The Statistical Analysis of Resting-State fMRI Data

A Comparison of Averaging Methods

Christian Klein

Bachelor thesis Psychology
Institute of Psychology
Faculty of Social and Behavioural Sciences – Leiden University
Unit: Department of Statistics
Date: 11/07/2017
Bachelorproject number: 79
Student number: 1499076
First examiner: Dr. W. D. Weeda, Dr. T. F. Wilderjans

Abstract

Functional magnetic resonance imaging (fMRI) studies examining differences in functional connectivity between the brain’s functional (resting-state) networks have become pivotal for investigating the dynamical functionality of the brain across various mental states and neuropsychiatric patient populations. These resting-state group comparison studies rely on correlations between blood-oxygen-level-dependent (BOLD) signals that have been obtained by summarising the fMRI time-series from every voxel within a network. Regarding this analysis, an important question concerns the method by which BOLD signals are averaged across all the voxels in a resting-state network. To our knowledge, however, no simulation studies addressing this problem have been conducted. The current study, therefore, compares the effects of four averaging methods (mean, median, PCA and CCA) on the ability of a simulated resting-state group comparison analysis to detect a significant difference in inter-network functional connectivity between two groups. Variables representing the group sizes, the size of the networks, the signal-to-noise ratio and the true between-groups inter-network functional connectivity difference were taken as between-subjects independent factors. Results from statistical analysis of the output from the simulation demonstrated better performances for the mean and median approaches compared to CCA whilst the PCA strategy invariably failed to aid detection of the inter-network functional connectivity difference between the groups. Furthermore, although constrained to small region sizes, a beneficial effect of increasing signal-to-noise ratio was found. Researchers investigating group differences in resting-state functional connectivity are therefore recommended to implement the mean and median approaches when summarising fMRI time-series across networks and to minimize or account for the amount of noise in their studies. Future research should seek to replicate the current findings and determine the superior method between the mean and median approaches.

Introduction

Functional magnetic resonance imaging (fMRI) exploits fluctuations in blood-oxygen levels associated with the electrochemical activities of large bodies of neurons to infer the amount of activation relative to a baseline in various neural structures (Damoiseaux et al, 2006; Fox & Raichle, 2007; Lindquist, 2008; Van Dijk, Hedden, Venkataraman, Evans, Lazar, & Buckner, 2010). That is, the so-called blood oxygenation level-dependent (BOLD) contrast is utilised as a proxy to investigate neuronal activities throughout the cerebrum. Typically, fMRI studies have sought to determine the extent of neuronal activation within regions-of-interest (ROIs) in response to varying demands of function (functional specialisation) via analysis of the BOLD contrast in response to a stimulus presented within the scanner or a task that a participant must perform (Damoiseaux et al, 2006; Fox & Raichle, 2007; Friston, 1994; Friston, 2009). However, brain areas do not process information in isolation. Instead, they communicate with each other and share the output of their own specialisations (functional integration) in a complex network composed of several sub-networks, each varying in their contributions to cognitive processes and phenomenal experience (Damoiseaux et al, 2006; Fair et al, 2007; Fox et al, 2005; Laird et al, 2011; Sporns, 2011; Thompon & Fransson, 2015; Van den Heuvel et al, 2010; Van Dijk et al, 2010; Yu, Sui, Kiehl, Pearlson, & Calhoun, 2013). Therefore, many contemporary researchers prefer to use fMRI to examine the co-activation in haemodynamic activity (functional connectivity) between ROIs and networks at rest rather than straightforward task-based activation (Cohen et al, 2008; Damoiseaux et al, 2006, Fox & Raichle, 2007, Fox et al, 2005; Van den Heuvel et al, 2010). Moreover, apart from representing an invaluable source for understanding the functional architecture of the brain, resting-state research is ideal for investigating neurodegenerative diseases and other neuropsychiatric conditions that are considered to involve the disorganization of the brain’s networks and their functional connections (Alonso-Solís et al, 2015; Calhoun, Eichele, & Pearlson, 2009; Church et al, 2009, Power et al, 2011; Seeley, Crawford, Zhou, Miller, & Greicius, 2009, Van den Heuvel et al, 2010; Yu, Sui, Kiehl, Pearlson, & Calhoun, 2013).

In resting-state studies, subjects are placed in the fMRI scanner and instructed to simply relax, eyes either open or closed, whilst spontaneous, low-frequency (0.01-0.1Hz) fluctuations in their BOLD activity are recorded over time (Biswal, Yetkin, Haughton, & Hyde, 1995; Cordes et al, 2001; Damoiseaux et al, 2006; Fox & Raichle, 2007; Power et al, 2011; Thompson & Fransson, 2015; Van Dijk et al, 2010). These resting-state time-series are then correlated between a previously defined ROI, often called a seed region, and every other voxel (volumetric pixel) in the brain (Biswal et al, 1995; Cohen et al, 2008; Damoiseaux et al, 2006; Kang et al, 2011; Lindquist, 2008; Poldrack, Mumford, & Nichols, 2011; Van Den Heuvel & Hulshoff, 2010). The output from this analysis represents a functional connectivity map for the seed region from which the brain's intrinsic functional networks, in which the seed region serves as a network node, may then be derived (Cohen et al, 2008; Damoiseaux et al, 2006; Kang et al, 2011; Lindquist, 2008; Poldrack et al, 2011; Van Den Heuvel & Hulshoff, 2010). Once these resting-state networks have been established, estimates of functional connectivity between them may be obtained by averaging the fMRI time series from every voxel within a network and correlating the resulting averaged signal with one from another network (Alonso-Solís et al, 2015; Carhart Harris et al, 2013). The strength of this inter-network functional connectivity may then be compared between subject groups, affording examination of how the brain's functional network dynamics differ across conscious states, clinical populations and developmental milestones (Alonso-Solís et al, 2015; Calhoun et al, 2009; Carhart-Harris et al, 2013; Church et al, 2009; Cohen et al, 2008; Çetin et al, 2014; Damoiseaux et al, 2006; Fukunaga et al, 2006; Nir, Hasson, Levy, Yesherun & Malach, 2006; Seeley et al, 2009; Van den Heuvel et al, 2010; Yu et al, 2013).

The procedure described above is known as the seed-voxel-based correlational method for assessing functional connectivity since the fMRI time series representing the seed region is correlated with the time series from every other voxel in the brain (Biswal et al, 1995; Cohen et al, 2008; Damoiseaux et al, 2006; Kang et al, 2011; Lindquist, 2008; Poldrack et al, 2011; Van Den Heuvel & Hulshoff, 2010). In order to obtain a summarised BOLD signal for a ROI, the time series from every voxel within that region must be averaged in a manner which best captures the mutual, coherent activity between every voxel in that region. Furthermore, for clinical, developmental and consciousness studies which aim to compare differences in inter-network functional connectivity between groups, summarised BOLD signals reflecting the shared activity of entire resting-state networks must be obtained. Therefore, a key question concerns the method by which the signals from individual voxels in a ROI or a resting-state network are averaged to produce an overall signal for that region or network (Poldrack et al, 2011).

Currently, four approaches to acquiring a global, temporal signal across multiple voxels are considered viable. The first approach is simply the computation of the mean signal across all voxels in a ROI or network. A problem with this method, however, is that the resulting averaged signal shall reflect an amalgam of multiple signal sources and not necessarily the latent, mutual signal shared by the voxels (Poldrack et al, 2011). Another simple method to determine a global signal for a ROI is to take the median of all the voxels from the region under investigation. Although, this approach runs into the same predicament as that of taking the mean signal. That is, the median signal of all the voxels may not necessarily represent the underlying pattern of activation.

The remaining approaches involve more complex computation but perhaps yield a more accurate representation of the underlying signal across multiple voxels over time. Among these is the canonical correlation analysis (CCA) which takes linear combinations of two or more sets of variables and applies the weights in such a manner that the correlation between the sets is maximised (Lazar, 2008). CCA, therefore, exposes the underlying correlation structure between the variable sets and may yield a more precise approximation of the overall signal within a ROI or network when applied to fMRI (Lazar, 2008). A final approach utilises the data reduction technique known as principal components analysis (PCA). PCA aims to identify the latent composition of the variables in question by generating linear combinations of those variables, known as components, in such a way that each component explains more variance in the original variable set than its successor (the first component, therefore, always explains the most variance) and is uncorrelated to all other components (Lazar, 2008; Poldrack et al, 2011). In the context of averaging the time-series from every voxel in a ROI or network, PCA, with its ability to reveal underlying dimensions behind a set of variables, could yield a more valid estimation of the mutual signal shared by multiple voxels.

So far there has been an absence of simulation studies investigating the effects of these averaging methods on the estimation of functional connectivity between ROIs and resting-state brain networks (Poldrack et al, 2011). This lack of investigation, together with the rapidly growing interest regarding how resting-state brain networks vary in functional connectivity between one another across various conscious states and neuropsychiatric patient populations, presents a fitting opportunity to compare these averaging methods in the context of a simulated comparison between two groups whom differ on their levels of inter-network functional connectivity between two networks. Thus, the current study aims to compare the efficacy of these averaging approaches to detect a significant difference in inter-network functional connectivity between two groups under varying levels of typical resting-state fMRI and BOLD-related variables, namely, the size of the groups, the size of the networks, the signal-to-noise ratio and the true correlation of spontaneous BOLD activity between the neural networks. These variables were chosen since they represent factors likely to affect the estimation of functional connectivity and therefore the ability to detect an inter-network functional connectivity difference between sets of subjects in resting-state studies that aim to compare two or more groups.

The present study maintains an exploratory approach with no a priori hypotheses being made regarding which averaging methods will prove superior or how the independent variables will interact with each other. However, logical expectations concerning the direction of some effects can be considered. Firstly, as signal-to-noise ratio increases, it follows that the difference between the groups should be easier to detect. Furthermore, detection should be easier at a large true difference in inter-network functional connectivity between the groups. Lastly, with larger group sizes, detection should be facilitated due to beneficial effects on the statistical power of the analysis.

The sections that follow describe the methodology of the current study including how the simulation was programmed (conceptually) and the statistical analysis of the results from the simulation. Afterwards, results of the statistical analysis shall be reported before a discussion considering the interpretation and implications of the findings for resting-state studies. Limitations of the study and directions for future research will, thereafter, be considered.

Method

A 3 x 3 x 3 x 3 x 4 design varying the levels of the between-groups inter-network functional connectivity difference, group sizes, size of the networks, signal-to-noise ratio and the averaging method was simulated using RStudio 3.3.2. Firstly, fMRI data with the structure of two functionally connected resting-state brain networks was generated for two groups. Single regions were programmed in place of networks, however, as this allowed for a more straightforward and parsimonious code whilst holding little consequence for the results or validity of the study. Thus, the connectivity weights between the regions served as a proxy for inter-network functional connectivity. The software was instructed to generate resting-state data of length 360000 for each region in such a manner that the correlations between the regions, and thus the difference in inter-network functional connectivity between each group, varied depending on the level of the factor representing the difference between the groups. The difference in functional connectivity between the two groups varied from low (.75 - .5 = .25) to moderate (.75 - .25 = .5) to high (.75 – 0 = .75). These levels were chosen because they reflect a large range of differences likely to be present in resting-state studies that compare two or more groups. With the use of a custom ‘resting-state data generation’ function it was possible to program random gaussian noise inherent to the data. This is important for resting-state data, since noise is essentially a constituent of the signal rather than a separate component that needs to be added. For every level of the signal-to-noise ratio variable, the signal-to-noise ratio in the voxels varied between low (1), medium (5) and high (10). Again, to represent the broad range of noise levels possible in resting-state group comparison studies. The same function was used to incorporate region size and the size of the groups into the data generation process as varying factors. The size of the regions varied between 50, 100 or 200 voxels as these sizes represent realistic anatomical region sizes whilst simultaneously placing little strain on the computational demands of the simulation. The number of subjects in each group cycled between 10, 30 and 60. These levels reflect the full range of sample sizes likely to be present in resting-state studies comparing multiple groups.

Before correlational analysis, the signals in each region were averaged with one of the four methods (mean, median, PCA, CCA) dependent on the level of the method factor. The regression between the averaged signals from each region was then computed for each group. Correlations outputted by this regression represented estimates of inter-regional functional connectivity between the regions for each group. In order to detect a difference in inter-regional functional connectivity between the two groups, t-tests were performed on the functional connectivity estimates for each group. The significance values yielded from these t-tests, under every condition of the independent variables, were taken as the outcome measure for statistical analysis. Initially, 50 simulations for each level of the independent variable representing the difference in functional connectivity between the groups were performed by the R software. Unfortunately, due to ambiguous issues with the stability of the code, far fewer simulations with useable output were available for analysis. Under ordinary circumstances, this would represent a major concern since the number of simulations normally takes the role of sample size in simulation studies. However, due to the aforementioned custom function, the sample sizes in the present study were already intrinsic to the resting-state data, thus lowering the importance of maintaining a large number of simulations.

Statistical Analysis

In order to examine whether the independent variables influenced the simulation's ability to detect a significant difference in inter-network functional connectivity between the two groups, a 3 x 3 x 3 x 3 factorial repeated measures ANOVA was performed on the t-test p-values generated by the simulation study. The size of the inter-network connectivity difference, the size of the groups, the number of voxels composing regions and the signal-to-noise ratio were set as between-subjects factors whilst the averaging method used to obtain the regional signals represented the within-subjects factor. Estimated marginal means and interaction plots were examined to determine the direction of significant effects. An α of .05 was set as the threshold for significant values. All statistical analysis was conducted using IBM's SPSS. 1848 cases were analysed in total.

Results

Table 1 shows the ANOVA table for the within-subjects factor 'method' as well as the between-subjects factors 'group difference', 'group sizes', 'region sizes' and 'signal-to-noise-ratio'. Due to the substantial number of cases analysed, the repeated measures assumption of normality was rendered irrelevant. Furthermore, Mauchly's test of sphericity, the repeated measures equivalent of a homogeneity of variance/covariance test, failed to return test statistics. This was most likely due to the low amount of variance observed across the dependent values. However, the epsilon values were small, indicating that the assumption of sphericity had been violated. This was not surprising since the probability of violating sphericity increases with each additional repeated measure or, in our case, with each signal averaging method used. On account of examining 4 separate averaging methods in the current study (mean, median, PCA, CCA), a severe violation of sphericity was highly probable and adjustment would have to made to reduce the probability of a type I error. Thus, a Greenhouse-Geisser correction was performed to aid in interpreting the results more conservatively. The epsilon value for the Greenhouse-Geisser correction shared the same value as that of the lower bound adjustment, normally the most conservative method of correction. Thus, the results were analysed in the most conservative manner possible.

Table 1. *ANOVA table of between and within-subjects effects on t-test significance values.*

Effect	SS	df	MS	F-value	p-value	Partial Eta Squared
INCdiff	.014	2	.007	.332	.718	.000
Nvox	.062	2	.031	1.493	.225	.002
Nsub	.071	2	.036	1.705	.182	.002
SNR	.047	2	.024	1.127	.324	.001
INCdiff * Nvox	.034	4	.009	.408	.803	.001
INCdiff * Nsub	.085	4	.021	1.015	.398	.002
INCdiff * SNR	.043	4	.011	.515	.724	.001
Nvox * Nsub	.074	4	.019	.888	.470	.002
Nvox * SNR	.397	4	.099	4.760	.001*	.011
Nsub * SNR	.063	4	.016	.760	.551	.002
INCdiff * Nvox * Nsub	.086	8	.011	.518	.844	.002
INCdiff * Nvox * SNR	.279	8	.035	1.668	.101	.007
INCdiff * Nsub SNR	.100	8	.013	.600	.779	.003
Nvox * Nsub * SNR	.220	8	.027	1.317	.230	.006
INCdiff * Nvox * Nsub * SNR	.627	16	.039	1.879	.018*	.017
Error	36.872	1767	.021			
Method	344.513	1	344.513	5503.345	.000*	.757

Method * INCdiff	.042	2	.021	.332	.718	.000
Method * Nvox	.187	2	.093	1.493	.225	.002
Method * Nsub	.213	2	.107	1.705	.182	.002
Method * SNR	.141	2	.071	1.127	.324	.001
Method * INCdiff * Nvox	.102	4	.026	.408	.803	.001
Method * INCdiff * Nsub	.254	4	.064	1.015	.398	.002
Method * INCdiff * SNR	.129	4	.032	.515	.724	.001
Method * Nvox * Nsub	.222	4	.056	.888	.470	.002
Method * Nvox * SNR	1.192	4	.298	4.760	.001*	.011
Method * Nsub * SNR	1.90	4	.048	.760	.551	.002
Method * INCdiff * Nvox * Nsub	.259	8	.032	.518	.844	.002
Method * INCdiff * Nvox * SNR	.836	8	.104	1.668	.101	.007
Method * INCdiff * Nsub * SNR	.301	8	.038	.600	.779	.003

Method * Nvox * Nsub * SNR	.659	8	.082	1.317	.230	.006
Method * INCdiff * Nvox * Nsub * SNR	1.882	16	.118	1.879	.018*	.017
Error(Method)	110.615	1767	.063			

Note. p-values marked with ‘*’ represent significant values. ‘INCDiff’, ‘Nsub’ and ‘SNR’ correspond to ‘functional connectivity difference’, ‘group sizes’ and ‘signal-to-noise ratio’, respectively.

Among the significant effects displayed in table 1, by far the most powerful is the main effect of the sole within-subjects factor ‘Method’ (*F* (1, 1767) = 5503.345, *p* < .001). This effect returned a large effect size (eta squared = .757), demonstrating that the averaging approach used to obtain an overall regional signal from the voxels did indeed influence the simulation’s efficacy to detect a significant difference in functional connectivity between the two groups. Observation of the estimated marginal means for each averaging method, presented in Table 2 below, reveal a clear inability of the PCA approach to detect a significant difference between the groups (larger means correspond to deficiency in detecting a difference). Furthermore, both the mean and the median approaches to voxel averaging appear superior to the CCA approach. A post-hoc comparison of the estimated marginal means using a Bonferroni correction to account for multiple testing confirmed a significant difference between the PCA approach and all other approaches (*p* < 0.001 for all comparisons). The comparison between the mean and median approaches with CCA, however, did not reach significance (*p* = .139).

Table 2. *Estimated marginal means for the within-subjects factor ‘Method’.*

Method	Estimated Marginal Mean (SE)
Mean	1.375xE-013 (.000)
Median	2.044xE-013 (.000)
PCA	.505 (.007)
CCA	1.921xE-008 (.000)

Prior to the analysis, beneficial effects on detection of signal-to-noise ratio, true functional connectivity difference and groups sizes were expected to occur. However, none of the main effects from these independent factors reached significance by themselves. Instead, effects of between-subjects variables depended on levels of each other. Among these interaction effects, the most pronounced significant F-ratios stemmed from the 4-way interaction between all the between-subjects independent variables (*F* (16, 1767) = 1.879, *p* < .05) and the 5-way interaction between all of the independent factors (*F* (16, 1767) = 1.879, *p* < .05) (for both effects, eta squared = .017). Due to both PCA's poor performance and a severe lack of variance in the data, however, interaction plots in SPSS were rendered uninterpretable for these more complex interactions. Therefore, these effects were excluded from the interpretation phase.

Among the less powerful significant interaction effects was the interaction between region size and signal-to-noise ratio (*F* (4, 1767) = 4.760, *p* < .01) as well as the 3-way interaction between these same variables with the averaging method used (*F* (4, 1767) = 4.760, *p* < .01) (for both effects, eta squared = .011). Inspection of the estimated marginal means (see Appendix A) and interaction plot (see Figure 1) for the interaction between region size and signal-to-noise demonstrates the presence of a beneficial effect of increasing signal-to-noise ratio on detection, dependent on the size of the region. Specifically, it appears that at region sizes larger than 50 voxels, the positive effect of signal-to-noise ratio dissipates. The interaction becomes more interesting, however, when the averaging method is considered. Observation of th estimated marginal means (see Appendix B) and interaction plots (see Appendix C) for the interaction between the averaging method with the size of the region and signal-to-noise ratio reveals that within a small region size, the ability of the mean and median approaches to detect the difference increases as the amount of signal to noise increases. CCA's ability to detect the difference, on the other hand, is more apt to improve contingently with the signal-to-noise ratio within larger region sizes. The detection ability of PCA was unchanged under all these conditions, failing to detect the difference on every occasion.

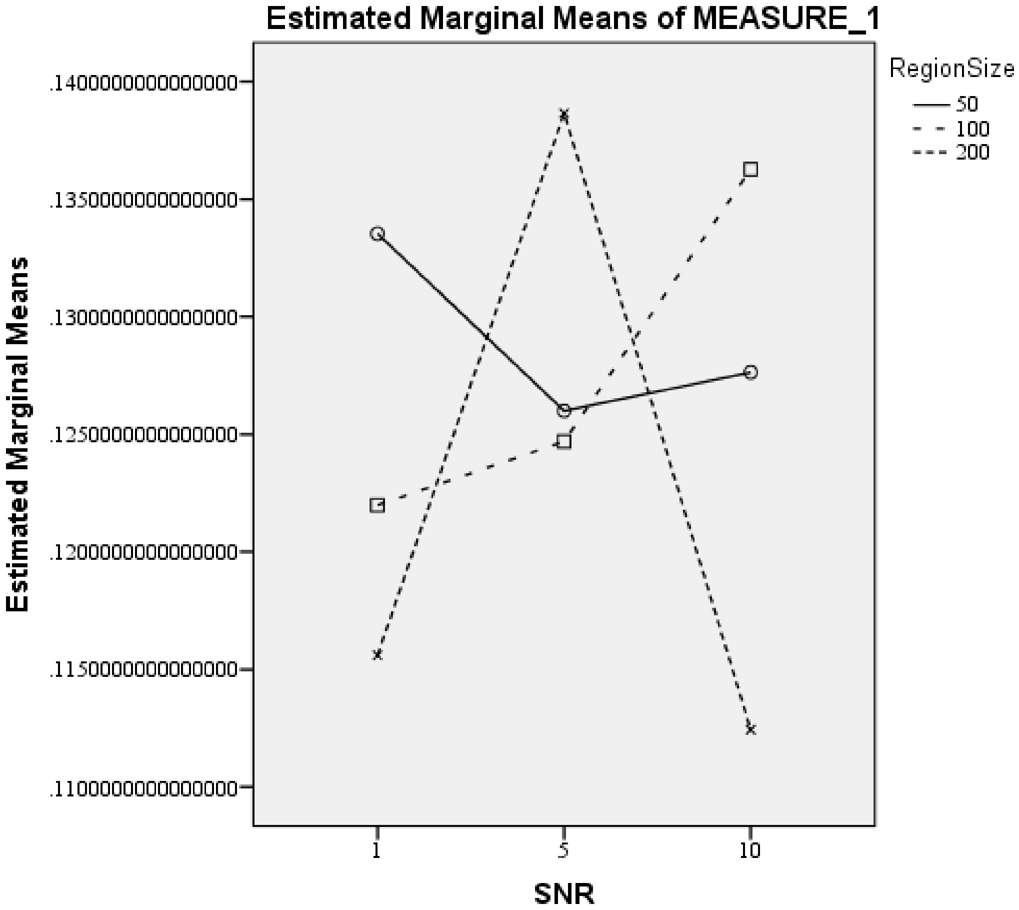

Figure 1. Interaction plot for signal-to-noise ratio and region size.

Discussion

The current study set out to compare and evaluate, under varying levels of typical fMRI conditions, the effectiveness of four recognised statistical approaches (mean, median, PCA and CCA) to averaging resting-state BOLD signals across brain regions. In the context of a simulated comparison between two groups, the effects of each averaging method on the simulation's ability to detect the difference in inter-network functional connectivity between the groups was investigated. Results from statistical analysis of the simulation's output demonstrated solid performances of the mean, median and CCA approaches whilst the PCA method invariably failed to detect the difference. Furthermore, the expected beneficial effect of increasing the signal-to-noise ratio was more conspicuous for the mean and median approaches within small region sizes whereas with CCA, the effect was greatest at large regions. The overall effect of signal-to-noise ratio, however, was only relevant within small regions.

In light of these findings, various implications can be made for future research endeavouring to unravel the relations between inter-network functional connectivity, phenomenal experience, neurocognitive functionality and neuropsychiatric disorders. Firstly, resting-state studies should always refrain from using PCA to obtain an overall underlying signal for a region. Instead, either the mean, median or CCA should be used. Furthermore, the mean and median approaches are slightly preferable to CCA, producing smaller significance values even within larger region sizes where the positive effect of signal-to-noise ratio is most pronounced for CCA. Implications should also be noted for the importance of reducing the amount of noise present in the data of resting-state studies. Although the beneficial effect of increasing signal-to-noise ratio was largely constrained to small region sizes, it remains a relevant factor which researchers are able to influence to maximise the likelihood of detecting a difference under certain conditions. Noise contributions to the BOLD signal most often arise from physiological sources such as cardiac, respiratory and autonomic processes (Fukunaga et al., 2006; Liu, 2016; Triantafyllou, Polimeni, & Wald, 2011). Moreover, thermal noise resulting from scanner instabilities and subject motion further account for a portion of the total noise (Fukunaga et al., 2006; Liu, 2016; Triantafyllou et al., 2011). Various techniques may be utilised by researchers to minimise, or at least account for, contributions from all these sources of noise. For instance, thermal noise can be attenuated by performing studies at high magnetic field strengths and applying multi-channel brain coil detector arrays with multi-channel receivers (Bodurka et al., 2003; Fukunaga et al., 2006; Triantafyllou et al., 2011; de Zwart et al., 2004). In fact, depending on the exact neuroanatomical location, this arrangement allows for between two to six-fold increases in signal-to-noise ratio compared with a standard setup (Fukunaga et al., 2006; de Zwart, van Gelderen, Kellman, & Duyn, 2002; de Zwart et al., 2004). Furthermore, noise induced by head motion can be dealt with by including inflatable cuffs between the MRI detector array and the subjects' heads (Fukunaga et al., 2004). Lastly, signal fluctuations arising from cardiac and respiratory cycles, although immune to direct influence, may be assessed and accounted for by slowing down the repetition time between scans for a certain number of participants (Fukunaga et al., 2004).

The current study, as with any investigation, does suffer from some key limitations. Firstly, being a simulation, the study suffers from a lack of ecological validity. That is, due to the artificial nature of the simulated environment, results may not be fully generalisable to real-life settings in which the data tends to be much 'dirtier' than that in simulation studies (Hallgren, 2013). Furthermore, although reflecting realistic sizes of some brain regions, the highest level of the variable representing the size of the regions was set to only two hundred voxels to comply with both computational and temporal constraints. Regions investigated in resting-state studies, however, may contain many more voxels. Moreover, it is possible that the remaining independent variables behave differently within regions larger than two hundred voxels. Due to the small scale of the ‘region size’ factor, however, these hypothetical patterns were never afforded the opportunity to be detected. Finally, since t-test significance values were used to assess the effects of the four averaging methods on the simulation’s ability to detect a difference in inter-network functional connectivity between two groups, the current study provides no information regarding the effects of the averaging methods on the accuracy of functional connectivity estimation. Researchers seeking to determine the best method for precise estimation of functional connectivity would be advised to compare known connectivity values with the functional connectivity estimates for each method.

Future resting-state fMRI simulation studies should strive to replicate the current findings with a larger scale of region size settings. That is, researchers with more computational power and time at their disposal should compare the effects of the averaging methods and their patterns of interaction with the other independent variables within region sizes consisting of over a thousand voxels. Furthermore, focus should be given to determining the superior between the mean and median approaches. This could be achieved by comparing them at extreme, adverse fMRI conditions to observe which method proves more robust.

To conclude, the current simulation study compared the effects of four averaging approaches (mean, median, PCA, CCA), under varying levels of typical resting-state fMRI and BOLD-related variables, on the detection of a difference in inter-network functional connectivity between two groups. Statistical analysis of the output from this simulated comparison revealed a clear inability of PCA to promote detection of the functional connectivity difference across every condition. Furthermore, the mean and median approaches aided detection marginally more than CCA, especially within small region sizes where the beneficial effect of increasing signal-to-noise ratio is most pronounced for these methods and is most relevant by itself. Resting-state researchers aiming to compare group differences in inter-network functional connectivity are therefore recommended to use the mean and median approaches whilst employing efforts to maintain high levels of signal-to-noise ratio.

References

Alonso-Solís, A., Vives-Gilabert, Y., Grasa, E., Portella, M, J., Rabella, M., Sauras, R, B., . . . Corripio, I. (2015). Resting-state functional connectivity alterations in the default network of schizophrenia patients with persistent auditory verbal hallucinations. *Schizophrenia Research, 161*(2-3), 261-268.

Biswal, B., Yetkin, F. Z., Haughton, V. M., & Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magnetic Resonance in Medicine, 34*(4), 537-541.

Bodurka, J., Ledden, P. J., van Gelderen, P., Chu, R., de Zwart, J. A., Morris, D., & Duyn, J. H. (2003). Scalable multichannel MRI data acquisition system. *Magnetic Resonance In Medicine, 51(1)*, 165-171.

Calhoun, V. D., Eichele, T. & Pearlson, G. (2009). Functional Brain Networks in Schizophrenia: A Review. *Frontiers in Human Neuroscience, 3*, 17.

Carhart-Harris, R. L., Leech, R., Erritzoe, D., Williams, T. M., Stone, J. M., Evans, J., . . . Nutt, D. J. (2013). Functional Connectivity Measures After Psilocybin Inform a Novel Hypothesis of Early Psychosis. *Schizophrenia Bulletin, 39*(6), 1343-1351.

Çetin, M. S., Christensen, F., Abbott, C. C., Stephen, J. M., Mayer, A. R., Cañive, J. M, . . . Calhoun, V. D. (2014). Thalamus and posterior temporal lobe show greater inter-network connectivity at rest and across sensory paradigms in schizophrenia. *NeuroImage, 97*, 117-126.

Church, J. A., Fair, D. A., Dosenbach, N. U. F., Cohen, A. L., Miezin, F. M., Petersen, S. E., & Schlagger, B. L. (2009). Control networks paediatric Tourette syndrome show immature and anomalous patterns of functional connectivity. *Brain, 132*(1), 225-238.

Cohen, A. L., Fair, D. A., Dosenbach, N. U. F., Miezin, F. M., Dierker, D., Van Essen, D. C., . . . Petersen, S. E. (2008). Defining functional areas in individual human brains using resting functional connectivity MRI. *NeuroImage, 41*(1), 45-57.

Cordes, D., Haughton, V. M., Arfanakis, K., Carew, J. D., Turski,, P. A., Moritz, . . . Meyerand, M. E. (2001). Frequencies contributing to functional connectivity in the cerebral cortex in “resting state” data. *American Journal of Neuroradiology, 22*, 1326–1333.

Damoiseaux, J. S., Rombouts, S. A. R. B., Barkhof, F., Scheltens, P., Stam, C. J., Smith, S. M., & Beckmann, C. F. (2006). Consistent resting-state networks across healthy subjects. *Proceedings of the National Academy of Sciences of America, 103*, 13848-13853.

Fair, D. A., Dosenbach, N. U. S., Church, J. A., Cohen, A. L, Brahmbhatt, S., Miezin, F. M., . . . Schlaggar, B. L. (2007). Development of distinct control networks through segregation and integration. *Proceedings of the National Academy of Sciences of the United States of America, 104*(33), 13507-13512.

Fox, M. D., & Raichle, M. E., (2007). Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nature. Reviews. Neuroscience, 8*, 700-711.

Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Essen, D. C., & Raichle, M. E. (2005). From The Cover: The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences, 102*(27), 9673-9678.

Friston, K. J. (1994). Functional and Effective Connectivity in Neuroimaging: A Synthesis. *Human Brain Mapping, 2*, 56-78.

Friston, K. J. (2009). Modalities, modes, and models in functional neuroimaging. *Science, 326*, 399-403.

Fukunaga, M., Horivitz, S. G., van Gelderen, P., de Zwart, J. A., Jansma, J. M., Ikonomidou, V. N., . . . Duyn, J. H. (2006). Large-amplitude, spatially correlated fluctuations in BOLD fMRI signals during extended rest and early sleep stages. *Magnetic Resonance Imaging, 24*, 979-992.

Hallgren, K. A. (2013) Conducting Simulation Studies in the R Programming Environment. *Tutorials in Quantitative Methods for Psychology, 9*(2), 43-60.

Kang, J., Wang, L., Yan, C., Wang, J., Liang, X., & He, Y. (2011). Characterizing dynamic functional connectivity in the resting brain using variable parameter regression and Kalman filtering approaches. *NeuroImage, 56*(3), 1222-1234.

Lazar, N. A. (2008). *The Statistical Analysis of Functional MRI Data*. Springer.

Laird, A. R., Mickle Fox, P., Eickhoff, S. B., Turner, J. A., Ray, K. L., Reese McKay, D., . . . Fox, P. T. (2011). Behavioral Interpretations of Intrinsic Connectivity Networks. *Journal of Cognitive Neuroscience, 23*(12), 4022-4037.

Lindquist, M. A. (2008). The Statistical Analysis of fMRI Data. *Statistical Science, 23*, 439-464.

Liu, T. T. (2016). Noise contributions to the fMRI signal: An overview. *NeuroImage, 143*, 141-151.

Nir, Y., Hasson, U., Levy, Y., & Malach, Y. R., (2006) Widespread functional connectivity and fMRI fluctuations in human visual cortex in the absence of visual stimulation. *Neuroimage, 30*(4), 1313–1324.

Poldrack, R. A., Mumford, J. A., & Nichols, T. E. (2011). *Handbook of Functional MRI Data Analysis*. Cambridge: Cambridge University Press.

Power, J. D., Cohen, A. L., Nelson, S. M., Wig, G. S., Barnes, K. A., Church, J. A., . . . Schlagger, B. L. (2011). Functional Network Organization of the Human Brain. *Neuron, 72*, 665-678.

Seeley, W. W., Crawford, R. K., Zhou, J. Miller, B. L., & Greicius, M. D. (2009). *Neuron, 62*, 42-52.

Sporns, O. (2011). The human connectome: a complex network. *Annals of The New York Academy Of Sciences, 1224*, 109-125.

Thompson, W. H., & Fransson, P. (2015). The frequency dimension of fMRI dynamic connectivity: Network connectivity, functional hubs and integration in the resting brain. *NeuroImage, 121*, 227-242.

Triantafyllou, C., Polimeni, J. R., & Wald, L. L. (2011). Physiological noise and signal-to-noise ratio in fMRI with multi-channel array coils. *NeuroImage, 55*, 597-606.

Van den Heuvel, M. P., & Hulshoff Poll, H. E. (2010). Exploring the brain network: A review on resting-state fMRI functional connectivity. *European Neuropsychopharmacology, 20*(8), 519-534.

Van Dijk, K. R. A., Hedden, T., Venkataraman, A., Evans, K. C., Lazar, S. W., & Buckner, R. L. (2010) Intrinsic Functional Connectivity As a Tool For Human Connectomics: Theory, Properties, and Optimization. *Journal of Neurophysiology, 103(1),* 297-321.

Yu, Q., Sui, J., Kiehl, K. A., Pearlson, G., & Calhoun, V. D. (2013). State-related functional integration and functional segregation brain networks in schizophrenia. *Schizophrenia Research, 150*,450-458.

de Zwart, J. A., van Gelderen, P., Kellman, P., & Duyn, J. H. (2002). Application of Sensitivity-Encoded Echo-Planar Imaging for Blood Oxygen Level-Dependent Functional Brain Imaging. *Magnetic Resonance In Medicine, 48*, 1011-1020.

de Zwart, J. A., Ledden, P. J., van Gelderen, P., Bodurka, J., Chu, R., & Duyn, J. H. (2004). Signal-to-Noise Ratio and Parallel Imaging Performance of a 16 Channel Receive-Only Brain Coil Array at 3.0 Tesla. *Magnetic Resonance In Medicine, 51*, 22-26.

Appendix A

Table 3. *Estimated marginal means for the interaction between region size and signal-to-noise ratio*

Region Size	SNR	Estimated Marginal Mean (SE)
50	1	.134 (.005)
	5	.126 (.005)
	10	.128 (.005)
100	1	.122 (.005)
	5	.125 (.005)
	10	.136 (.005)
200	1	.116 (.005)
	5	.139 (.005)
	10	.112 (.005)

Appendix B

Table 4. *Estimated marginal means for the 3-way interaction between region size, signal-to-noise ratio and averaging method.*

Region Size	SNR	Method	Estimated Marginal Mean (SE)
50	1	Mean	3.651xE-013 (.000)
		Median	9.293xE-013 (.000)
		PCA	.534 (.020)
		CCA	5.300xE-012 (.000)
	5	Mean	1.084xE-013 (.000)
		Median	1.212xE-013 (.000)
		PCA	.504 (.020)
		CCA	1.049xE-011 (.000)
	10	Mean	5.157xE-015 (.000)
		Median	4.278xE-015 (.000)
		PCA	.511 (.020)
		CCA	1.308xE-013 (.000)
100	1	Mean	1.629xE-014 (.000)
		Median	1.830xE-014 (.000)
		PCA	.488 (.020)
		CCA	1.024xE-008 (.000)
	5	Mean	2.234xE-013 (.000)
		Median	1.946xE-013 (.000)
		PCA	.499 (.020)
		CCA	2.246xE-013 (.000)
	10	Mean	6.430xE-014 (.000)
		Median	6.479xE-014 (.000)
		PCA	.545 (.020)
		CCA	7.339xE-013 (.000)
200	1	Mean	2.964xE-013 (.000)
		Median	3.539xE-013 (.000)
		PCA	.462 (.020)
		CCA	1.600xE-007 (.000)

5	Mean	7.761xE-014 (.000)
	Median	8.141xE-014 (.000)
	PCA	.555 (.020)
	CCA	2.576xE-009 (.000)
10	Mean	8.067xE-014 (.000)
	Median	7.139xE-014 (.000)
	PCA	.450 (.020)
	CCA	4.664xE-011 (.000)

Appendix C

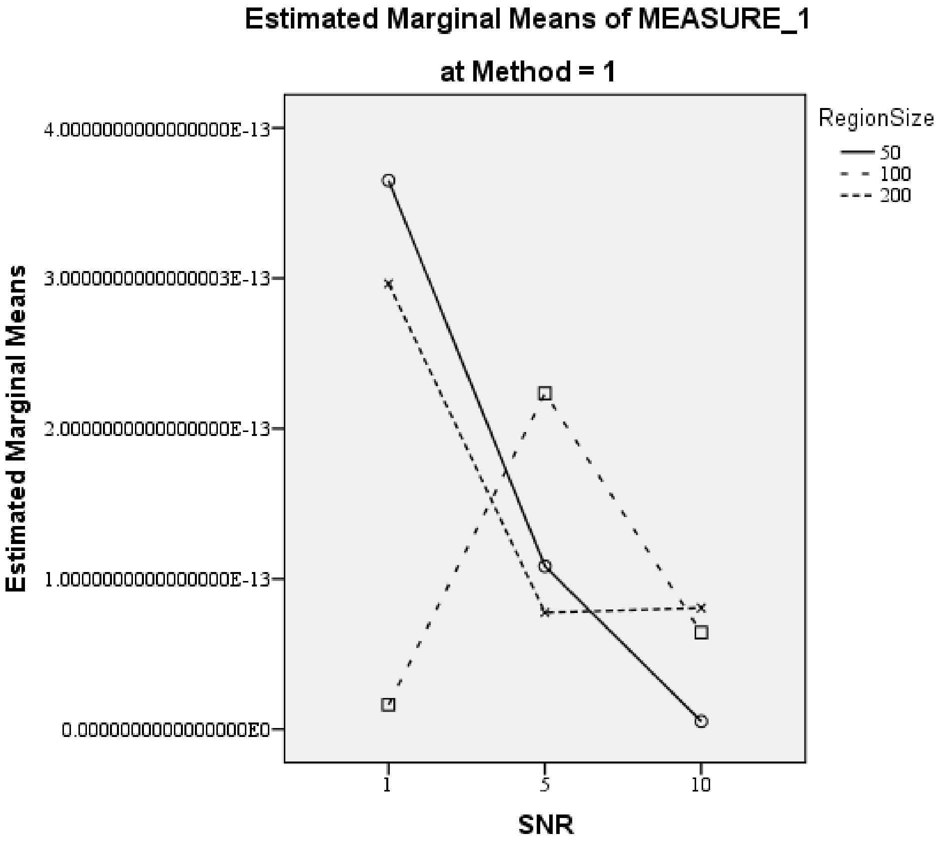

Figure 2a: Interaction plot for region size and signal-to-noise ratio using the 'mean' method.

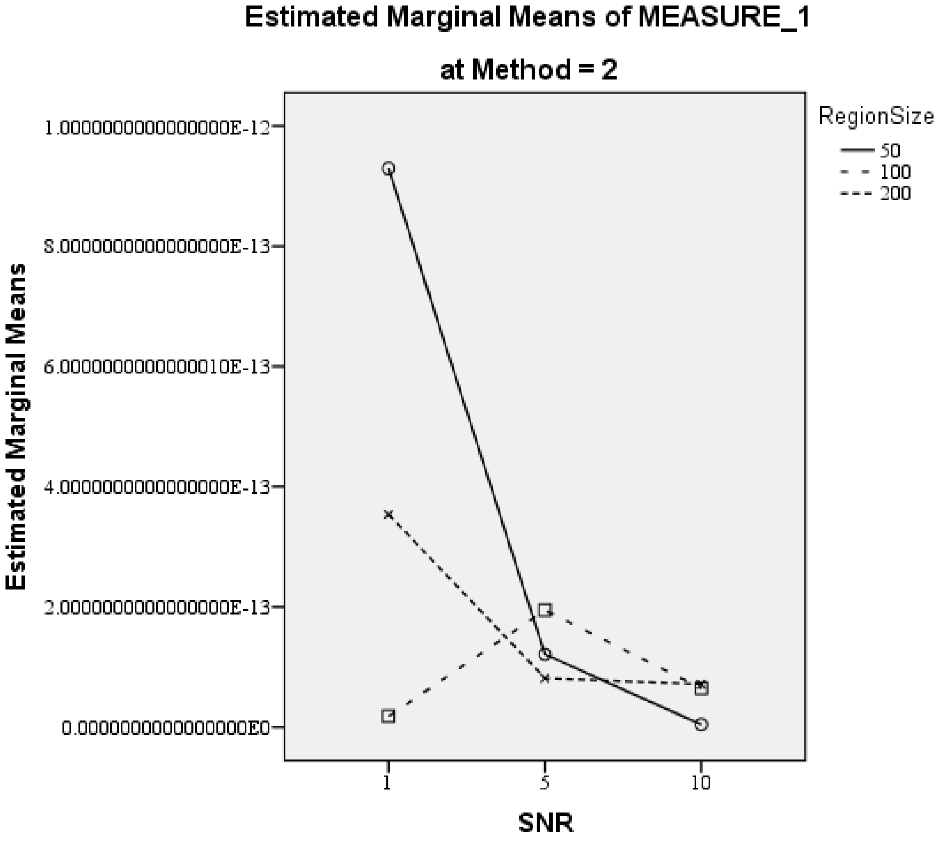

Figure 2b: Interaction plot for region size and signal-to-noise ratio using the ‘median’ method.

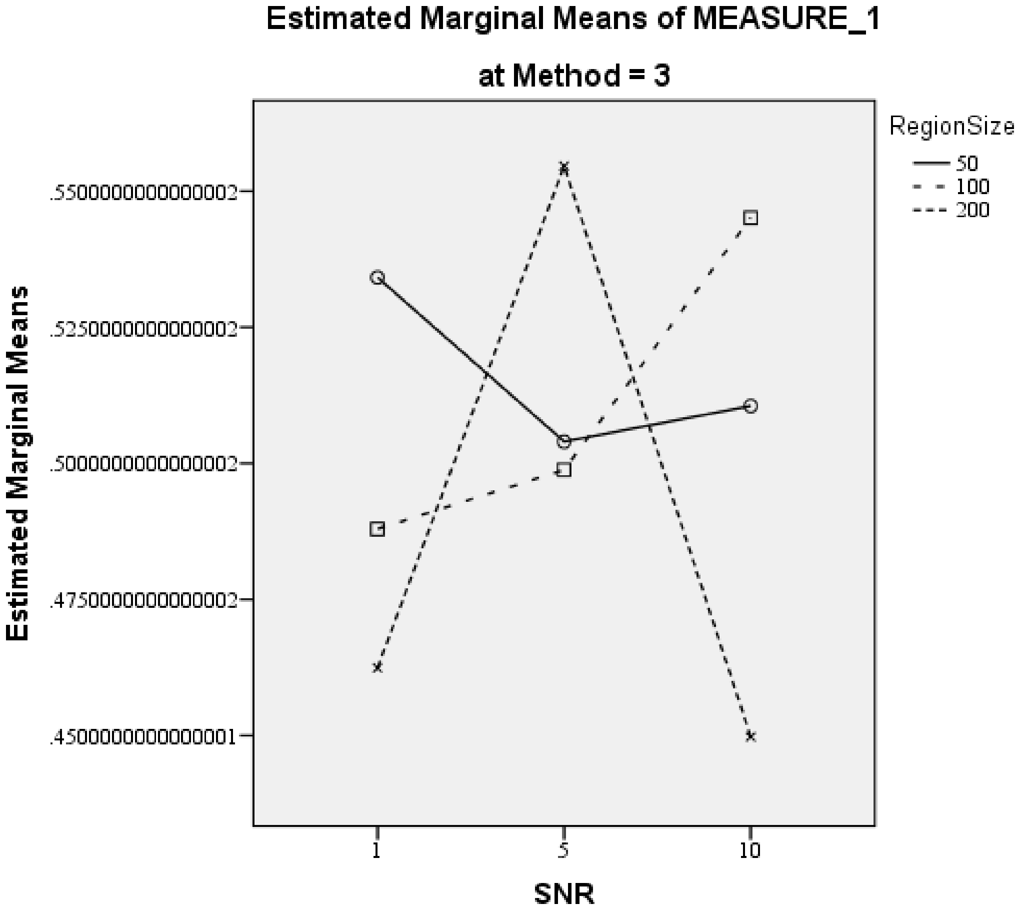

Figure 2c: Interaction plot for region size and signal-to-noise ratio using the ‘PCA’ method.

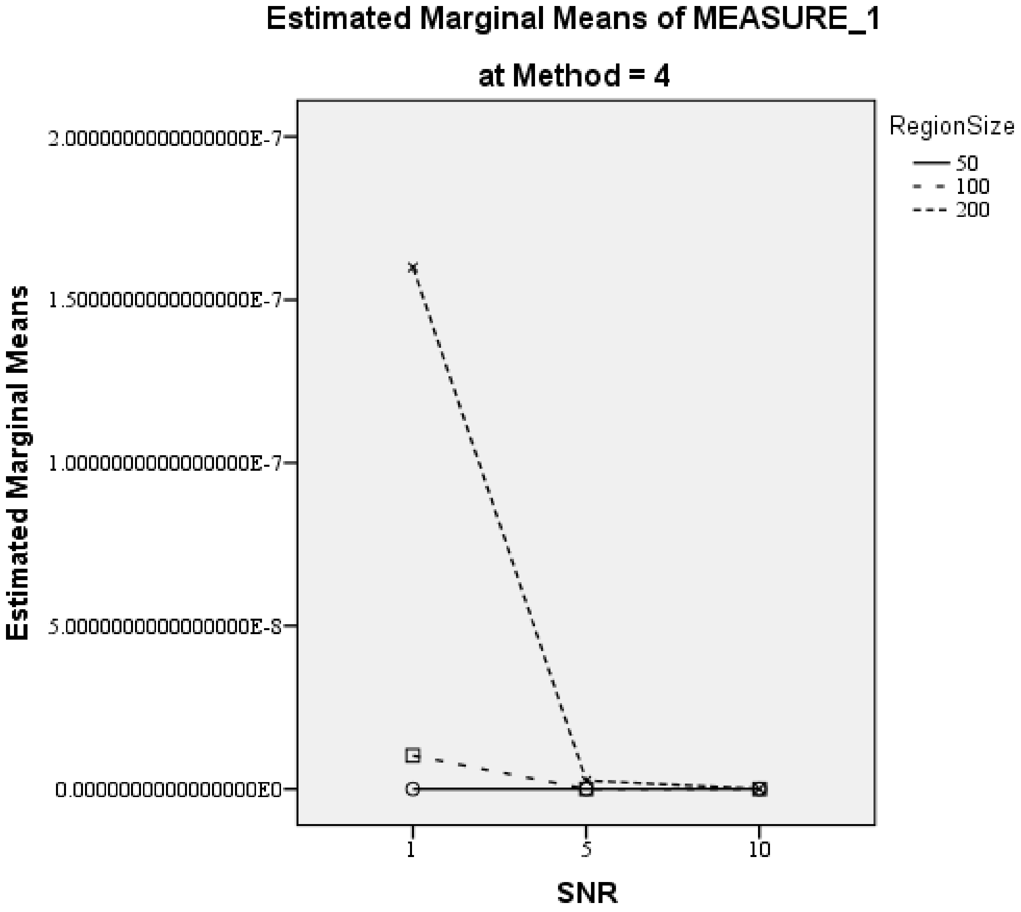

Figure 2d: Interaction plot for region size and signal-to-noise ratio using the ‘CCA’ method.