SEED-BASED ANALYSIS ON MULTI-SITE RELIABILITY OF RESTING STATE FMRI DATA

by

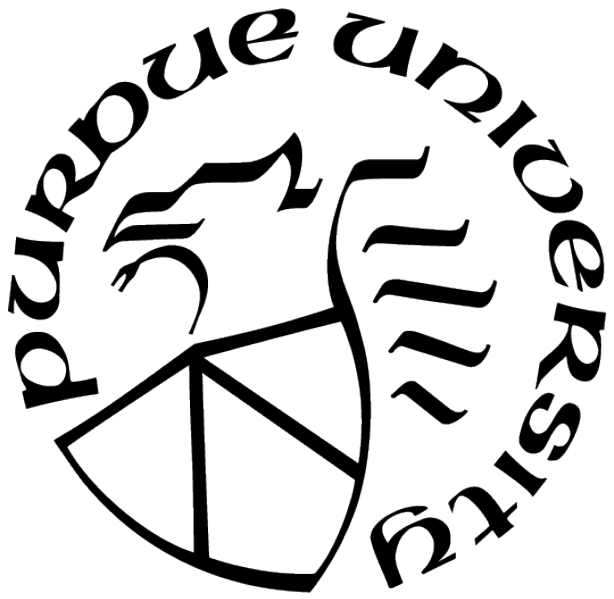
Ruihong Lyu

A Thesis

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Master of Science in Electrical and Computer Engineering



School of Electrical and Computer Engineering

West Lafayette, Indiana

May 2021

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Thomas M. Talavage, Chair

School of Electrical and Computer Engineering

Dr. Charles A. Bouman

School of Electrical and Computer Engineering

Dr. Mary Comer

School of Electrical and Computer Engineering

Approved by:

Dr. Dimitrios Peroulis

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to Prof. Thomas M. Talavage for his guidance and support over the past two years. Much appreciation to the members of my Advisory Committee Charles A. Bouman and Mary Comer for taking the time to consider this work. My gratitude also goes to Prof. Joaquin Goni Cortes, Prof. Joseph Rispoli, Dr. Kausar Abbas, Dr. Sumra Bari, Dr. Ikbeom Jang for all their academic help. I am also grateful to all my lab mates and friends, Andres Felipe Llico Gallardo, Antonia Susnjar, Arthur Terlep, Bradley Fitzgerald, Pratik Kashyap, Somosmita Mitra, Xiaoyu Ji, Xuewen Yu and Yin Jin. All their help and supports are too precious to express by words.Thanks for those who acquire and share the data. I am grateful for my parents Songnan Lyu and Yuhing Zhang who are always supportive and encouraging. I am grateful for all these people, without you, this work would not be possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Data acquisition for Magnetic Resonance Imaging (MRI) is usually expensive and time-consuming. Multi-site study enables pooling more data with less cost. However, the reliability of multi-site study is not guaranteed since the data acquired from different sites always introduces site related variations. Further, these variation can not be fully resolved even using the same imaging protocols. In this thesis, we propose a seed-based image processing and statistical analyzing pipeline which mitigates the variations brought by sites to a statistically insignificant level. We collect data from a same group of subjects on two different scanners where each subject undergoes two imaging session on each site. Seed-based correlations of BOLD timeseries are used to access the connectivity between the human brain regions and seed region. The results imply that images collected from the four visits generate similar results of seed-based connectivity. The variance brought by site-related factors, machine, visit and interaction are proved to be insignificant by ANOVA test. Moreover, principle component analysis (PCA) are performed in a manner that data are reconstructed where subject identifiability is maximized. It is shown that reconstructed data introduces less variance from interaction of machine and visit.

1. INTRODUCTION

Multi-site study has become more important in recent years. During multi-site study, data are acquired from multiple locations. Data acquisition is not a constraint compared with single-site study since they can be accessed through shared database. Multi-site study is a powerful way to enhance the scale of study while lowering the cost comparing with single-site study[1]. While multi-site study has simplified the data collecting process, it on the other hand, introduced the site-dependent variations. The data collected from multiple sites should pass quality check prior to conduct further analysis [2]. This thesis is focused on evaluating the reproducibility of functional magnetic resonance imaging (fMRI) data acquired from different sites. 16 healthy subjects are scanned twice on two different machines. A seed-based as well as a functional connectome (FC)-based methods were investigated. Evaluation of data reproducibility is based on the brain connectivity map produced from each scans. If MR data collected from different sites produce similar connectivity results, the scanner variation are considered minor.

The brain was divided into several regions of interest (ROIs) by standard brain template. In the anatomically seed-based method, the seed region was manually selected. The correlations between timeseries within each ROI and seed ROI was evaluated.The number of highly correlated regions to seed regions as well as the labels of those regions were recorded for the 4 scans on 2 different scanners. The number of highly correlated regions was passed to ANOVA analysis to access the variance as well as its significant level from each site-related factor. In the FC-based analysis, data in FC matrix that represents correlation of regions which encompass seed region and the rest of brain regions were selected. The number of highly correlated regions with regions that encompass seed region were passed to ANOVA analysis. Principal component analysis (PCA) was then applied to the original FC matrices. The same FC-based approach were repeated based on the reconstructed FC matrices. The number of regions were then passed to ANOVA. The results of ANOVA analysis based on original and reconstructed FC matrices were then cross compared for evaluating the denoising effect of PCA reconstruction.

2. BACKGROUND

2.1 Multi-site Study

Functional magnetic resonance imaging (fMRI) data acquired from single site always takes long time. Typical resting state fMRI (rs-fMRI) scans usually takes 5-7 minutes in most case [3]. Based on different type of study, fMRI scans can take up to 30 minutes during one session. Further more, each subject might undergo several imaging sessions for a study. This time-consuming process largely limits the efficiency of data acquisition. Longer the scanning time, more subject movements are introduced. But most fMRI scans are sensitive to motion [4]. Most of the time, not all data collected pass quality check. The motion effect further increase the difficulty of data acquisition. Besides, MRI study is sensitive to time as well. For instance, when subjects are scanned twice for quality check, conditions have to be satisfied that subject's brain isn't exposed to any injury between the two imaging session. The shorter the period in between the two sessions, the more accurate quality comparison. In summary, collecting fMRI data from single site is time consuming and not so efficient. In addition, MRI machines can cost from $325,000 to more than $500,000 which brings huge obstacles to funds.

Multi-site study is a promising solution for breaking the limitations of data acquisition from single site. By pooling data from multiple sites, the amount of time and funds is highly reduced while the amount of data acquired are increased. While multi-site study benefits data acquisition, it undermines the reliability of the data. Compared with single site study, multi-site study can induce more potential variations. Those variations can possibly come from types of scanners, different coils, different image sequences and so on. Even with the same type of scanners, the physics of each machine are different which can bring unexpected noise. Lots of efforts are made in sequence and protocols design which still can not fully resolve the issues[1].

2.2 Resting State fMRI

fMRI is a non-invasive imaging method that detects the fluctuations of MRI signal during the scanning time. The fluctuations are assumed to be related to blood-oxygen-level-

dependent changes [5]. According to C.J.Gauthier *et al.*[6], when certain region of the brain is performing a task, oxygen will be supplied to this region from nearby vessels and thus decrease the deoxygenated hemoglobin level and increase the blloOLD level. The correlation of fluctuations of MRI signal related to blood-oxygen-level-dependent (BOLD) variations are used for evaluating the brain connectivity.

Resting state fMRI means when subjects are undergoing the imaging session, they are not doing any specific task, for example typing or thinking. Subjects are keeping conscious during the imaging session as well. The low fluctuations detected during the resting state can be used to analyze the connectivity across regions in the brain. Brain regions that have highly correlated MRI signals during the resting state form the default mode network (DMN) [7]. DMN is the intrinsic property of human brain. Although the strength of functional connectivity across nodes in DMN might be different in different age, sex groups [8]. The important nodes are the same across healthy brains. In our method, we compare the resting state brain connectivity map with DMN. If the regions that are connected in resting-state are similar to DMN, our anatomically-based seed method is reasonable.

2.3 Seed-based Analysis

Seed region is defined as a small region in certain part of brain within which same functionality is preserved. Seed regions can be selected in multiple ways. For a task related group analysis, seed regions can be defined as the regions that are most active among all subjects during the task. Seed can be defined anatomically if it is easy to delineated from the image [9]. When defining seeds anatomically, seed is usually selected based on prior assumptions [10]. For instance, when conducting research on resting state brain connectivity, seeds are usually chosen to be the important nodes in DMN. The voxel values within seed region are averaged and used to represent the value of seed region. Seed region are usually small, such as a 5mm radius spherical region which makes the seed-based analysis representative and intuitive [10].

The posterior cingulate cortex (PCC), temporal parietal junction (TPJ), and medial prefrontal cortex (MPFC) are often considered as the most important nodes in DMN. [11]. Medial prefrontal cortex (MPFC) plays a key role in self-referential processing in healthy

subjects [12]. When conducting resting state analysis, a seed is often placed in those regions such that the connectivity maps generated can be compared with DMN.

2.4 Previous Work

To ensure the reproducibility of data acquired from different scanners, several metrics were developed.

At first, signal to noise ratio (SNR) of images collected from multiple sites was used to compare the qualities of imaging. The SNR were tuned to be similar across different sites such that images are of same quality. As pointed out by Vincent A. Magnotta *et al.* [13], signal to noise ratio (SNR) and contrast to noise ratio (CNR) are important estimators for image quality. Basal ganglia and cortical area voxels were used as regions of interest (ROIs). When conducting the multi-site study, the scanner parameters, such as repetition time (TR), time to echo (TE), field of view (FOV) and slice thickness were kept the same as long as can be provided by the scanner software. The result showed that SNR and CNR were similar across 1.5T scanners. While for scanners with larger magnetic field such as 3T ad 4T, SNR and CNR as well as variance across scanners would increase. As is obvious, not only SNR and CNR among 3T and 4T scanners vary extensively, but also those two parameters can not fully ensure the similarity of image data. In summary, Vincent A. Magnotta *et al.* [13] provided the cornerstone of multi-site fMRI data pooling.

The next step of collecting reproducible data on different scanners was through sequence/protocol harmonization. The core idea was to come up with a universal imaging protocol which produced similar images of the same subject brain across different scanners. The same subject was scanned on different MRI scanners multiple times. Each time, parameters such as TR, TE, type of sequence were adjusted. The combination of parameters that resulted in best uniformity of visualization of images were used for further data acquisition.

One concrete example was what William T. Clarke *et al.* have done in protocol harmonization across 7T MRI scanners[14]. They tested several well-established imaging sequences such as MPRAGE, MP2RAGE, GRE, multi-echo GRE and Dual-echo GRE on 5 scanners of three different type: Siemens Magnetom Terra, Siemens Magnetom 7T and Philips Achieva 7T. Every scanner used the same 32-channel receive (1Tx32Rx) head coil. Raw data was

then reconstructed using the scanner's own built-in software with slight changes. In order to test the reproducibility of the reconstructed data, the same subject was scanned one time at five different sites and the subject was scanned another four times at one site for determining intra-site stability. In this study, not only temporal SNR was used, but more statistical methods as well as visualizations of brain images were utilized to compare the data acquired from multiple sites. Their results showed that the harmonized sequence yields images with better contrast of grey-matter compared with images generated by the default sequence from each scanner. In addition, across-site as well as within-site temporal SNR were similar. Coefficient of variation of spatial activation extent and BOLD signal is similar for both inter-site and intra-site evaluations. The result is promising but still there exists several limitations. First, the designed protocol is only applicable to certain type of vendors, models as well as head coils. The protocols cannot be used in different type of scanners. Second, the physics of each machine are different even from the same vendor. Third, system stability can be different, for instance, 3T Siemens MAGNETOM Prisma MRI scanner has a magnetic field of 2.9T which is less than listed 3T. Third, systems are maintained differently which leads to more divisions in scanner physics as time goes by.[1]. Fourth, artifacts exist in T1 images collected from certain scanner while not the others [15].

At the same time, image processing as well as analyzing techniques of multi-site data were explored as important methods for assuring the robustness of images collected from different sites. The key idea was it was not the visualization of raw images that counted the most, but the unity of information conveyed by the images that mattered.

As Sumra Bari *et al.* [1] proposed, an image processing pipeline took the images collected from multiple sites as input and produced functional connectomes (FC), the pair-wise correlations across whole brain parcellations as output. The FC matrices were than used as the brain connectivity estimation. Identifiability matrices constructed based on FCs were used to evaluate the self uniqueness of same subject across different scans as thus evaluate the cross scanner differece. The idea is that if same subjects have higher self idenfiability acorss 4 scans compared with other subjects, then the data collected from multiple scanners don't differ much. In their study, 23 healthy subjects were scanned twice on two different sites with each scan on a single site being divided into test and re-test scans. Images collected from

sites were processed by Analysis of Functional NeuroImages (AFNI) platform and calling functions from FMRIB Software Library (FSL) through MATLAB code. Functional connectome (FC) matrices were created based on shen278 [16] atlas without the 30 cerebellum regions. Then the initial 248x248 FC matrix corresponding to each subject at each scan was vecotorized and stacked together into a new FC matrix, resulting in a 72-column FC matrix. The new FC matrix was later on reconstructed based on PCA. Four identifiability matrices were created by correlating the FCs from two scans within sites and FCs from two scans across sites. Four dentifiability matrices for reconstructed FCs were computed under the same manner. Since it is the fact that an individual's FC is much more correlated with itself compared with other individuals, the identifiability matrix should have higher values along the diagonal. The result showed that for identifiability matrices based on original FCs, the diagonal elements appeared to be of higher values when evaluating the two scans within sites, but it is not as obvious for the two cross-site identifiability matrices. For identifiabilitiy matrices based on principal component reconstructed FCs, the diagonals for all within and across site matrices are obviously higher. This paper indicates that PCA reconstructed FC matrix is robust and able to enhance the fingerprinting of individuals.

The original FC itself did not show much robustness for cross-site analysis following Sumra Bari *et al.*'s results [1]. Furthermore, literature shows that correlation across nodes of DMN is showing more robustness without further enhancement. Kausar Abbas *et al.* [17] conducted a seed-based study on 22 high school American football athletes as well as 10 noncollision-sport athletes. The PCC regions is chosen as seed region. Instead of evaluating the pair-wise correlations of all regions in the brain. The relationship between seed region and the rest of the brain was evaluated. The 22 high school American football athletes were scanned three times: before season, during season and after season. The brain was divided into 116 ROIs based on the Automated Anatomical Labeling atlas. Instead of accessing the correlation of whole brain, a seed ROI is created by a 12 mm radius spherical region placed at the posterior cingulate/precuneus which is an important node in DMN. Then pair-wise correlations of the 116 ROIs with seed region was then evaluated. The highly correlated regions (False discover rate adjusted p-value $\leq$0.05) of each subject at each scan was selected and counted. The result showed that players of collision-sport's number of highly correlated

regions significantly differed from the noncollision-sport athletes before the season starts. The during season and post season scans showed that the number of highly correlated regions were varying compared with baseline scan of the players.

Although this study was not conducted on multi-sites, the result implies that seed-based analysis for rs-fMRI results maintains its robustness in DMN framework.

In order to quantify the robustness of data collected from multiple scanners, repeated measures ANOVA test is performed in a similar way mentioned in Stephanie Noble *et al.* [18]. Eight traveling subjects were scanned twice at eight different sites. Posterior cingulate cortex (PCC), right motor cortex(RMC), and left thalamus (LT) were selected as seed regions. Subject, site, scanner manufacturer and the dates were used as independent factors for ANOVA analysis. The result in Stephanie Noble *et al.* [18] showed that subject difference domain the variations in seed-based connectivity as well as whole-brain atlas based connectivity. In contrast, only minor effect was found for the other independent conditions.

Based on the findings of Sumra Bari *et al.* [1] Kausar Abbas *et al.* [17] and Stephanie Noble *et al.* it is valuable to combine the two data processing methods for analyzing multi-site data. Instead of using FC matrix for evaluating the whole brain connectivity, seed-based analysis is adopted. PCA reconstruction will be conducted based on the seed-based correlations as well to further enhance the robustness of seed-based correlations. Besides, ANOVA is added such that quantitative assessment of how each site-dependent factors contribute to the difference of brain connectivity could be established.

3. METHODS

3.1 Study Subjects

A cohort of 23 students (12 males and 11 females; ages 18-28 years)[1], both undergraduate and graduate from Purdue University main campus undertook two imaging sessions at two different sites. The two scans on each site were back-to-back. The scans taken on different sites were less than 21 days apart. None of the subjects reported any history of neurological disorders. Finally, scans of 16 subjects from the 23 participants were used for data analysis. Three subjects' data were discarded because of excessive motion during any of the four imaging sessions. Two subjects' data were discarded because of failing to register T1 image to the standard brain template in MNIspace. Another two subjects were rejected because of missing cerebellum when warping from MNI space beck to standard anatomical space. Details will be given in section 3.3.

3.2 Data Acquisition

At site 1, subjects undertook imaging on a 3T General Electric Signa HDx machine with a 16-channel brain array (Nova Medical). At site 2, subjects underwent scans on a 3T GE Discovery MR750 machine with a 32-channel brain array (Nova Medical).

On both sites, each imaging session was composed of one T1 weighted scan and two rs-fMRI scans (each lasts 9 minutes and 48 seconds). The pulse sequence parameters for both T1 weighted and rs-fMRI scans were the same for scans conducted on two sites. The pulse sequence used for all high resolution T1 weighted scan was 3D fast spoiled gradient recalled echo sequence. Echo time (TE) of 1.976 msec and repetition time (TR) of 5.7 msec were used in order to achieve the desired amount of volume. Flip angle was set to 73°. Isotropic resolution was 1 mm to keep the image quality for future template registration, extraction and segmentation. Blipped echo-planar imaging sequence was used for all rs-fMRI scans. TE was set to 26 msec and TR was set to 2000 msec. Flip angle was 35°. Total of 34 slices were required with a 20x20cm field of view (FOV) and voxel size 3.125 x 3.125 x 3.80mm. 294 volumetric data points were acquired. All the data were saved in the form of DICOM files with a file extension .MRDC.

3.3 Image Pre-processing

Functions from MRIcron[19], AFNI[20], FSL[21] and MBWSS[22] were called through MATLAB in the image pre-processing step. AFNI API was also used for manual correction of some data when necessary. Standard MNI brain templates including ch2bet.nii[19], shen278.nii[16] were downloaded. ch2bet[19] is a standard anatomical brain templates in MNI space. Shen278[16] is a standard brain parcellation template in MNI space which divides the grey matter of the brain into 278 regions. AFNI was used to select two 5mm radius spherical ROIs placed at left and right medial prefrontal cortex on shen278[16] template. The result templates were called left_mpf.nii and right_mpf.nii. Both of them were in standard MNI space with only gray matters of the seed regions.

T1 weighted images were first transferred from DICOM format to NIfTI format (MRIcron *dcm2nii*). Then NIfTI images were read into 3D volumetric data (FSL *MRIread*) and denoised *MRIDenoisingONLM*). Last step of image denoising was to correct for the RF energy field (B1) inhomogeneity (FSL *fsl_anat*) by applying non-local denoising filters blockwise [21] such that all parts of the image has same level of brightness. Binary masks of the brain were generated based on the denoised and bias-field corrected T1 weighted image (MBWSS *scalperWS –refine*). In some cases, brain mask generated by algorithm was not flawless. A few slices need to be manually adjusted when brain masks do not extract the brain precisely. This process was conducted using AFNI API by fixing the boundaries and holes in the brain mask in those problematic slices by hand. After achieving the correct binary mask, the mask was then applied to the processed T1 weighted image (*fslmath –mul*) so that skull was discarded and only brain remains in the image for further image processing steps.

Warping of T1 weighted brain images were required since different people have different shapes of brain which led to ambiguity if segmenting regions directly in T1 anatomical space. A two-step registration was introduced. First, a non-linear transformation which includes both affine and warp was applied to T1 images to align the anatomical data to standard MNI template. The inverse of this operation was calculated and saved. Second, the saved inverse operation was applied to standard MNI templates including shen278[16], left_mpf and right_mpf. This step brings the standard MNI space template to each subject's T1

anatomical space. The reason we conduct further image processing on anatomical space instead of MNI space was because the images are usually more blurring in the MNI template space. For some images, this inverse operator (AFNI *INV*) created image cut-off issue. To fix this problem, the two-step registration was modified by first transforming from standard MNI template to T1 image. Then this transform was then directly applied to other templates in MNI space to bring them to T1 anatomical space.

3.4 Multi-Site Evaluation of Seed-Based Connectivity Robustness

We are going to use two different methods generating mpf seed regions to access the functional connectivity across multiple measurements and different MR systems. In the first case, seed ROIs were defined by the 5mm radius spherical placed in left mpf and right mpf regions. While in the second method, seed region was defined as six shen278[16] parcellations that cover the mpf region. Although the definitions of seed regions were different, the seed ROIs selected in the first method were within the functional regions defined in the second method. The reason we define the seed region as selected shen278[16] parcellations was that we would like to explore the effect of principal component reconstruction of functional connectome matrix have in denoising the data. In the second method, seed regions data were read from raw FC matrix directly. During PCA step, seed regions data were selected from corresponding columns of reconstructed FC matrix. For both analysis methods, 11 p-values starting from 0.0002 to 0.05 with an 0.005 increment were used as criteria to evaluate the robustness of our analysis. The first p threshold is set to 0.0002 based on bonferroni correction since multi-comparison were conducted at 290 time points.

3.4.1 Evaluation of Anatomically-Based Seed Definition

Shen278[16] template in T1 anatomical space was first resampled to have the same resolution as the rs-fMRI images. Then the resampled shen278[16] parcellation was used to divide the whole brain into 278 functional regions in rs-fMRI space. The 4D timeseries data was averaged over all voxels within each region at each time point. This resulted in a 278x290 matrix of which each row represented the region value at 290 time points. The same process

was repeated for left mpf as well as right mpf seed region. Since seed region was treated as a single ROI, we will have two row vectors of length 290. Each element in the vector was the average of all voxel values within the seed region at certain time point.

Correlations as well as corresponding p-values between seed ROIs and 278 regions were calculated based on the timeseries data (MATLAB *corr*). The output matrix is of dimension 2x278. The first row of the matrix records the correlations between seed ROI and 278 brain regions, the second row represents the p-value for each correlation. A region is considered as highly correlated to seed if its p-value is less than p threshold.

In addition, out of the selected highly correlated regions, thirty regions corresponding to cerebellum were discarded. The cerebellum was not used for analysis since not all subjects had correct warping of cerebellum regions.

After getting rid of the cerebellum region, the regions marked as highly correlated as well as the total number of highly correlated regions based on all 11 p-thresholds were saved. This process was repeated for all 16 subject in all 4 scans. The saved number of highly correlated regions statistics were used in ANOVA test introduced in session 3.5.1. The regions selected were drawn into 3D brain graphs for visual comparison with DMN.

3.4.2 Functional Connectome Based Analysis

Function connectomes were calculated by finding the pair-wise correlations of 278 time-series data. Seed region consisted of six shen278[16] regions that encompass mpf region. Function connectomes were used in principal component reconstruction. Besides the thirty cerebellum regions, brain stem region, with label 247 was also deleted for further data analysis, leaving a 247x247 FC matrix. The brain stem needs to be deleted since most of the voxel values in timeseries within that region were not available. Using matrix containing NaN values for PCA reconstruction creates error in the program. Brain stem did not create a problem for seed ROI based correlation since for MATLAB operator *corr*, NaN were set to 0 by default.

Then the pair-wise correlation values between those six seed regions and other 241 regions were extracted from the FC matrix. P-values were calculated based on the method

discussed in session 3.4.3. Within 241 regions, a region was selected as highly correlated if it was highly correlated to at least one of the six seed regions based on statistical test. Furthermore, the original FC matrix was reconstructed based on PCA introduced in session 3.4.4. The highly correlated regions were selected based on the reconstructed matrix. The number of highly correlated regions was recorded. This process was repeated for 16 subjects among 4 imaging sessions as well. The saved number of highly correlated regions statistics were used in ANOVA test introduced in session 3.5.1. Noting that significance of correlations was evaluated using a Pearson correlation coefficient and a Fisher Transform via the MATLAB commande *tcdf*.

3.4.3 Effect on Multi-Site Stability of Identifiability Enhancement

As mentioned by Sumra *et al*, when analyzing the brain connectivity based on function connectomes, reconstructing FC matrices at which a maximum self uniqueness is achieved will largely improve the self identifiability across scans collected from different sites [1]. In other words, PCA reconstruction of FC matices will denoise the rs-fMRI data in a sense that subject variation domains the total variance of the data. We would like to access the effect of reconstruction based on self-identifiability enhancement here as well. The reconstruction was based on the information provided by all 64 (16 subjects, 4 visits) FC matrices. FC matrix is symmetric. Elements along the diagonal were all 1. Considering these two properties, upper triangle elements, besides diagonal elements, of each matrix were chosen and vectorized. Each 247x247 FC matrix produced a vector of 30381 elements. The total 64 vectors were stacked together, forming a 64x30381 shaped matrix.

The number of principal components was determined by identifiability. Self-identifiability matrix was established by correlate the timeseries from the same subject across 4 scans. Others-identifiability was defined as the average correlations among different subjects in 4 scans. The aim of principal component reconstruction was to improve the self-identifiability of each subject, meaning maximizing the uniqueness of each person. Following this logic, the number of principal components was selected where the difference between self-identifiability and others-identifiability was maximized.

MATLAB command *pca* was used for reconstructing the original 64x30381 matrix. The

number of principal components used varies from 1 to 64. During each reconstruction, difference between self-identifiability and others-identifiability was calculated. In our experiment, the self uniqueness was most obvious when number of principle components was selected as 19.

There raised one problem when reconstructing FC matrix based on PCA—some entries of the reconstructed matrix could possibly have an absolute value that is greater than 1 which makes the correlation not meaningful in real world. In response to this, Fisher transform (z-transform) was used [23]. Instead of reconstructed based on the correlations, we applied z-transformation on original data (MATLAB *atanh*). After having the reconstructing matrix of which each element was a z-score for corresponding correlation value, the matrix was then transformed back to the original correlations domain (MATLAB *tanh*). Since correlation values r within the original FC matrix was always between -1 and 1, the Fisher transformation brought the original data to range [-inf, inf]. The inverse Fisher transformed z-scores went back to the domain [-1, 1], which overlapped the range of correlations.

3.5 Statistical Analysis

3.5.1 Two-way ANOVA Test with Repeated Measures

The goal of this thesis is to estimate the reliability of multi-site data. The reliability of data was evaluated through ANOVA (Analysis of Variance) of two independent factors: machine and scan. Each factor has 2 groups, we denote them as machine 1, machine 2 and scan 1, scan 2. We are also interested in investigating whether there are interactions within those two factors. Since the same 16 subjects underwent two different conditions—machine and scan, we adopted two-way ANOVA with repeated measures for our data analysis. The statistical software SPSS [24] is used for this analysis.

As mentioned earlier, the aim is not only to evaluate how much impact does each factor have on the total variance of the data mean but also how much change does those two independent factors bring to each other. We used software SPSS for ANOVA test. Statistics including F value, significant value as well as degree of freedom were reported. In ANOVA, we set the significance threshold to be 0.05. Any factor variance that has a significance more

than 0.05 is considered to be insignificant. If some factors are recognized as significant, more statistical test will be performed as confirmation.

4. RESULTS

4.1 Results Based on Seed ROIs Method

Table 4.1 represents the total number of highly correlated regions over all 16 subjects for each imaging session. Each row compares statistics across 4 scans. This table contains 11 rows corresponding to data achieved based on 11 p thresholds. There is no significant variation across 4 imaging sessions based on result of ANOVA test.

Table 4.1. Total # of Regions

p threshold	machine 1,scan 1	machine 1,scan 2	machine 2,scan 1	machine 2 ,scan 2
.0002	2443	2690	2634	2563
.005	2916	3117	3014	3026
.010	3036	3225	3103	3144
.150	3125	3284	3044	3196
.200	3181	3324	3193	3266
.250	3229	3360	3228	3307
.300	3269	3386	3264	3332
.350	3289	3396	3293	3364
.400	3317	3420	3316	3381
.450	3348	3441	3340	3398
.500	3371	3460	3360	3415

Figure 4.1-4.3 describe the ranking of each 16 subjects in the 4 scans based on the number of highly correlated regions. The subject with the smallest number of highly correlated regions has ranking 1. Figures contain results evaluated on 11 different p-thresholds. From the figures, rankings of subjects are stable across 4 scans. Based on the scattered plot, there is no pattern showing that either machine concludes more regions than the other. It is true for all 11 p thresholds.

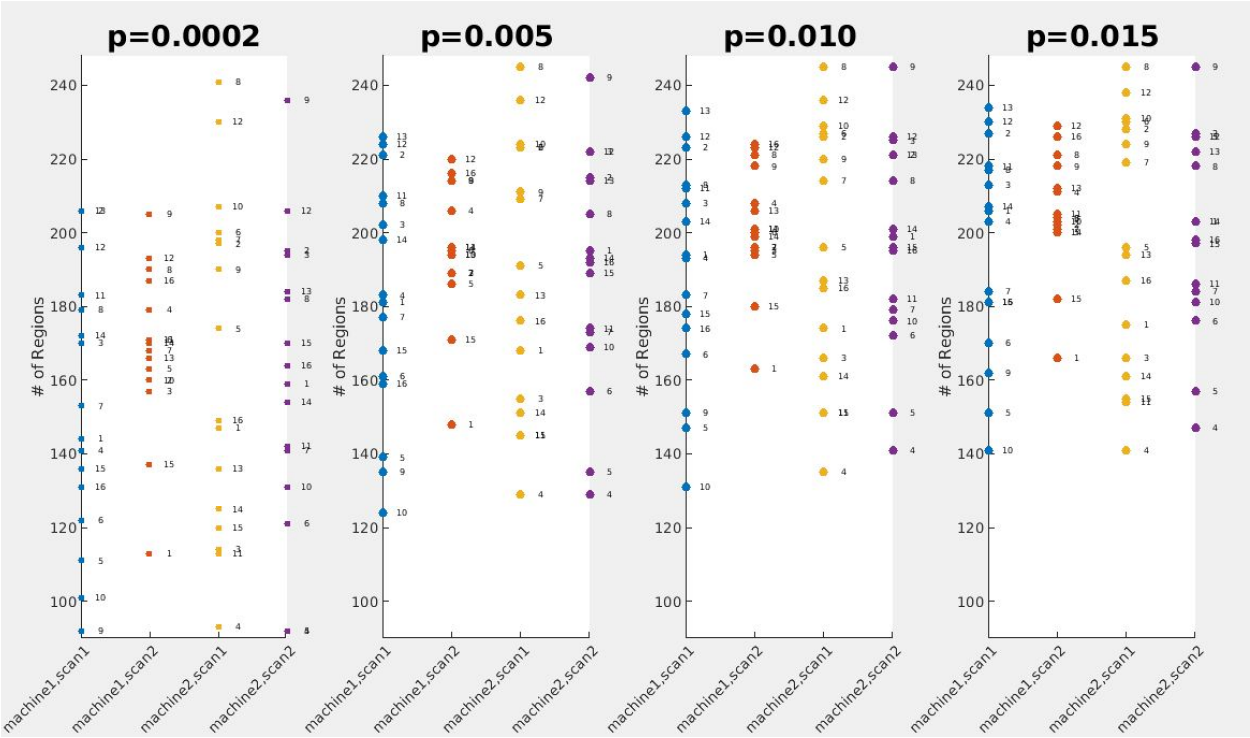


Figure 4.1. Subject ranking based on # of regions

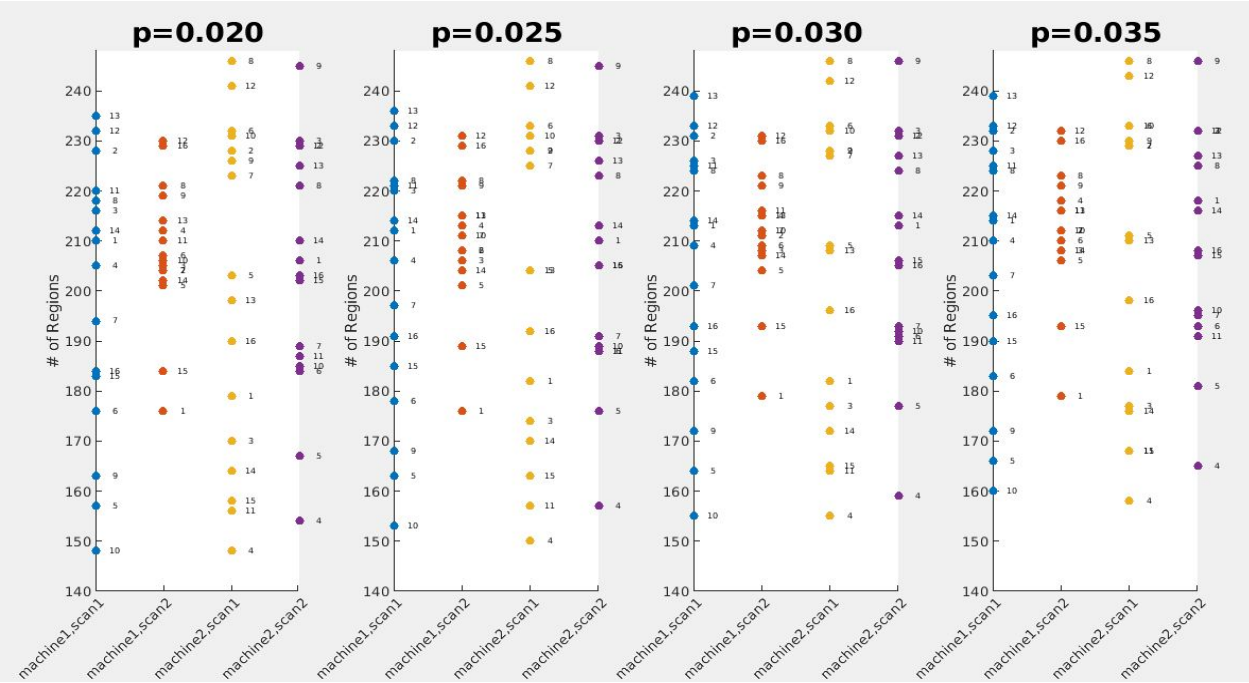


Figure 4.2. Subject ranking based on # of regions

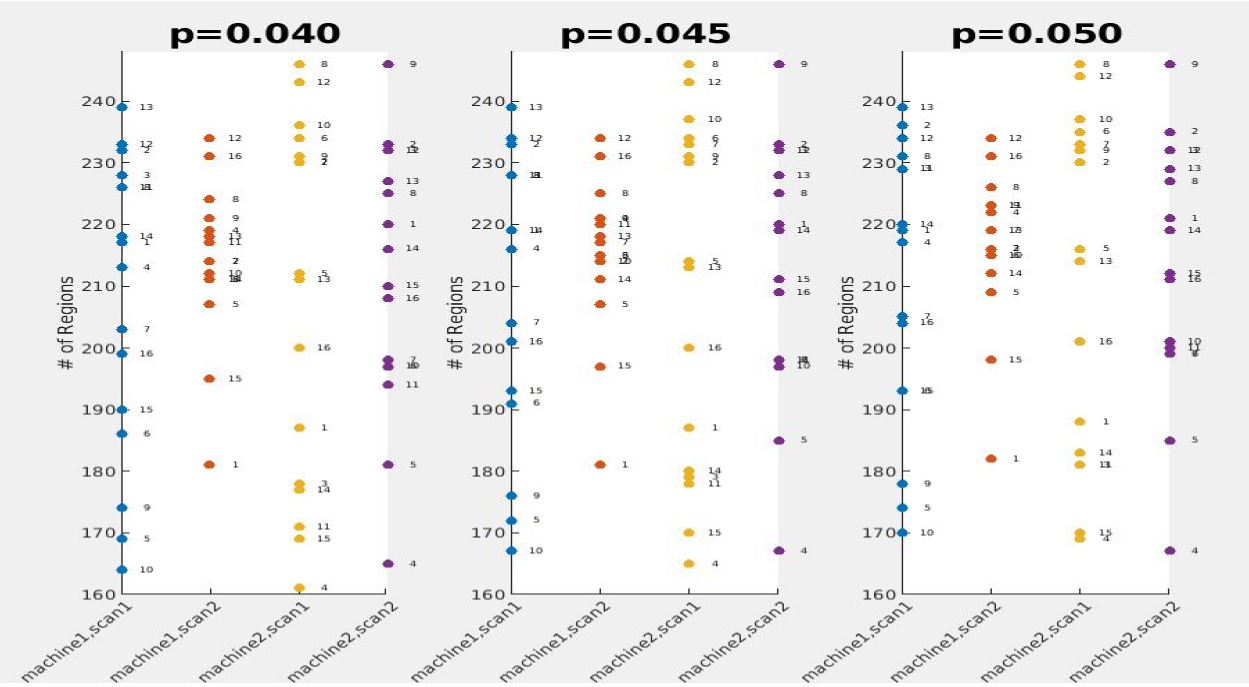


Figure 4.3. Subject Ranking Based on # of Regions

Figure 4.4 describes the relationship of the average number of highly correlated regions of total 16 subjects with logarithm of 11 p over 4 scans. Points on x-axis labeled as p_{th} have actual value $log(p_{th})$. The standard deviations of machine 1, scan 1 and machine 2, scan 2 are plotted as error bars on the graph as well. The two scans from machine 1 have the highest as well as lowest average number of regions across 11 p thresholds. The 2 scans from machine 2 recognize number of regions in between. This means that there isn't an noticeable variance brought by machine. Also, based on the statistical test result from ANOVA (table 4.2), there isn't a significant machine or visit variance.

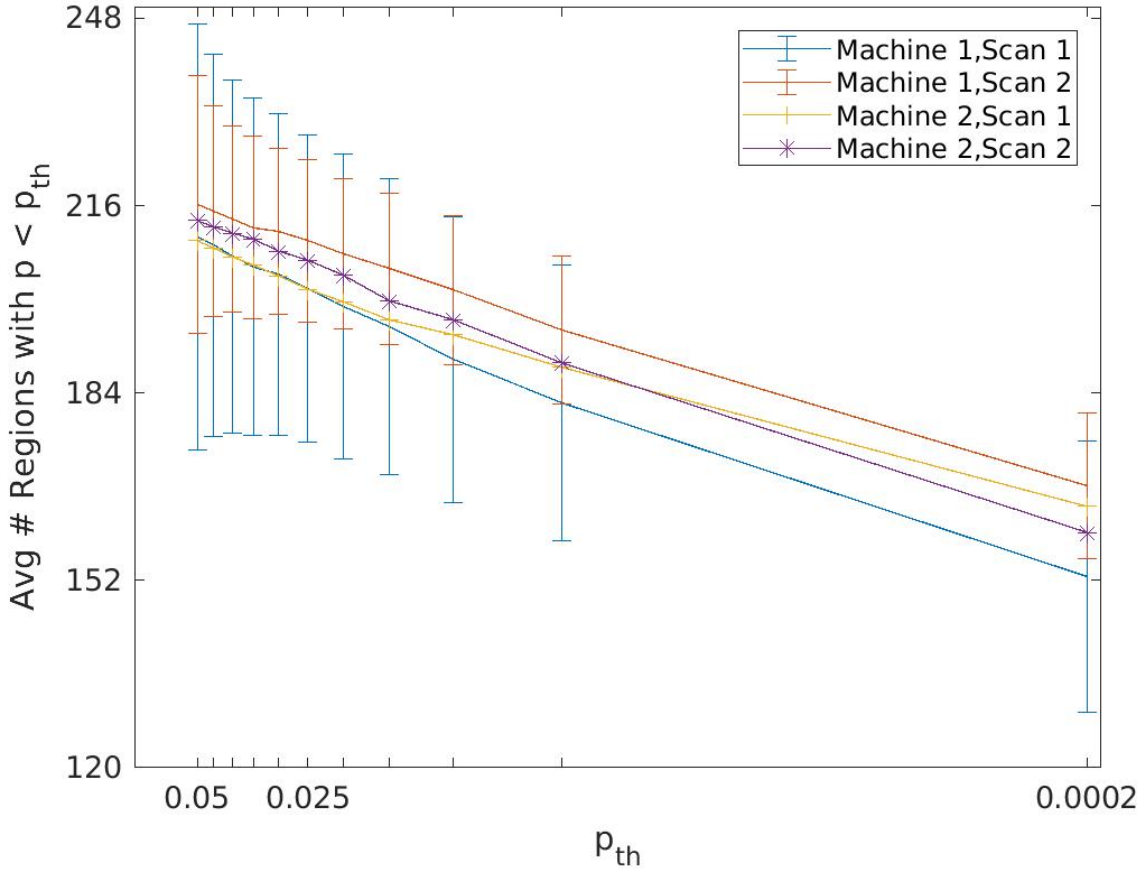


Figure 4.4. Average # of Regions With Respect to P Thresholds

Table 4.2 summarizes the results of ANOVA test based on seed ROI method. Significance value columns represent the p thresholds of the F-test for each factor machine, visit and interaction between machine and visit. If the significant value is less than 0.05, it means these factors are important. Results show that across 11 p-values, there isn't a recognizable factor that contributes to the total variance.

Table 4.2. Summary of ANOVA Test F(1,15)

	Machine		Scan		Interaction	
P threshold	F	Sig.	F	Sig.	F	Sig.
.0002	0.057	0.841	0.612	0.446	1.057	0.320
.005	0.001	0.976	1.570	0.229	0.531	0.477
.010	0.004	0.949	2.712	0.120	0.371	0.552
.015	0.107	0.748	2.776	0.116	0.206	0.656
.020	0.052	0.823	2.961	0.106	0.098	0.758
.025	0.076	0.787	3.239	0.092	0.061	0.809
.030	0.095	0.762	2.727	0.119	0.056	0.817
.035	0.022	0.883	2.740	0.119	0.032	0.861
.040	0.046	0.833	2.623	0.126	0.038	0.849
.045	0.081	0.779	2.095	0.168	0.033	0.858
.050	0.103	0.753	1.974	0.180	0.033	0.858

Figure 4.5 is an example of DMN [25].Regions marked in blue are regions that are known to be highly correlated to MPF seed region. If we cross compare this figure with figure 4.6-4.8, we will find that the important nodes in DMN marked in blue here are included in the regions recognized by our seed-based method.

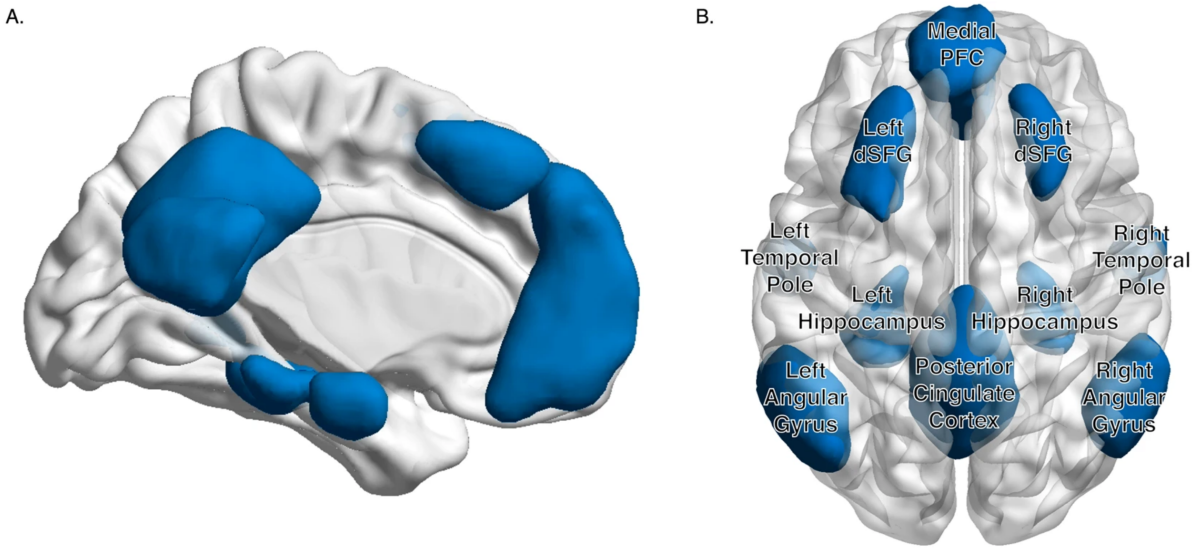


Figure 4.5. Regions of Default Mode Network (DMN)that are Known to Be Highly Correlated With MPF Seed Region [25]

Figure 4.6-4.8 represent the brain regions that are marked as highly correlated out of 4 scans. If the region is recognized 4 times, 3 times, 2 times, 1 time and 0 time, the region will be painted as yellow, red, green, blue and white respectively. The three graphs use 0.02%, 2.52% and 5% as p-value sequentially. Six views—front, back, left, right, top and bottom are used for representation. The regions marked by red and yellows are regions that are statistically more possible to be highly correlated to MPF. Those regions marked by yellow and red are highly overlapped with DMN shown in figure 4.8. This indicates that the 4 scans provide accurate information for detecting brain connectivity.

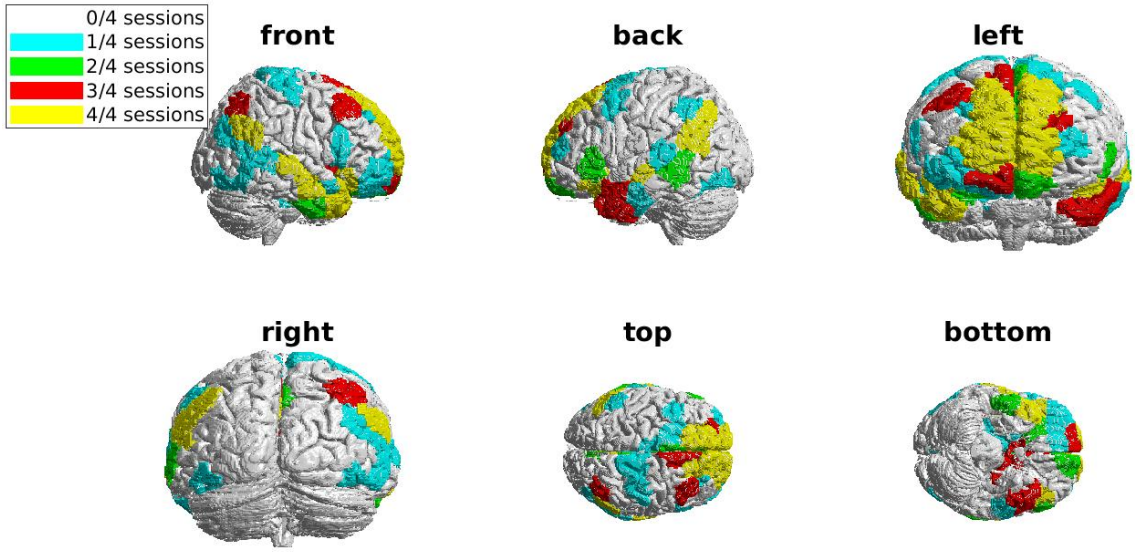


Figure 4.6. Regions from shen278[16] Parcellation Exhibiting Correlation with the MPF Seed That Exceeds A Significance Threshold of $p<0.0002$

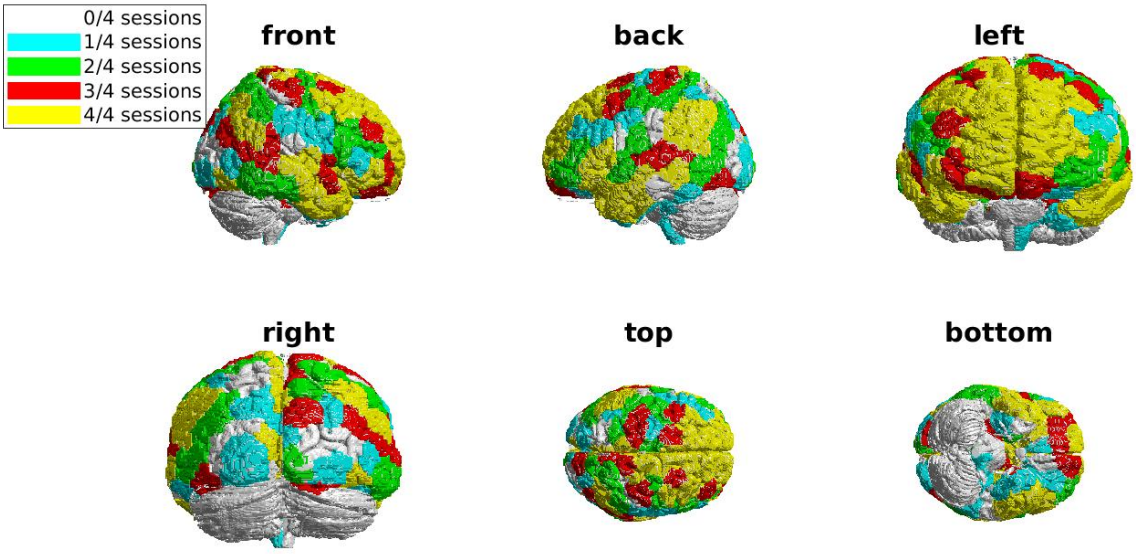


Figure 4.7. Regions from shen278[16] Parcellation Exhibiting Correlation with the MPF Seed That Exceeds A Significance Threshold of p<0.025

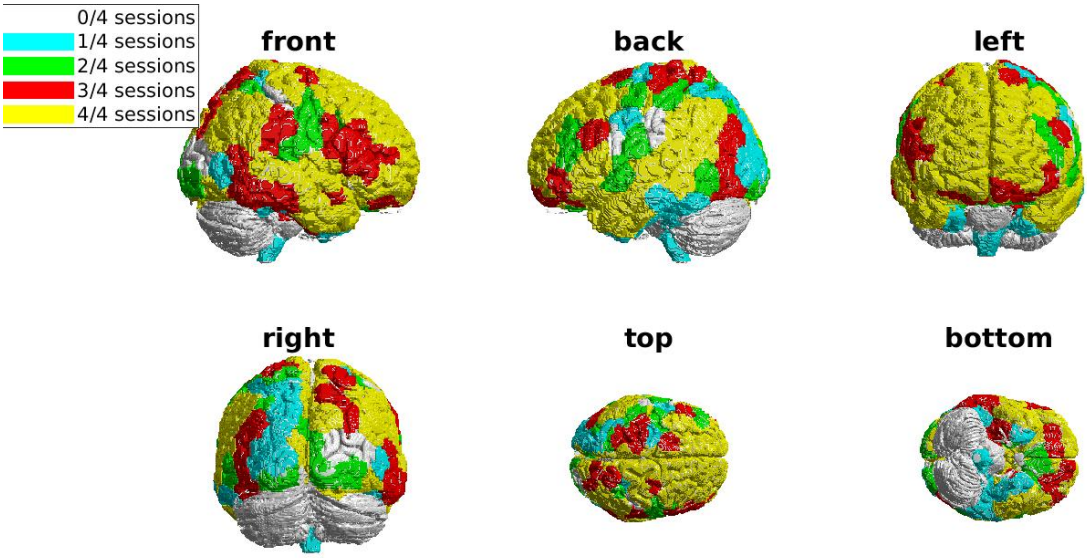


Figure 4.8. Regions from shen278[16] Parcellation Exhibiting Correlation with the MPF Seed That Exceeds A Significance Threshold of p<0.05

4.2 Results Based on Function Connectomes Method

Table 4.3 presents the results of ANOVA analysis based on original FC matrices that defines seed region with shen278 [16] atlas that encompass MPF. F value states the variance of means contribute by each factor. Significance value columns represent the p values of the F-test for each factor machine, visit and interaction between machine and visit. If these p-values are less than 0.050, it means these factors are significant at 0.050 level. As shown, none of the factors have a significant effect in variance contribution.

Table 4.3. Summary of ANOVA Test Based on Original FC F(1,15)

	Machine		Scan		Interaction	
P thresholds	F	Sig.	F	Sig.	F	Sig.
.0002	0.015	0.903	0.017	0.898	0.011	0.917
.005	0.037	0.851	0.078	0.783	0.136	0.717
.010	0.003	0.956	0.055	0.817	0.145	0.709
.015	0.001	0.973	0.041	0.843	0.150	0.704
.020	0.003	0.957	0.058	0.813	0.188	0.671
.025	0.001	0.971	0.046	0.833	0.132	0.722
.030	0.003	0.956	0.047	0.832	0.184	0.674
.035	0.000	0.985	0.100	0.756	0.131	0.723
.040	0.002	0.969	0.098	0.759	0.119	0.735
.045	0.007	0.937	0.067	0.800	0.120	0.733
.050	0.006	0.940	0.078	0.784	0.170	0.686

Table 4.4 presents the results of ANOVA analysis based on self-identifiability enhanced FC matrices. Seed region is defined as shen278 [16] atlas that encompass MPF. F value states the variance of means contribute by each factor. Significance value columns represent the p values of the F-test for each factor machine, visit and interaction between machine and visit. As shown, none of the factors have a significant effect in variance contribution.

Table 4.4. Summary of ANOVA Test Based on Reconstructed FC F(1,15)

	Machine		Scan		Interaction	
P thresholds	F	Sig.	F	Sig.	F	Sig.
.0002	1.485	0.242	0.810	0.382	0.000	0.984
.005	0.584	0.457	0.214	0.650	0.001	0.978
.010	0.606	0.449	0.771	0.006	0.009	0.926
.015	0.639	0.437	0.089	0.770	0.001	0.974
.020	0.828	0.377	0.088	0.771	0.002	0.969
.025	0.800	0.385	0.107	0.748	0.005	0.946
.030	0.798	0.396	0.082	0.778	0.010	0.923
.035	0.681	0.422	0.073	0.791	0.009	0.928
.040	0.626	0.441	0.061	0.808	0.004	0.950
.045	0.651	0.432	0.072	0.792	0.003	0.960
.050	0.696	0.417	0.070	0.795	0.010	0.920

4.2.1 Analysis of the Effect of Effort to Enhance Identifiability

If we compare talble 4.3 with table 4.4. We will notice from the F values that at all p thresholds, machineas well as scan variance becomes larger after identifibility enhancement. Although F values are becoming larger, their corresponding significance value tells that that none of them are statistically significant. One thing to notice is that the variance from the interaction term is decreased. This means when fix the machine, the difference between 2 scans on the same machine is reduced. After enhancing the identifiability of subjects, both machine give a more stabilized output across scans.

5. DISCUSSION

5.1 Study Overview

In order to access the reliability of multi-site study, we proposed a seed-based approach for analyzing brain connectivity. If the scans on different machines provided similar brain connectivity, for example, similar regions as well as similar number of regions connected to MPF, we concluded that our image processing as well as data analysing pipeline to be reliable for multi-site study. Specifically, the same 23 subjects were scanned on two different type of scanners: 3T General Electric Signa HDx and 3T GE Discovery MR750. Two back-to-back scans were conducted on each scanner for all 23 subjects. 16 subjects' scans passed quality check and were used for further analyzing.

The whole brain was segmented into 278 functional parts using shen278 brain atlas [16]. 30 cerebellum regions were discarded because of warping issue. The BOLD timeseries signals of each region was used for evaluating the correlations between the 248 regions and the spherical seed MPF defined by AFNI. Since all images collected by scanners were first warped to MNI space, in which MPF was divided into left MPF and right MPF. Regions that were highly correlated to either of them were marked. The criteria for high correlation was based on Pearson's correlation coefficient. 11 p-values were used to enhance the robustness of our approach. Table 3.1, Table 3.2, Figure 3.1 and Figure 3.4 present the details of the robustness of our approach.

To access the actual variance brought by machine and visit and their interaction, we used ANOVA to quantitatively measure those effects. Table 3.3 summarizes the result of ANOVA test.

Besides regular seed-based method, we also explored the effect of PCA reconstruction proposed by Sumra *et. al* [1] had on further aligning data collected from different sites. Seed-based approach was altered by defining the seed-region based on shen278 atlas instead of using AFNI. Since reconstruction was based on 64 FC matrices, defining seed regions as parcellations enabled us to read pair-wise correlations directly from the original and the reconstructed matrices.

We used ANOVA test to evaluate the variance brought by different independent factors:

machine and visit and their interactions. Comparison of test results based on original FC matrices as well as reconstructed matrices are shown in Table 3.4.

5.2 Findings

Our seed based approach was robust since the total number of regions was similar across 4 scans. Based on Table 4.1, across 11 p-values, the total number of highly correlated regions sum over 16 subjects were similar across 4 scans. By comparing the numbers collected from different scanners, there was no obvious trend showing that one of the scanners always picks more regions than the other. Actually, machine variance was less obvious than variance brought by visits which were discussed in details when analyzing ANOVA results. Also, the observation across 11 p-values were similar—there did not exist a certain p-value at which showed significant difference across 4 scans. This indicated that our method was robust regardless of using strict or loose criteria.

Figures 4.1-4.3 sorts subjects at 4 scans by number of highly correlated regions. The difference brought by machine, visit and interaction should all mitigate at each one out of the four rankings since for each comparison, each subject was compared against its peers under the same condition. The ranking of each subject should be dominated by subject difference. The variance across rankings of the same subject at 4 scans theoretically stemed from random noise. As we can see, for all 16 subjects, evaluated at 11 p-values, 16/16 of them had at least 2/4 rankings that were similar (with difference less than 2). At least 5/16 subjects had 3/4 rankings that were similar (with difference less than 2). Observations were similar across 11 p-values. This table indicated that our image and data processing pipeline were reliable across all subjects, scanners and visits.

In Figure 4.4, average number of highly correlated regions of 16 subjects at each scan was plotted with respect to $-log(p)$. Furthermore, it was obvious from the plot that images collected from machine 1 on first scan provided the lowest number of highly correlated regions and the images collected on machine 1, scan 2 always provided the highest number of regions by our algorithm. The data provided by the other machine produces results in between. This indicated that there was no obvious machine difference between the 2 scanners. One interesting phenomenon was that based on the graph, the second scans produced more

number of highly correlated regions on both machines.

Figures 3.2-3.4 are brain plots with selected regions marked in different colors. It is shown that as we make the test threshold (p-value) looser, more regions of the brain will be colored. In addition, the regions that were already recognized at a tighter p-value were recognized more times out of the 4 scans. This again enhanced the robustness of our analysis. The regions colored yellow were those regions that were recognized 4 out of 4 times, which were statistically more important. Those regions mostly overlapped with the important nodes in DMN. Since MPF was an important node in DMN, the consistency of our study with DMN suggests that our pipeline worked reasonably well in extracting essential information from different sources of data.

In summary, previous findings were mostly based on visual analysis. To provide statistical evidence, Table 4.2 offers statistics on how much impact each factor has on the total mean variance of the data. None of the significant value were less than 0.05, suggesting neither factors nor interaction was important such that contributes to variance of the data in a non-negligible sense. Based on the ANOVA test, we confirmed that seed-based analysis on multi-site brain connectivity study was robust since it reduced all the variances factor brought by site to negligible ranges.

We are also interested in measuring how much effect PCA reconstruction has on mitigating the variance brought by site. Table 3.4 lists the ANOVA test results based on original FC matrices and the reconstructed matrices. If we define MPF regions by shen278 atlas, based on original FC matrices, none of the factors or interaction produced a significant variance contribution which was consistent with seed ROI based method. The result based on PCA reconstructed FC matrices indicates that the variance brought by machine will increase while the variance introduced by interaction will decrease. This meant that on the same machine, the difference between visits would decrease. This actually implied that PCA reconstruction can improve the stability of data collected on a single machine. While the variance from machine increased, it was still non-significant.

5.3 Advantages of Proposed Method

We now summarize the positive findings. First, seed-based analysis, no matter whether defining seed regions by 10mm spherical region or by shen278 atlas, exhibited robustness in analyzing brain connectivity from multi-site data. In our designed test on 11 p thresholds, input data collected from multiple sites generated similar number of highly correlated regions regarding MPF across all p thresholds. Moreover, not only the number of selected regions were close, but also the regions themselves were the same. The regions marked as highly correlated from the 4 scans were also important nodes in DMN which strongly supported the accuracy of our pipeline. Second, the PCA reconstruction improved the reproducibility of data collected on the same machine. Third, seed-based analysis manifested more robustness compared with using FC matrix sorely. As mentioned by Sumra *et al* [1], correlating FC matrices from the same sites produced strong correlations for the same subjects while correlating original FC matrices from multiple site did not show a strong identifiability for the same subjects. This indicated that by using whole brain connectivity, sites can bring visible difference to the result. However, seed-based method did not produce any obvious site related differences. In addition, seed-based method could also benefit researchers who do not aim at comparing data from multi-site, but need to pool data from different places for their own purposes.

5.4 Limitations of Current Work

Although seed-based method showed robustness in conducting multi-site study, it was not flawless. Only one seed region was used which cannot represent the whole brain connectivity. Another interesting phenomenon we noticed in Figure 3.1 was that the second scan always caught more regions than the first scan on both scanners. The reason behind it was not yet clear. This may happen because the MRI machine took more time to warm up or because there were more neural activities happening in subject's brain during the second scan which produced more highly correlated regions. Or this observation was totally random. Besides, PCA reconstruction did lower down the variance introduced by interaction, but at the same time it increased the machine as well as visit variances which was not ideal.

6. CONCLUSION

In this study, we evaluated the multi-site reliability of seed-based functional connectivity, both using an anatomically defined and a constructed seed region. The latter is further evaluated for the benefit of identifiability enhancement.

Results show that the there are not any significant difference in total number of regions across four different scans estimated at all p-values. The averaged number of regions over four scans exhibit a closely linear relationship with respect to $log(p)$. The plots are not unique to machines, meaning there is not a pattern that some machine recognized more regions than the other. Regions recognized by four scans are almost identical. ANOVA test shows that the variations induced by site-related factors are insignificant based on F-test ($p<0.05$). Identifiability enhancement will bring the effect of interaction down but machine and visit up. These findings suggest that our proposed method is reliable for multi-site data pooling since the variance from site-related factors are low.

Possible extension of future work can be done by increasing the number of seed regions. Introducing more seed regions will better cover the connectivity of the whole brain. At the same time, the reason that the second scans always pick up more regions than the first scans can be explored by involving more subjects and more sites as well as increasing the number of scans on each site. Besides that, the reason that identifiability enhancement increases machine and visit variance can be explored.

REFERENCES

[1] S. Bari, E. Amico, N. Vike, T. M. Talavage, and J. Goñi, "Uncovering multi-site identifiability based on resting-state functional connectomes," *Neuroimage*, vol. 202, 2019. DOI: https://doi.org/10.1016/j.neuroimage.2019.06.045.

[2] M. G. Kahn, M. A. Raebel, J. M. Glanz, K. Riedlinger, and J. F. Steiner, "Pragmatic framework for single-site and multisite data quality assessment in electronic health record-based clinical research," *Med Care*, vol. 50, pp. 9–21, 2012. DOI: https://doi.org/10.1097/MLR.0b013e318257dd67.

[3] R. M. Birn, E. K. Molloy, R. Patriat, T. Parker, T. B. Meier, G. R. Kirk, V. A. Nair, M. E. Meyerand, and V. Prabhakaran, "The effect of scan length on the reliability of resting-state fmri connectivity estimates," *Neuroimage*, vol. 83, pp. 550–558, 2 2013. DOI: https://doi.org/10.1016/j.neuroimage.2013.05.099.

[4] F. Godenschweger, U. Kägebein, D. Stucht, U. Yaracha, A. Sciarra, R. Yakupov, F. Lüsebrink, P. Schulze, and O. Speck, "Motion correction in mri of the brain," *Phys Med Biol.*, vol. 61, pp. 32–56, 5 2016. DOI: https://doi.org/10.1088/0031-9155/61/5/R32.

[5] N. K. Logothetis, "The underpinnings of the bold functional magnetic resonance imaging signal," *Journal of Neuroscience*, vol. 23, no. 10, pp. 3963–3971, 2003, ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.23-10-03963.2003. [Online]. Available: https://www.jneurosci.org/content/23/10/3963.

[6] C. Gauthier and A. Fan, "Bold signal physiology: Models and applications," *Neuroimage*, vol. 187, pp. 116–127, 2019. DOI: https://doi.org/10.1016/j.neuroimage.2018.03.018.

[7] S. J. Broyd, C. Demanuele, S. Debener, S. K. Helps, C. J. James, and E. J. Sonuga-Barke, "Default-mode brain dysfunction in mental disorders: A systematic review," *Neuroscience & Biobehavioral Reviews*, vol. 33, no. 3, pp. 279–296, 2009, ISSN: 0149-7634. DOI: https://doi.org/10.1016/j.neubiorev.2008.09.002. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0149763408001504.

[8] L. E. Mak, L. Minuzzi, G. MacQueen, G. Hall, S. H. Kennedy, and R. Milev, "The default mode network in healthy individuals: A systematic review and meta-analysis," *PubMed*, vol. 7, pp. 25–33, 1 2017. DOI: https://doi.org/10.1089/brain.2016.0438.

[9] J. O'Reilly, M. Woolrich, T. Behrens, S. Smith, and H. Johansen-Berg, "Tools of the trade: Psychophysiological interactions and functional connectivity," *Social cognitive and affective neuroscience*, vol. 7, pp. 604–9, May 2012. DOI: 10.1093/scan/nss055.

[10] H. Lv, Z. Wang, E. Tong, L. Williams, G. Zaharchuk, M. Zeineh, A. Goldstein-Piekarski, T. Ball, C. Liao, and M. Wintermark, "Resting-state functional mri: Everything that nonexperts have always wanted to know," *American Journal of Neuroradiology*, 2018, ISSN: 0195-6108. DOI: 10.3174/ajnr.A5527. eprint: http://www.ajnr.org/content/early/2018/01/18/ajnr.A5527.full.pdf. [Online]. Available: http://www.ajnr.org/content/early/2018/01/18/ajnr.A5527.

[11] S. Wang, L. J. Tepfer, A. A. Taren, and D. V. Smith, "Functional parcellation of the default mode network: A large-scale meta-analysis," *Scientific Reports*, vol. 10, no. 1, p. 16 096, 2020, ISSN: 2045-2322. DOI: 10.1038/s41598-020-72317-8.

[12] J. H. Jang, W. H. Jung, D.-H. Kang, M. S. Byun, S. J. Kwon, C.-H. Choi, and J. S. Kwon, "Increased default mode network connectivity associated with meditation," *Neuroscience Letters*, vol. 487, no. 3, pp. 358–362, 2011, ISSN: 0304-3940. DOI: https://doi.org/10.1016/j.neulet.2010.10.056.

[13] V. A. Magnotta, L. Friedman, and F. BIRN, "Measurement of signal-to-noise and contrast-to-noise in the fbirn multicenter imaging study," *J Digit Imaging*, vol. 19, pp. 140–147, 2 2006. DOI: https://doi.org/10.1007/s10278-006-0264-x.

[14] W. T. Clarke, O. Mougin, I. D. Driver, C. Rua, A. T. Morgan, M. Asghar, S. Clare, S. Francis, R. G. Wise, C. T. Rodgers, A. Carpenter, K. Muir, and R. Bowtell, "Multi-site harmonization of 7 tesla mri neuroimaging protocols," *Neuroimage*, vol. 206, 1 2020. DOI: https://doi.org/10.1016/j.neuroimage.2019.116335.

[15] M. D. Fox, A. Z. Snyder, J. L. Vincent, M. Corbetta, D. C. Van Essen, and M. E. Raichle, "The human brain is intrinsically organized into dynamic, anticorrelated functional networks," *Proceedings of the National Academy of Sciences*, vol. 102, no. 27, pp. 9673–9678, 2005, ISSN: 0027-8424. DOI: 10.1073/pnas.0504136102. eprint: https://www.pnas.org/content/102/27/9673.full.pdf. [Online]. Available: https://www.pnas.org/content/102/27/9673.

[16] X. Shen, E. S. Finn, D. Scheinost, M. D. Rosenberg, M. M. Chun, X. Papademetris, and R. T. Constable, "Using connectome-based predictive modeling to predict individual behavior from brain connectivity," *Nature protocols*, vol. 12, pp. 506–518, 3 2017. DOI: https://doi.org/10.1038/nprot.2016.178.

[17] K. Abbas, T. E. Shenk, V. N. Poole, E. L. Breedlove, L. J. Leverenz, E. A. Nauman, T. M. Talavage, and M. E. Robinson, "Alteration of default mode network in high school football athletes due to repetitive subconcussive mild traumatic brain injury: A resting-state functional magnetic resonance imaging study," *Brain Connectivity*, vol. 5, 2014. DOI: https://doi.org/10.1089/brain.2014.0279.

[18] S. Noble, D. Scheinost, E. S. Finn, X. Shen, X. Papademetris, S. C. McEwen, C. E. Bearden, J. Addington, B. Goodyear, K. S. Cadenhead, *et al.*, "Multisite reliability of mr-based functional connectivity," *Neuroimage*, vol. 146, pp. 959–970, 2017.

[19] C. Holmes, R. Hoge, L. Collins, R. Woods, A. W. Toga, and A. C. Evans, "Enhancement of mr images using registration for signal averaging," *J Comput Assist Tomogr.*, vol. 22, pp. 324–333, 2 1998. DOI: https://doi.org/10.1097/00004728-199803000-00032.

[20] R. W. Cox, "Afni: Software for analysis and visualization of functional magnetic resonance neuroimages," *Computers and Biomedical Research*, vol. 29, no. 3, pp. 162–173, 1996. DOI: https://doi.org/10.1006/cbmr.1996.0014.

[21] P. Coupé, J. V. Manjón, E. Gedamu, D. Arnold, M. Robles, and D. L. Collins, "Robust rician noise estimation for mr images," *Med Image Anal.*, vol. 14, pp. 483–493, 4 2010. DOI: https://doi.org/10.1016/j.media.2010.03.001.

[22] R. Beare, J. Chen, C. Adamson, T. Silk, D. Thompson, J. Yang, V. Anderson, M. Seal, and A. Wood, "Brain extraction using the watershed transform from markers.," *Frontiers in Neuroinformatics*, vol. 7, no. 32, Dec. 2013. DOI: doi:10.3389/fninf.2013.00032.

[23] E. Amico and J. Goñi, "The quest for identifiability in human functional connectomes," *Scientific Reports*, vol. 8, no. 1, 2018. DOI: https://doi.org/10.1038/s41598-018-25089-1.

[24] IBM Corp., *Ibm spss statistics for windows*, version 25.0, Armonk, NY: IBM Corp, 2017. [Online]. Available: https://hadoop.apache.org.

[25] E. M. McCormick and E. H. Telzer, "Contributions of default mode network stability and deactivation to adolescent task engagement," *Scientific Reports*, vol. 8, 1 2018. DOI: https://doi.org/10.1038/s41598-018-36269-4.