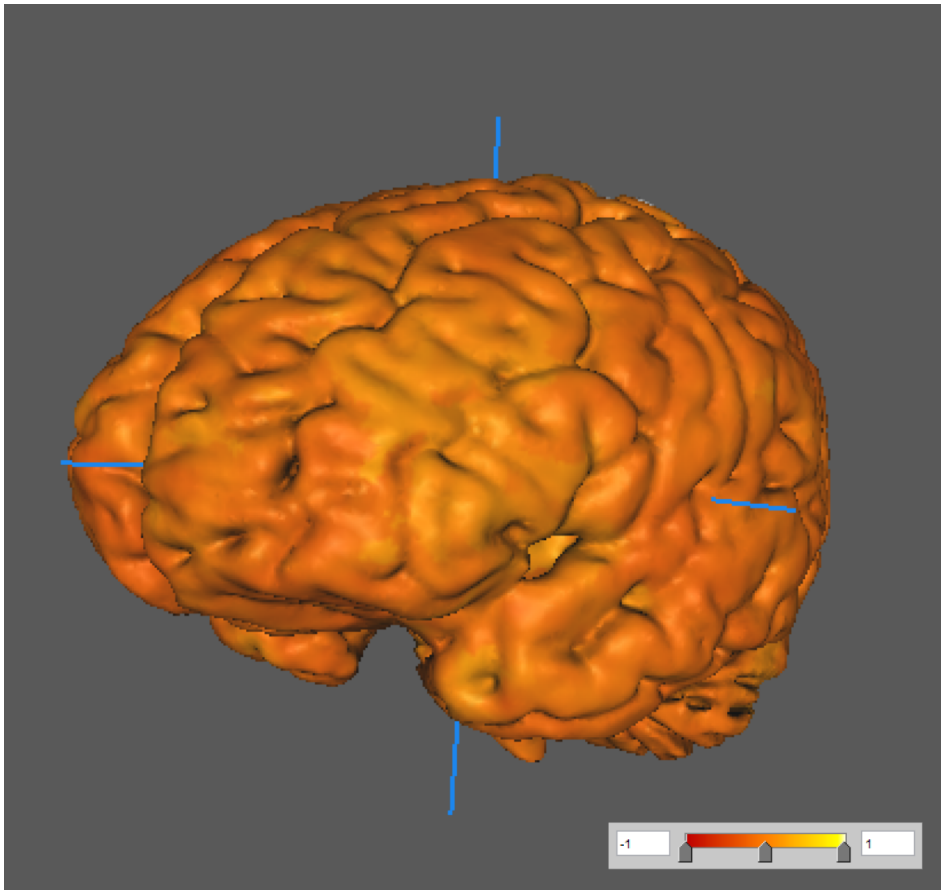

Figure 2.4: Estimated connectivity of a Healthy Subject (ABIDE)

Possible subject outliers are shown in their summary scans, in which ROIs that are more active than designated in the S-MTGPR model are shown in red. These areas are primarily

focused in the region that is considered to be the temporal lobe. This is in accordance with the theory that subjects diagnosed with ASD have a deficit in processing within this lobe [4, 13, 30, 32, 33, 26].

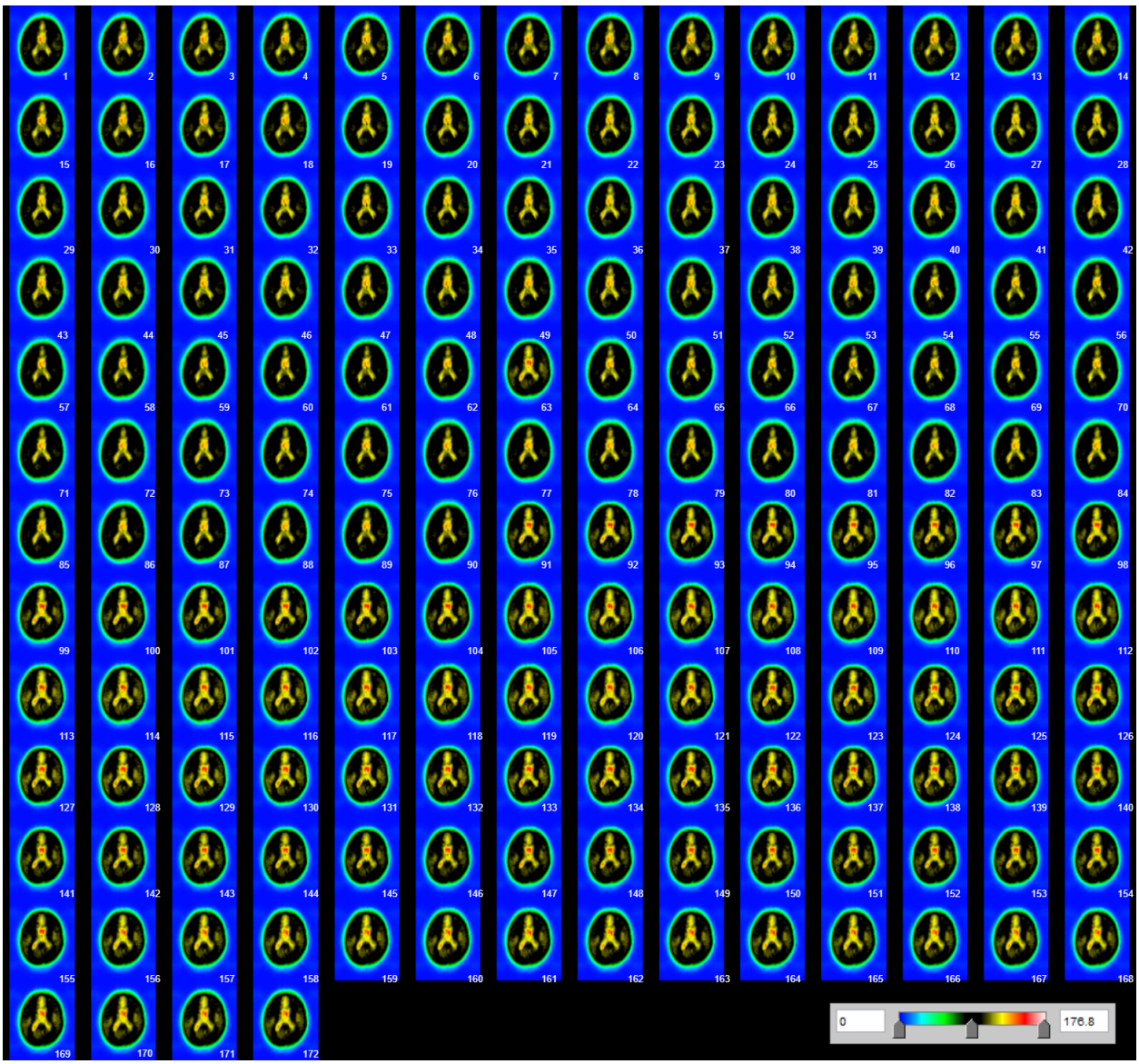

Figure 2.5: Predicted Yhat Layout (ABIDE)

After estimation, 10 k-folds resulted in the lowest RMSE for the model at 12.1292. The estimated mean connectivity value for the created S-MTGPR model is 22.9549 for this set. The estimated standard deviation for this model is 0.1162. If using 1.5 standard deviations as

the acceptable cut off limit, healthy subjects should fall within the range (22.8387, 23.0711).

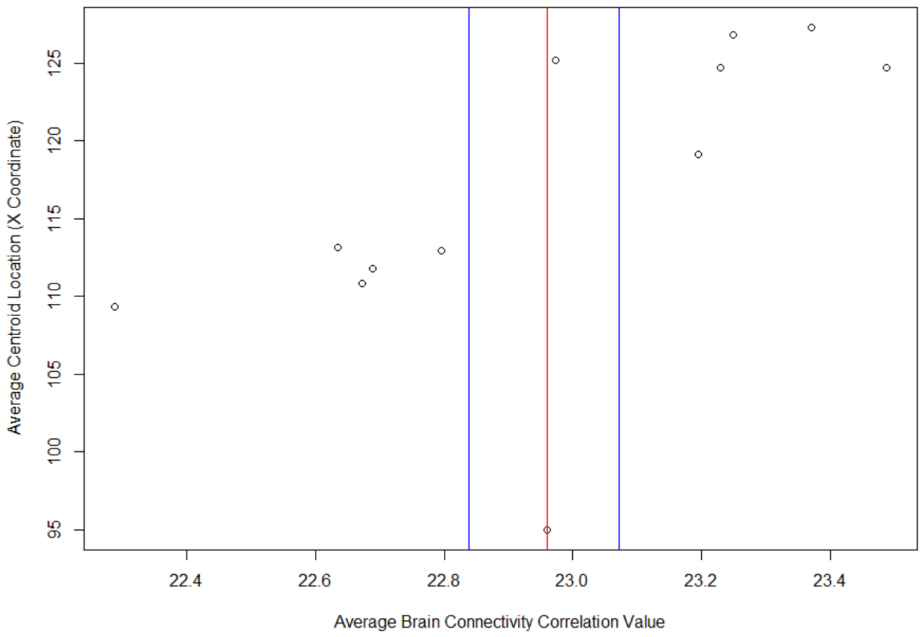

Figure 2.6: Associated Scatterplot of Volume Activity by Centroid

The full table of reported volume activity means per subject can be found in Appendix B. 138 subjects were flagged in this group as deviating significantly from the normative model and showing potential for ASD. Of these 138, 72 subjects were correctly diagnosed with ASD and 17 subjects were overlooked. The ABIDE set has a sensitivity rate of 87% and a specificity rate of 78%.

Chapter 3

Conclusion

3.1 Sensitivity & Specificity

Sensitivity and specificity are used here instead of an overall model accuracy because it's important to look at both sides of the picture. Sensitivity gives us the true positive, while specificity gives us the true negative. Rather than looking at the blanket "accuracy" of the model, we must examine four different aspects: the percentage of subjects correctly diagnosed with a disorder, the percentage of subjects incorrectly diagnosed with a disorder, the percentage of subjects correctly marked as control, and the percentage of subjects incorrectly marked as control. This can all be summarized through the sensitivity and specificity rates, which is why these are employed in analysis of the S-MTGPR model. These values are calculated across the entire dataset.

-	Sensitivity	Specificity
PTSD	75%	95.7%
ABIDE	87%	78%

Table 3.1: Sensitivity & Specificity Ratings

These ratings are fairly high and show that S-MTGPR is a useful tool in aiding the diagnosis of heterogeneous disorders. While this has already been proven in reference to cortical thickness in ASD detection [33], this thesis has proved that S-MTGPR can be used in novelty detection of heterogeneous disorders when the BOLD signal is the primary measurement.

In each dataset, the non-imaging data correlated directly to the results. Individuals that deviate further from the mean in their non-imaging data (KD score and ADOS total) were

outliers in the S-MTGPR model, and ranged in the severity of their deviations in accordance with their non-imaging data. Using the ADOS total as the covariate in the S-MTGPR model was considered, but there was a large lack of data available - especially in the non-ASD group.

3.2 Interpretation

Normative modeling is not meant to be the "final say" in diagnosing subjects with disorders. It is merely meant to be a tool that can assist physicians in heterogeneous cases. The primary goal of this research was to determine if normative modeling methodology can distinguish highly heterogeneous disorders within a group. Two datasets were tested - ABIDE and Combat. Through the use of S-MTGPR, a normative model was created for both datasets. MSE was examined in the selection of K-folds, and the smallest MSE was selected and used to determine the number of folds. S-MTGPR also produced biased hyperparameters, and mapped the expected fischer bivariate correlation within the brain for an average patient within the sample group.

The average patient in the Combat dataset was relatively healthy and did not have a diagnosis of PTSD. Therefore, if a subject registered as an outlier, they could be inferred as having PTSD. This was proven true numerically through the examination of deviation scores and mapping of the individual patient scans in comparison to the normative model.

The subjects in the ABIDE dataset were roughly split 50/50 of controls and treatment. As such, if a subject registered as an outlier, they would be perceived as having ASD. This was also proven true numerically through the examination of deviation scores and mapping of the individual patient scans in comparison to the normative model.

If a sufficiently large data bank could be created and maintained, it is likely that disorders and diseases - even highly heterogeneous - could be identified through the use of normative modeling. It is relatively inexpensive and quick to calculate, and does not require treatment groups or subject correlation, making it highly ideal for universal identification.

Normative modeling provides fast and inexpensive results, but more importantly, the results are easy to understand and interpret. Though physicians and psychologists study statistics, it is important to remember they are not statisticians or data scientists. As such, using methodology that is straight forward and builds off basic principles of statistics should be used for interdisciplinary work.

3.3 Issues & Concerns

As cross-validation is used in model fitting, the dangers of overfitting needs to be acknowledged. While the neuroscience community as a whole largely disregards this issue, it is of prime concern in the statistical community. Additionally, the combat dataset is rather small and the model is selected through the minimization of MSE through k-fold testing. This affects the hyperparameters chosen, and introduces bias into the model in addition to possible overfitting.

It is very likely that the combat dataset is over-fitted, with a training size of 85% and a testing size of 15%. However, because similar results were obtained with a larger dataset using a 50% split, it should be considered that over-fitting is not a primary concern with this methodology.

Bibliography

[1] Akshoomoff N, Corsello C, Schmidt H. The Role of the Autism Diagnostic Observation Schedule in the Assessment of Autism Spectrum Disorders in School and Community Settings. Calif School Psychol. 2006;11:7-19. doi:10.1007/BF03341111

[2] Alfonso Nieto-Castanon, "CONN Toolbox: Minimal Preprocessing Pipeline," 2020.

[3] CONN functional connectivity toolbox (Whitfield-Gabrieli, S., and Nieto-Castanon, 2012; http://www.nitrc.org/projects/conn)

[4] Baraka K, Melo F, Veloso M, "Simulating Behaviors of Children with Autism Spectrum Disorders Through Reversal of the Autism Diagnosis Process," EPIA 2017.

[5] Jonathan Blumenthal, Alex P Zijdenbos, Elizabeth Molloy, Jay N Giedd, "Motion Artifact in Magnetic Resonance Imaging: Implications for Automated Analysis," NeuroImage, 2016.

[6] Beckmann C, Kia S, Marquand A, "Scalable Multi-Task Gaussian Process Tensor Regression for Normative Modeling of Structured Variation in Neuroimaging Data," stat.ML, 2018.

[7] Bonilla, E.V., Chai, K.M., Williams, C., "Multi-task Gaussian process prediction,"In: Advances in neural information processing systems,2008.

[8] Boyle P, Frean M, "Dependent Gaussian Processes," Neural Information Processing Systems Conference, 2004.

[9] Bremner J, Brett E, "Trauma-Related Dissociative States and Long-Term Psychopathology in PTSD," Journal of Traumatic Stress, Vol 10, 1997.

[10] Buckner, R.L., DiNicola, L.M. The brain's default network: updated anatomy, physiology and evolving insights. Nat Rev Neurosci 20, 593–608 (2019). https://doi.org/10.1038/s41583-019-0212-7

[11] Cardoso S, "Structure of the Neuron," Brain and Mind, 2003.

[12] Cameron Craddock, Yassine Benhajali, Carlton Chu, Francois Chouinard, Alan Evans, András Jakab, Budhachandra Singh Khundrakpam, John David Lewis, Qingyang Li, Michael Milham, Chaogan Yan, Pierre Bellec (2013). The Neuro Bureau Preprocessing Initiative: open sharing of preprocessed neuroimaging data and derivatives. In Neuroinformatics 2013, Stockholm, Sweden.

[13] Di Martino A, Yan CG, Li Q, et al. The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. Mol Psychiatry. 2014;19(6):659-667. doi:10.1038/mp.2013.78

[14] Hart H, Chantiluke K, Cubillo A, Smith A, Simmons A, Brammer M, Marquand A, Rubia K, "Pattern Classification of Response Inhibition in ADHD: Toward the Development of Neurobiological Markers for ADHD," Wiley Periodicals, 2013.

[15] Hillman EM. Coupling mechanism and significance of the BOLD signal: a status report. Annu Rev Neurosci. 2014;37:161-181. doi:10.1146/annurev-neuro-071013-014111.

[16] Kia S, Marquand A, "Normative Modeling of Neuroimaging Data using Scalable Multi-Task Gaussian Processes," Stat.ML, 2018.

[17] Koenigs M, Huey E, Raymont V, Cheon B, Solomon J, Wassermann E, Grafman J, "Focal Brain Damage Protects Against PTSD in Combat Veterans," Nat Neurosci, 2018.

[18] Kristin A. Galtetta, et al.,"The King–Devick test and sports-related concussion: Study of a rapid visual screening tool in a collegiate cohort," Journal of the Neuorlogical Sciences, 2011.

[19] Authored by the Lumen Learning Group, "Introduction to Psychology: Parts of the Brain Involved with Memory," Lumen Learning, 2020.

[20] MANGO (Research Imaging Institute, UT Health Science Center at San Antonio, TX, USA; http://ric.uthscsa.edu/mango/)

[21] Marquand A, Rezek I, Buitelaar J, Beckmaan C, "Understanding Heterogeneity in Clinical Cohorts Using Normative Models Beyond Case-Control Studies," BioPsych, 2015.

[22] Michael Notter, "Nipype Beginner's Guide," 2018.

[23] National Institute of Health, National Institute of Biomedical Imaging and Bioengineering, "Magnetic Resonance Imaging (MRI) Fact Sheet," 2019.

[24] O'Connor S, Agius M, "A Systematic Review of Structural and Functional MRI Differences Between Psychotic and Nonpsychotic Depression," Psychiatria Danubina, 2015.

[25] Patriat R, Reynolds R, Birn R, "An Improved Model of Motion-Related Signal Changes in fMRI," Neuroimage, 2017.

[26] Authored by the Queensland Government, "Brain Map: Temporal Lobes," Queensland Health, 2017.

[27] Smith J III, "Physical Audio Signal Processing," 1993.

[28] N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, M. Jolio, "Automated Anatomical Labeling of Activations in SPM Using a Macroscopic Anatomical Parcellation of the MNI MRI Single-Subject Brain," NeuroImage15, 273–289 (2002).

[29] Wojciech Mula, "B-spline Curve," Wikimedia Commons, 2013.

[30] Wolfers T, Beckmann C, Hoogman M, Buitelaar J, Franke B, Marquand A, "Individual Differences vs. the Average Patient Mapping the Heterogeneity in ADHD Using Normative Models," Psychological Medicine 1-10, 2019.

[31] R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

[32] Wolfers T, Rooij D, Oosterlaan J, Heslenfeld D, Hartman C, Hoekstra P, Beckmann C, Franke B, Buitelaar J, Marquand A, "Quantifying Patterns of Brain Activity Distinguishing Unaffected Siblings from Participants with ADHD and Healthy Individuals," NeuroImage: Clinical, 2016.

[33] Zabihi M, et al., "Dissecting the Heterogeneous Cortical Anatomy of Autism Spectrum Disorder Using Normative Models," Bio Psychiatry Cogn Neurosci Neuroimaging, 2019.

[34] Zhang S, Li X, Lv J, Jiang X, Guo L, Liu T. Characterizing and differentiating task-based and resting state fMRI signals via two-stage sparse representations. Brain Imaging Behav. 2016;10(1):21-32. doi:10.1007/s11682-015-9359-7

Appendix

Appendix A

ABIDE Data Tables

Subject	Mean	Control/ASD	ADOS
1	22.7955	1	13
2	22.7955	1	11
3	22.7955	1	14
4	22.7955	1	10
5	22.7955	1	6
6	22.7955	1	7
7	22.7955	1	6
8	22.7955	1	8
9	22.7955	1	10
10	22.7955	1	15
11	22.7955	1	18
12	22.7955	1	7
13	22.7955	1	10
14	22.7955	1	6
15	22.7955	1	10
16	22.7955	1	11
17	22.7955	1	10
18	22.7955	1	13
19	22.6347	1	12
20	22.6347	1	17
21	22.6347	1	13
22	22.6347	1	7
23	22.6347	1	7
24	22.6347	1	7
25	22.6347	1	14
26	22.6347	1	8
27	22.6347	1	9
28	22.6347	1	12
29	22.6347	1	9
30	22.6347	1	10
31	22.6347	1	15
32	22.6347	1	13
33	22.6347	1	14
34	22.6347	1	17
35	22.6347	1	21
36	22.6347	1	22
37	22.6893	1	7
38	22.6893	1	14
39	22.6893	1	8
40	22.6893	1	15

Table A.1: Subject, Volume Activity Mean, Group, & ADOS (ABIDE Subjects 1-40)

Subject	Mean	Control/ASD	ADOS
41	22.6893	1	5
42	22.6893	1	15
43	22.6893	1	19
44	22.6893	1	11
45	22.6893	1	12
46	22.6893	1	8
47	22.6893	1	10
48	22.6893	1	19
49	22.6893	1	10
50	22.6893	1	7
51	22.6893	1	7
52	22.6893	1	12
53	22.6893	1	9
54	22.2886	1	7
55	22.2886	1	10
56	22.2886	1	5
57	22.2886	1	10
58	22.2886	1	8
59	22.2886	1	11
60	22.2886	1	13
61	22.2886	1	8
62	22.2886	1	8
63	23.2305	0	8
64	22.2886	1	16
65	22.2886	1	15
66	22.2886	1	17
67	22.2886	1	5
68	22.2886	1	10
69	22.2886	1	18
70	22.2886	1	12
71	22.96	1	13
72	22.96	1	10
73	22.96	1	18
74	22.96	1	8
75	22.96	1	10
76	22.96	1	-
77	22.96	1	-
78	22.96	1	-
79	22.96	1	-

Table A.2: Subject, Volume Activity Mean, Group, & ADOS (ABIDE Subjects 41-79)

Subject	Mean	Control/ASD	ADOS
80	22.96	1	-
81	22.96	1	-
82	22.96	1	-
83	22.96	1	-
84	22.96	1	-
85	22.96	1	-
86	22.96	1	-
87	22.96	1	-
88	22.6732	1	-
89	22.6732	1	-
90	22.6732	1	-
91	23.3702	0	-
92	23.3702	0	-
93	23.3702	0	-
94	23.3702	0	-
95	23.3702	0	-
96	23.3702	0	-
97	23.3702	0	-
98	23.3702	0	-
99	23.3702	0	-
100	23.3702	0	-
101	23.3702	0	-
102	23.3702	0	-
103	23.3702	0	-
104	23.3702	0	-
105	23.4877	0	-
106	23.4877	0	-
107	23.4877	0	-
108	23.4877	0	-
109	23.4877	0	-
110	23.4877	0	-
111	23.4877	0	-
112	23.4877	0	-
113	23.4877	0	-
114	23.4877	0	-
115	23.4877	0	-
116	23.4877	0	-
117	23.4877	0	-
118	23.4877	0	-

Table A.3: Subject, Volume Activity Mean, Group, & ADOS (ABIDE Subjects 80-118)

Subject	Mean	Control/ASD	ADOS
119	23.4877	0	-
120	23.4877	0	-
121	23.4877	0	-
122	23.1956	0	-
123	23.1956	0	-
124	23.1956	0	-
125	23.1956	0	-
126	23.1956	0	-
127	23.1956	0	-
128	23.1956	0	-
129	23.1956	0	-
130	23.1956	0	-
131	23.1956	0	-
132	23.1956	0	-
133	23.1956	0	-
134	23.1956	0	-
135	23.1956	0	-
136	23.1956	0	-
137	23.1956	0	-
138	23.1956	0	-
139	22.9733	0	-
140	22.9733	0	-
141	22.9733	0	-
142	22.9733	0	-
143	22.9733	0	-
144	22.9733	0	-
145	22.9733	0	-
146	22.9733	0	-
147	22.9733	0	-
148	22.9733	0	-
149	22.9733	0	-
150	22.9733	0	-
151	22.9733	0	-
152	22.9733	0	-
153	22.9733	0	-
154	22.9733	0	-
155	22.9733	0	-
156	23.2502	0	-
157	23.2502	0	-
158	23.2502	0	-
159	23.2502	0	-
160	23.2502	0	-

Table A.4: Subject, Volume Activity Mean, Group, & ADOS (ABIDE Subjects 119-160)

Subject	Mean	Control/ASD	ADOS
161	23.2502	0	-
162	23.2502	0	-
163	23.2502	0	-
164	23.2502	0	-
165	23.2502	0	-
166	23.2502	0	-
167	23.2502	0	-
168	23.2502	0	-
169	23.2502	0	-
170	23.2502	0	-
171	23.2502	0	-
172	23.2502	0	-

Table A.5: Subject, Volume Activity Mean, Group, & ADOS (ABIDE Subjects 161-172)