Schwaferts, Patrick:

Effects of Signal Dropout in Resting-State fMRI Data of the Human Connectome Project on Functional Connectivity

Masterarbeit, Sommersemester 2017

Fakultät für Mathematik, Informatik und Statistik

Ludwig-Maximilians-Universität München

https://doi.org/10.5282/ubm/epub.41007

Master Thesis

Master Statistics

Ludwig-Maximilians-Universität Munich

Effects of Signal Dropout in Resting-State fMRI Data of the Human Connectome Project on Functional Connectivity

Patrick Schwaferts

Advisor:

Prof. Dr. Volker Schmid

Ludwig-Maximilians-Universität Munich

External Advisors:

Prof. Dr. Russ Greiner

Dr. Sunil Kalmady

University of Alberta

Munich, 24. July 2017

Abstract

The Human Connectome Project (HCP) utilizes an eight factorial multi-band gradient-echo echo-planar-imaging sequence together with a right-to-left (RL) or left-to-right (LR) phase encoding direction in their resting-state (rs) functional magnetic resonance imaging (fMRI) acquisition setting, which is accompanied by signal dropouts, i.e. signal loss in certain brain areas. These dropouts differ in location between rs fMRI data of both phase encoding directions and their impact on rs functional connectivity (FC) as well as coping strategies suggest by the HCP are assessed in this thesis. Results indicate systematic FC differences between RL and LR data and the existence of brain regions with extraordinary dropout based impairment in FC. A relationship between signal loss and FC-impairment could be found, unless a coarse brain parcellation (AAL) was utilized and brain areas with strong FC-impairment were excluded. Accordingly, caution should be applied on rs fMRI (FC) results based on the HCP, if detailed parcellations are used or dropout areas are not accounted for. In order to reduce the spatial heterogeneity of the influence of these dropouts on FC results, either FC matrices of both phase encoding directions can be averaged or time-series of both phase encoding directions can be concatenated. With respect to the latter, however, it is crucial to implement the concatenation of time-series after and not before (or during) preprocessing of the data.

Contents

1. Introduction **1**
1.1. Human Connectome Project (HCP) . . . 1
1.2. Multi-band EPI . . . 1
1.3. Signal Dropouts in HCP . . . 2
1.4. Aims . . . 3

2. Methods **4**
2.1. Data . . . 4
2.2. Preprocessing . . . 4
2.3. Measures . . . 7

3. Analyses and Results **10**
3.1. Regions with impaired FC . . . 10
3.2. Influence of Signal Dropouts on FC . . . 13
3.3. Coping Strategies . . . 21

4. Discussion **32**
4.1. Summary of Results . . . 32
4.2. Guidelines . . . 33
4.3. Studies being potentially affected by HCP Signal Dropouts . . . 34
4.4. Limitations . . . 35
4.5. Conclusion . . . 36

References **37**

List of Figures **41**

List of Tables **41**

List of Abbreviations **42**

A. Subject IDs **43**

B. Preliminary Assessment **43**

C. Investigation using the MS444-no-smooth preprocessing pipeline **45**

D. Considerations on Coping Strategies **46**

1. Introduction

1.1. Human Connectome Project (HCP)

The Human Connectome Project (HCP, `www.humanconnectome.org/`, [1, 2]) by the WU-Minn Consortium is an extensive initiative to advance our understanding of human brain connectivity. Its main tasks are both to develop improved neuroimaging methods as well as to provide "a data set of unprecedented size and quality" [2, p.1175] and by now this project includes data of about 1200 subjects, different acquisition methods and various tasks. A considerable part of this data (about one hour per subject) is gathered using functional magnetic resonance imaging (fMRI) and based on the resting-state (rs) paradigm, in which participants are supposed to let their mind wander and not perform any specific task while being scanned. Although, usually fMRI data suffers under a trade-off between temporal and spatial resolution (as described below), the HCP managed to achieve high characteristics in both of these resolution types (repetition time $TR = 0.72s$, $2mm$ isotropic voxels) by using an eight factorial multi-band echo-planar imaging (EPI) sequence. This acquisition protocol as well as its drawbacks will be described in the next sections.

1.2. Multi-band EPI

In MRI hydrogen protons are first excited by an electromagnetic pulse, which is in the radio-frequency range (RF-pulse), and subsequently emitted electromagnetic signals are recorded (readout). Unfortunately, signals - having amplitude, phase and frequency - originate from multiple locations simultaneously and overlap. Their amplitude is of interest, as it is said to reflect neural activation [4] (using the blood oxygenation level dependent (BOLD-)contrast in fMRI [3]), so the other two properties can be used in encoding the origin of the signal, which is done as follows [5]:

> *Slice Encoding:* Because of a magnetic gradient applied along the z-axis during excitation, only one slice of tissue perpendicular to this axis is excited by the RF-pulse. Accordingly, signals resulting from this pulse have the same z position and all different slices have to be sampled subsequently in order to acquire a $3D$ image of the brain.
>
> *Encoding:* Applying a magnetic gradient after excitation and before readout along the y-axis will induce a gradient in the phase of the signals along this axis. Consequently, the phase of the signals within a slice differs along the

y-axis. This is usually along the anterior-posterior direction of the brain, but changed to the right-left (or left-right) direction in the HCP [6].

Frequency Encoding: Using another magnetic gradient during readout along the x-axis affects the frequency of the signal in an analogous manner.

Therefore, every signal in one slice has a different combination of phase and frequency respective to its location, which allows to disentangle these overlapping signals. The repetition time *TR* is the time between two consecutive RF-pulses. As in (structural) MRI each slice is excited by its own RF-pulse, *TR* gives the time to acquire one brain slice. However, fMRI typically employs the echo-planar-imaging (EPI) pulse sequence, in which the whole brain is excited by just one RF-pulse, and all slices are sampled consecutively after this single excitation pulse. Consequently, *TR* in the EPI sequence is the time to acquire the whole brain image, which is dependent on the number of slices. Sampling at a higher spatial resolution will increase the number of slices and accordingly *TR* as well, decreasing the temporal resolution of subsequent brain volumes (i.e. 3D brain images in fMRI). This is the trade-off between spatial and temporal resolution in fMRI (see also [7]).

In order to increase the temporal resolution - for a given spatial resolution - signals of multiple slices within one EPI sequence can be sampled simultaneously. This procedure is called multi-band EPI, in which the original signal needs to be reconstructed, for example by the SENSE algorithm [8]. The number of slices sampled simultaneously is referred to as multi-band factor M, which is $M = 8$ in the HCP rs-fMRI data [6, 9]. However, the use of multi-band EPI is associated with signal artifacts, which are described in the following section.

1.3. Signal Dropouts in HCP

The HCP's core tenets include acquiring data at both high temporal and spatial resolution [2], which was accomplished by the eight factorial multi-band gradient-echo EPI sequence described in the previous section. Although increasing acquisition speed, this imaging sequence is accompanied with signal dropouts, i.e. a reduced signal amplitude at certain locations in the brain. These locations are determined by the phase encoding direction, which has a right-to-left (or left-to-right) orientation in the HCP, resulting in a signal loss being strongest in the area around the orbitofrontal cortex (see Figure 1, top-left). To deal with this issue, rs-fMRI data was sampled using both a left-to-right (LR) phase-encoding setting and a right-to-left (RL) phase-encoding setting. Accordingly, locations of signal dropouts differed between these two phase encoding directions, which allows their investigation.

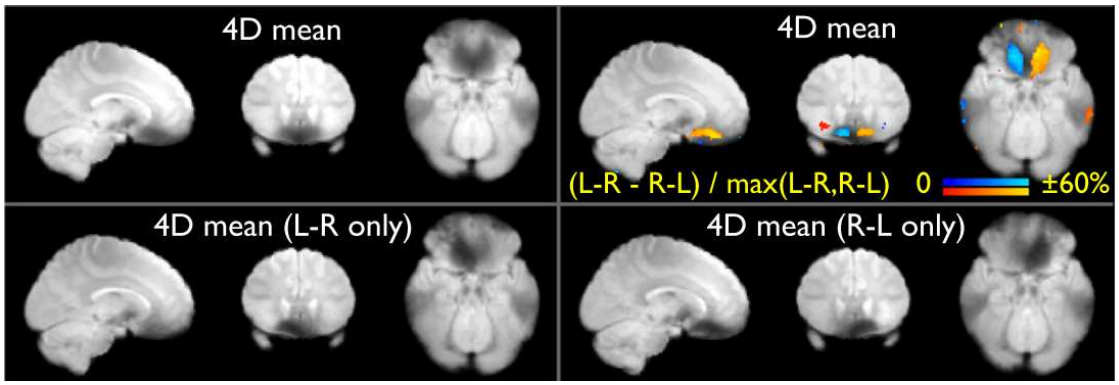

Figure 1: Mean signal over 20 subjects. Top-left: Avergae over all volumes of all subjects. Bottom-left: Only LR volumes. Bottom-right: Only RL volumes. Top-right: Relative mean differnce between LR and RL. Adapted from [6], Figure 4.

In a preliminary analysis by the HCP [6] resting-state data of 20 subjects (each with 2400 RL and 2400 LR volumes) was averaged and clearly showed dropout areas (see Figure 1, top-left). By considering each phase encoding direction separately, dropouts appear to be stronger in just one hemisphere: LR phase encoding data is characterized by a higher signal loss in the right hemisphere (see Figure 1, bottom-left) and RL phase encoding data shows more pronounced dropouts in the left hemisphere (see Figure 1, bottom-right). There are differences between the mean signal of both phase encoding directions of up to 60% (see Figure 1, top-right).

Smith *et al.* [6] report a lack of methods for combining analyses across LR and RL data and they simply suggest averaging LR and RL (functional) connectivity matrices or just concatenating time-series of both phase-encoding directions. However, for the latter they did not specify whether to concatenate voxel-wise time-series before preprocessing or region-wise time-series after preprocessing, so all three methods will be considered in this thesis.

1.4. Aims

As a basic understanding of signal dropouts in the HCP is missing, the aim of this master thesis is to investigate the influence of these signal dropouts on rs functional connectivity (FC). In addition, methods to cope with signal dropouts suggested by the HCP [6] should be assessed of how strong they are affected by signal dropouts.

2. Methods

2.1. Data

The basis of this investigation is rs-fMRI data of 100 subjects (43 males; 20 - 35 years, one subject older than 36 years; for a list of subject ids, see Appendix A) by the Human Connectome Project (HCP, by WU-Minn Consortium, `www.humanconnectome.org/`, [1, 2]) S500 release. This part of the data was collected for every subject $i \in \{1, ..., N = 100\}$ in two rs-fMRI sessions $j \in \{1, 2\}$ on separate days, each with two runs $k \in \{LR, RL\}$ of resting-state data of different phase encoding directions. During a run $T = 1200$ (functional) brain volumes were obtained in 14 minutes 24 seconds (one every $TR = 0.72s$) with a multi-band (factor 8) gradient-echo EPI imaging sequence resulting in $2mm$ isotropic voxels. Participants were instructed to look at a fixation cross in front of a dark background and to let their mind wander. All participants were scanned at the same $3T$ Siemens Connectome Skyra MRI scanner. A complete acquisition protocol can be found in [6].

2.2. Preprocessing

Rs-fMRI data was partly preprocessed by the HCP. Further preprocessing was performed using nilearn (version: 0.2.6 [10, 11]) based on python (version: 2.7.13[1]). It follows a description of the HCP based preprocessing as well as of the single preprocessing steps and in the end of this section an overview of all different pipelines in this thesis will be given. Preprocessing was performed in each run of each subject.

Preprocessing by HCP. Rs-fMRI data preprocessing by the HCP includes correction for (gradient) distortions and subject movement, bias field removal, registration of functional images to individual structural images, transformation to MNI152 atlas (by Montreal Neurological Institute) and intensity normalization to the global mean. No slice time correction was employed due to the short repetition time $TR = 720ms$ and the use of the multi-band technique. All these transformations preserve the isotropic voxel size of $2mm$. A more detailed description of these preprocessing steps is provided by Glasser *et al.* [12].

Concatenation (Voxel). Because Smith *et al.* [6] suggest averaging LR and RL FC-matrices or just concatenating time-series of different phase encoding directions to cope with signal dropouts, these procedures were evaluated for differences in

[1]Python Software Foundation. Python Language Reference, version 2.7. Available at http://www.python.org

their FC-matrices. For the latter (concatenation), RL and LR data of one session (preprocessed by HCP) were concatenated to get a long run of 2400 volumes (28 minutes 48 seconds) per session, which contains both RL and LR measurements. In pipelines with this preprocessing step (pipelines: MS444-concat-voxel and AAL-concat-voxel) further preprocessing and analyses were performed on this long run. Certainly, regressors for the nuisance regression were concatenated in these pipelines as well.

Spatial Smoothing. Data was smoothed spatially with a gaussian kernel of $4mm$ Full Width at Half Maximum (FWHM) to increase the signal-to-noise ratio, as it is assumed that signal is reflected by low and noise by high spatial frequencies [13]. However, with spatial smoothing boundaries between brain regions blur in a way that a region is influenced by measurements originating in other regions. In order to check the robustness of this investigation against the choice of kernel width, a separate analysis will be performed using a preprocessing pipeline without spatial smoothing (pipeline: MS444-no-smooth).

Parcellation. As this work aims to investigate, if signal dropouts affect FC, a brain parcellation is necessary, in which regions are small enough to be fully covered by signal dropouts. Utilizing coarser parcellations, it might be possible that even the brain areas being most affected by signal dropouts contain enough preserved signal, so that a possible effect of signal dropouts on FC cannot be detected. Accordingly, the Multiscale 444 parcellation (MS444, [14]) is used, consisting of $R_{MS} = 444$ distinct regions. Of course, results in this thesis should be generalized to rs-fMRI analyses of the HCP with coarser parcellations as well, so the investigation will be repeated with the widely used Automated Anatomical Labeling (AAL) brain atlas [15], which divides the brain into $R_{AAL} = 116$ disjoint regions (pipeline: AAL). Signals were averaged within each region, resulting in 444 or 116 time-series for each run, respectively.

Nuisance Regression. A lot of noise is present in fMRI data due to head movements as well as pulmonary and cardiovascular activity, which could be regressed out. To account for noise induced by head movements, 24 motion parameters can be included as independent variables in this regression on the mean activity time-series in each brain area [16]. Six motion parameters can be taken from the registration preprocessing step performed by the HCP, in which each volume is transformed to match a structural image via six parameter rigid body transformation [17]. Together with these six parameters, those of the previous time point as well as their squared values form those 24 motion variables in the regression. Note, that Power *et al.* [18] showed a better removal of motion related artifacts with a 36 motion param-

eter regression, but they also decided to use the 24 parameter model for a better comparison to other studies. Analogue considerations apply to this investigation. In order to cope with pulmonary and cardiovascular activity, the average signal of, respectively, white matter (WM) and cerebro-spinal-fluid (CSF) was also included in the nuisance regression, because they are supposed to contain influences of non-neural activity on the fMRI signal. Therefore, it is said that WM and CSF signal regression improve specicity of FC maps [19]. Regression of the global signal (mean signal per volume), however, was found to produce spurious negative correlations, which should be avoided [20, 21]. Jo *et al.* [22] argued that signal changes relevant for rs-fMRI in gray matter should only be reflected slightly in WM signals due to the propagation direction of the BOLD signal along the venous vasculature.

Temporal Filtering. Spontaneous fluctuations in the BOLD signal, which are relevant in rs-fMRI, are found to be strongest in the frequency range of $0.009Hz - 0.08Hz$ [23, 24], so a corresponding bandpass filter was applied. Note, that temporal filtering should be applied after nuisance regression in order to prevent the reintroduction of unwanted frequencies [25].

Concatenation (Region). Another possibility to concatenate time-series as a method of coping with signal dropouts would be for region-specific preprocessed time-series. Therefore, in additional pipelines each run was preprocessed separately and then both RL and LR time-series per region were concatenated before FC extraction (pipelines: MS444-concat-region and AAL-concat-region).

FC Extraction. For each pair of regions (r, r') with $r, r' \in \{1, 2, ..., R\}$, $r \neq r'$ and $R \in \{R_{MS}, R_{AAL}\}$ the Pearson correlation (FC-value) between their activity time-series is calculated, resulting in a symmetric 444×444 or 116×116 matrix specifying the functional connectivity (FC) for the MS444 or AAL parcellation, respectively.

Final Pipelines. Altogether the following preprocessing pipelines were used. Note, that all pipelines started with the data being already preprocessed by the HCP.

MS444: Spatial smoothing with $4mm$ FWHM, parcellation using Multiscale 444 atlas, nuisance regression, temporal filtering, FC extraction.

MS444-no-smooth: Parcellation using Multiscale 444 atlas, nuisance regression, temporal filtering, FC extraction.

AAL: Spatial smoothing with $4mm$ FWHM, parcellation using AAL atlas, nuisance regression, temporal filtering, FC extraction.

MS444-concat-voxel: Concatenation (voxel), spatial smoothing with $4mm$

FWHM, parcellation using Multiscale 444 atlas, nuisance regression, temporal filtering, FC extraction.

MS444-concat-region: Spatial smoothing with $4mm$ FWHM, parcellation using Multiscale 444 atlas, nuisance regression, temporal filtering, concatenation (region), FC extraction.

AAL-concat-voxel: Concatenation (voxel), spatial smoothing with $4mm$ FWHM, parcellation using AAL atlas, nuisance regression, temporal filtering, FC extraction.

AAL-concat-region: Spatial smoothing with $4mm$ FWHM, parcellation using AAL atlas, nuisance regression, temporal filtering, concatenation (region), FC extraction.

2.3. Measures

FC-Impairment. For every region $r \in \{1, 2, ..., R\}$ a measure of FC-impairment[2] $f_{r,drop}$ between acquisition runs *RL* and *LR* was calculated:

$$f_{r,drop} = \frac{1}{R-1} \sum_{r',r' \neq r} \left| \frac{1}{2N} \sum_{i,j} (fc^{RL}_{i,j,r,r'} - fc^{LR}_{i,j,r,r'}) \right| \tag{1}$$

with $R \in \{R_{MS}, R_{AAL}\}$ regions and $fc^{k}_{i,j,r,r'}$ being the FC-value (Pearson correlation) of subject i in session j between regions r and r' in run k. The terms inside the absolute value of equation 1 considering all possible pairs of regions (r, r')

$$FC^{RL-LR}_{a} = \left[\frac{1}{2N} \sum_{i,j} (fc^{RL}_{i,j,r,r'} - fc^{LR}_{i,j,r,r'}) \right]_{r,r'} \tag{2}$$

constitute the average difference matrix using atlas $a \in \{MS, AAL\}$: FC^{RL-LR}_{MS} (see Figure 5, top-left) and FC^{RL-LR}_{AAL} (see Figure 5, bottom-left).

Signal Loss. A measure[3] $l_{r,drop}$ of how strong a region $r \in \{1, 2, ..., R\}$ is affected by signal dropouts is calculated as absolute relative signal loss between acquisitions *RL* and *LR*:

$$l_{r,drop} = \left| \frac{s^{RL}_{r} - s^{LR}_{r}}{max\{s^{RL}_{r}, s^{LR}_{r}\}} \right| \tag{3}$$

[2]The index *drop* indicates that this measures is based on the *RL* - *LR* comparison, as opposed to the control measure (see below).

[3]Again, the index *drop* refers to the comparison between *RL* and *LR*.

with

$$s_r^k = \frac{1}{2NT}\frac{1}{\#voxels}\sum_{i,j,t,v_r} s_{i,j,k,t,v_r} \tag{4}$$

being the average signal s_{i,j,k,t,v_r} in a region r consisting of voxels v_r over all time-points t ($T = 1200$), all sessions j and all subjects i ($N = 100$) for run k. $\#voxels$ is the number of voxels v_r in region r.

Control Measures. As a control, both signal loss and FC-impairment measures were calculated comparing between session $j = 1$ and $j = 2$ using the same notation as above:

$$l_{r,cont} = \left| \frac{s_r^1 - s_r^2}{max\{s_r^1, s_r^2\}} \right| \tag{5}$$

with

$$s_r^j = \frac{1}{2NT}\frac{1}{\#voxels}\sum_{i,k,t,v_r} s_{i,j,k,t,v_r} \tag{6}$$

and

$$f_{r,cont} = \frac{1}{R-1}\sum_{r',r'\neq r}\left|\frac{1}{2N}\sum_{i,k}(fc^1_{i,k,r,r'} - fc^2_{i,k,r,r'})\right| \tag{7}$$

with $fc^j_{i,k,r,r'}$ being the FC-value of subject i in session j, run k and connection (r, r'). Again, the terms inside the absolute value of equation 7 for all pairs of regions (r, r') define the average difference matrix analogue to equation 2: FC^{s1-s2}_{MS} (see Figure 5, top-right) and FC^{s1-s2}_{AAL} (see Figure 5, bottom-right).

Method Difference. As mentioned above, the HCP [6] suggests two strategies for coping with signal dropouts in rs-fMRI investigations: averaging of FC-matrices $s = ave$ or concatenation of time-series over runs of both phase encoding directions. For latter, it was not specified whether this concatenation should be done with voxel-based time-series before preprocessing $s = concat(voxel)$ or with region-based time-series after preprocessing $s = concat(region)$, so both approaches are being investigated in this thesis.

In the averaging method FC values are calculated by

$$fc^{ave}_{i,j,r,r'} = \frac{1}{2}(fc^{RL}_{i,j,r,r'} + fc^{LR}_{i,j,r,r'}) \tag{8}$$

utilizing the notation as above and in the concatenation methods FC values $fc^{concat(voxel)}_{i,j,r,r'}$ and $fc^{concat(region)}_{i,j,r,r'}$ are extracted by the respective preprocessing pipeline (MS444-concat-voxel or AAL-concat-voxel vs. MS444-concat-region or AAL-concat-region).

If there is no difference between these methods, either of them can be applied

without concern, but if a difference exists, its relationship with signal dropouts can be assessed. To do this, a measure $d_{r,voxel}$ of difference in FC between both averaging and voxel-wise concatenation methods is calculated for each region $r \in \{1, 2, ..., R\}$:

$$d_{r,voxel} = \frac{1}{R-1} \sum_{r',r' \neq r} \left| \frac{1}{2N} \sum_{i,j} (fc^{ave}_{i,j,r,r'} - fc^{concat(voxel)}_{i,j,r,r'}) \right| \tag{9}$$

with $R \in \{R_{MS}, R_{AAL}\}$ regions and $fc^{s}_{i,j,r,r'}$ being the FC-value (Pearson correlation) of subject i in session j between regions r and r' using strategy $s \in \{ave, concat(voxel)\}$. The analogue measure $d_{r,region}$ was calculated for comparing the averaging method with the second version of the concatenation method as

$$d_{r,region} = \frac{1}{R-1} \sum_{r',r' \neq r} \left| \frac{1}{2N} \sum_{i,j} (fc^{ave}_{i,j,r,r'} - fc^{concat(region)}_{i,j,r,r'}) \right| \tag{10}$$

with notation as above and $s \in \{ave, concat(region)\}$.

3. Analyses and Results

Analysis were performed using R (version: 3.3.2, [26]): Weighted least squares regressions were calculated with the package "nlme" (version: 3.1.128) and for the algorithm of [27] to find the closest positive-definite matrix to a given matrix the package "corpcor" (version: 1.6.9) was used.

3.1. Regions with impaired FC

In order to find MS444 brain regions with impaired FC due to differences between RL and LR acquisition settings, outliers concerning the FC-impairment measure are detected based on a threshold criterion of $mean + 2.5sd$. A histogram for FC-impairment values (mean±sd: $4.18\% \pm 1.02\%$) of MS444 regions is depicted in Figure 2 (top) showing the existence of outlier values. Regions with the highest FC-impairment are listed in Table 1 and their location can be seen in Figure 3. Out of the 9 areas with a FC-impairment of more than 2.5 standard deviations above the mean, there are 4 regions $(349, 197, 379, 243)$, which coincide with the parts of the brain suffering the highest signal dropouts reported by [6] (see Figure 1). Based on the MS444-no-smooth pipeline for preprocessing, 9 outlier regions with a FC-impairment value above the threshold could be detected, of which 7 coincide with those listed in Table 1 (indicated with *; the other regions are 119 and 161), demonstrating the robustness of these results with respect to spatial smoothing.

Table 1: MS444 brain regions with highest FC-impairment values. Areas indicated with * are also found to be outliers (i.e. a FC-impairment value of $> mean + 2.5sd$) based on data from the MS-444-no-smooth pipeline.

Area	FC-impairment (in percentage points)	standard deviations above mean	signal loss (in %)
349 *	9.45	5.16	13.0
197 *	8.80	4.52	29.6
383 *	7.55	3.30	4.5
379 *	7.21	2.97	3.3
223 *	7.10	2.86	9.7
243 *	6.92	2.68	7.3
311	6.87	2.63	4.6
426 *	6.77	2.54	10.6
359	6.74	2.51	18.0

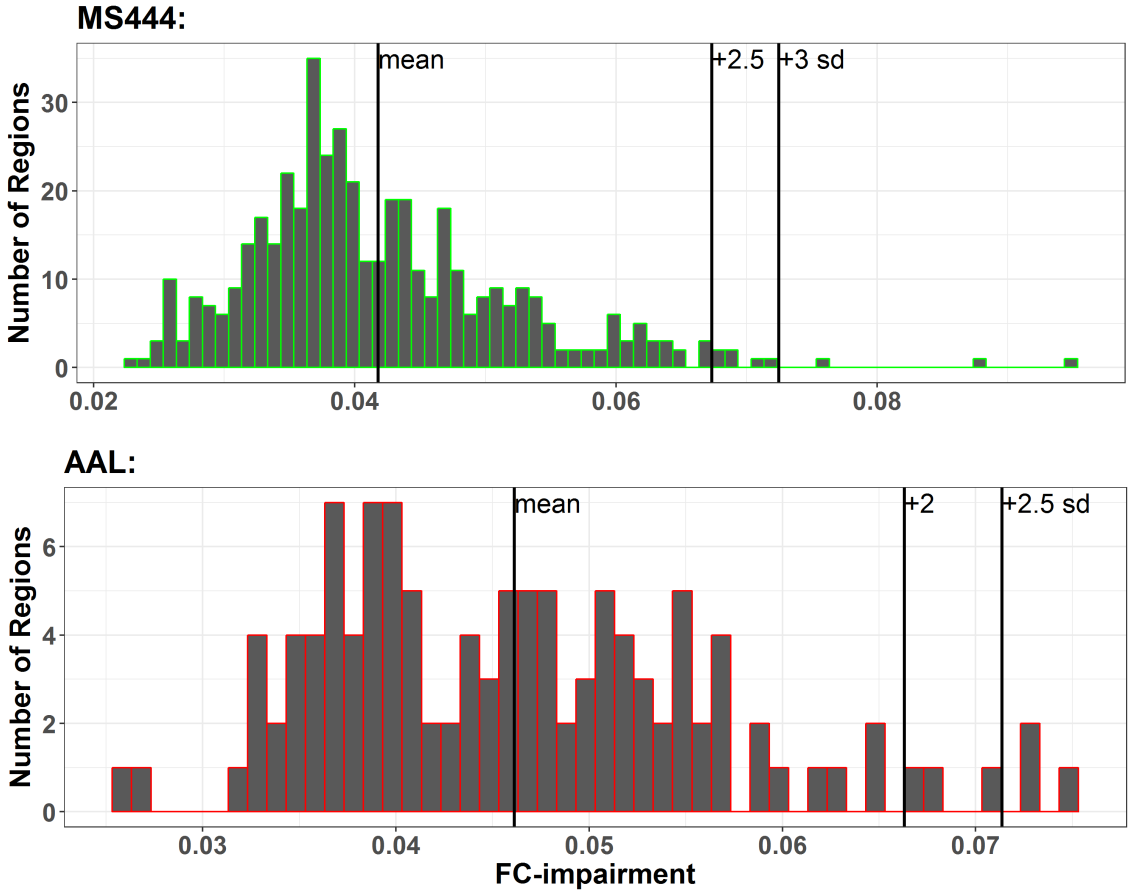

Figure 2: Histograms of FC-impairment values: MS444 and AAL. Top: The histogram is based on 444 FC-impairment values using the MS444 atlas. Black lines indicate the mean value and the values being 2.5 as well as 3 standard deviations above the mean. Utilizing the lower value ($2.5sd$) as threshold criterion results in 9 MS444 outlier regions. Bottom: The histogram addresses the AAL parcellation and is based on 116 FC-impairment values. Black lines indicate mean and values being 2 and 2.5 standard deviations above the mean. With the $2sd$ criterion, 6 outlier regions are defined in the AAL parcellation.

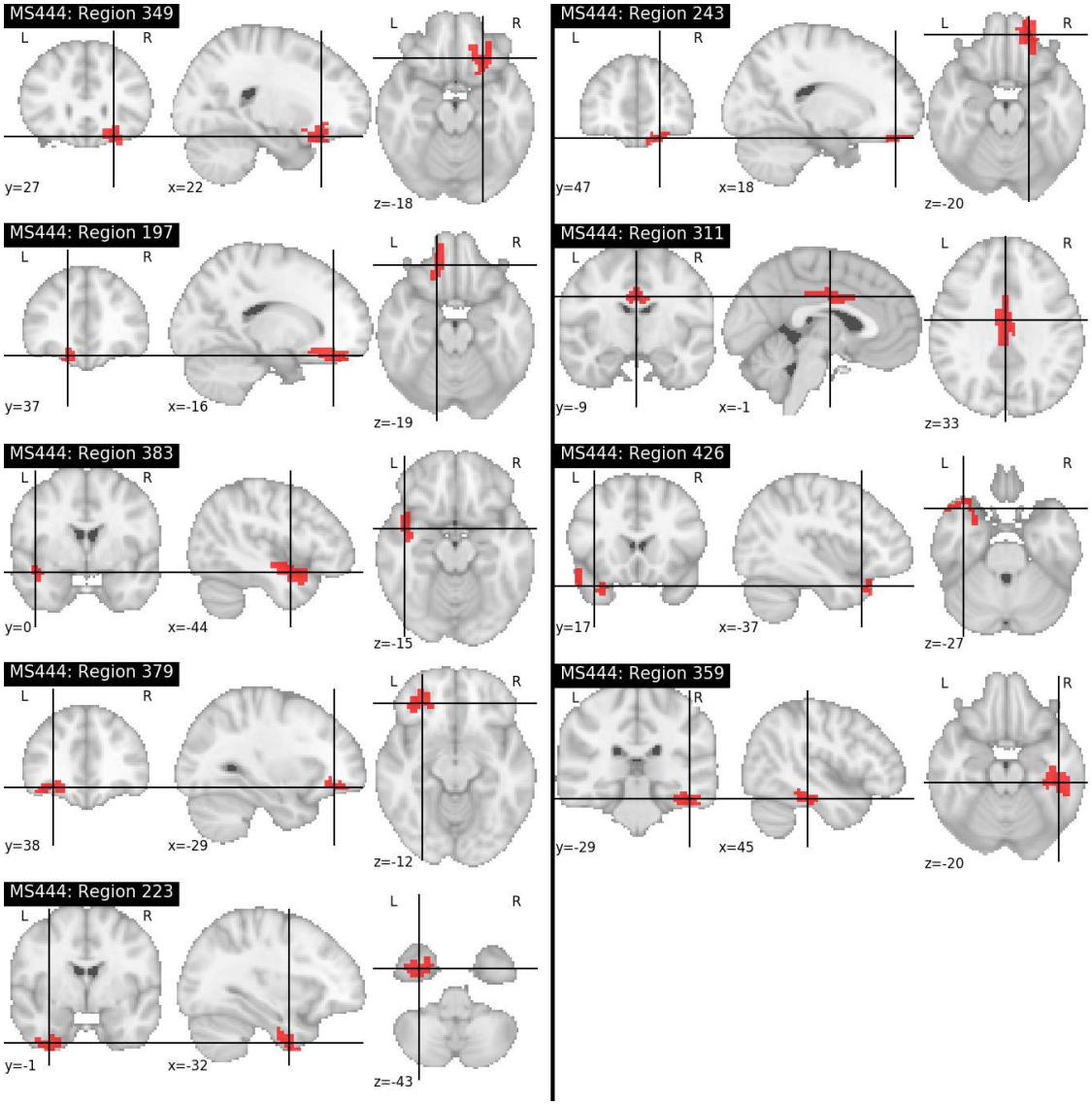

Figure 3: MS444 brain regions with strongest FC-impairment. These regions have a FC-impairment value of more than 2.5 standard deviations above the mean and refer to those listed in Table 1. There are 5 regions located in the left hemisphere (197, 383, 379, 223, 426), 3 regions in the right hemisphere (349, 243, 359) and one region in both hemispheres (311).

In the AAL brain parcellation, a threshold criterion of $mean + 2.5sd$ for the FC-impairment value (mean±sd: $4.61\% \pm 1.01\%$, see Figure 2, bottom) yields only three outlier regions, so regions with a FC-impairment higher than $mean + 2sd$ are considered as well. These regions are listed in Table 2 and depicted in Figure 4. Among them are also regions (21, 5, 22) corresponding to the strongest signal dropout regions described by [6].

Interestingly, most of the AAL (5/6) and many MS444 (5/8, as area 311 is located in both hemispheres) outlier regions are located in the left hemisphere. Possible explanations for this finding might be a difference in signal dropout strength or in FC between hemispheres (see e.g. [28]). However, the exact causes of this finding are unclear and demand further exploration.

Table 2: AAL brain regions with highest FC-impairment values.

Area	Location	FC-impairment (in percentage points)	standard deviations above mean	signal loss (in %)
21	Olfactory L	7.45	2.81	12.9
5	Frontal Sup Orb L	3.30	2.65	7.4
87	Temporal Pole Mid L	7.27	2.63	12.0
77	Thalamus L	7.04	2.40	1.98
22	Olfactory R	6.81	2.17	8.56
33	Cingulum Mid L	6.65	2.02	3.59

This analysis is based on outliers detected by a FC-impairment value of more than 2.5 (or 2 in the AAL analysis) standard deviations above the mean. Another common criterion for defining outliers is a value higher than 1.5 interquartile ranges above the $75th$ percentile [29]. In the MS444 analysis, this criterion would yield an outlier threshold of 6.46%, resulting in 14 outlier areas. Accordingly, the method applied in this work is more conservative. In the AAL analysis, however, the outlier threshold would be 7.27% with the interquartile range criterion, resulting in just 1 outlier region. An assessment of how this outlier criterion would have influenced other results in this thesis, is given at the end of the following section.

3.2. Influence of Signal Dropouts on FC

MS444. Furthermore, it is interesting to not only detect outlier regions characterized by a high FC-impairment value, but also to assess the influence of signal loss on FC-impairment considering all regions. To do this, a weighted least squares regression of signal loss on FC-impairment is calculated, including control measures to assess the

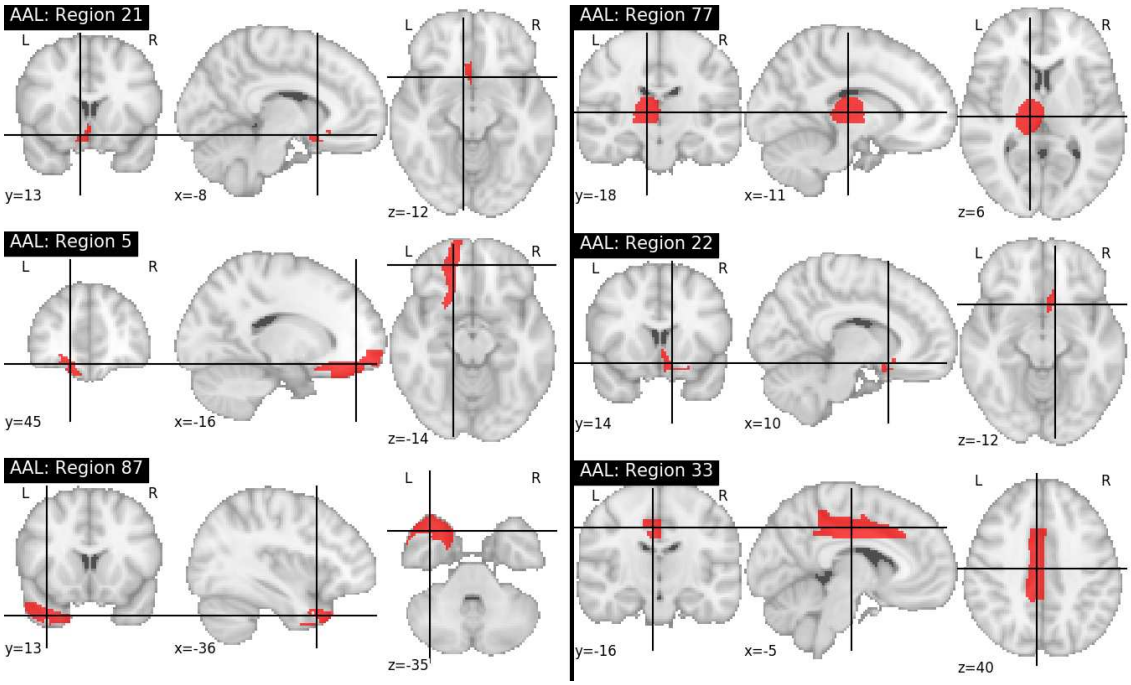

Figure 4: AAL brain regions with strongest FC-impairment. These regions have a FC-impairment value of more than 2 standard deviations above the mean and refer to those listed in Table 2. There are 5 regions located in the left hemisphere (21, 5, 87, 77, 33) and one regions in the right hemisphere (22).

baseline behaviour of the measures in use. In a preliminary assessment (see Appendix B), it appears that the distribution of FC-impairment values has a pronounced right-sided tail, which corresponds to the skewness in the histogram above (see Figure 2). Therefore, a log-linear model with a log-transformation of FC-impairment is deployed [30]. Another aspect revealed by this preliminary analysis (see Appendix B), is a heteroscedastic behaviour of FC-impairment values with respect to signal loss, which is modelled by specifying weights in the variance structure (see [31]).

In addition, FC-impairment values are related to each other by construction: The value of a region is the mean of all its absolute FC-difference values (see equations 1 and 7). So, this region has one of its FC-difference values in common with every other area, respectively. Therefore, each FC-difference value states how strong the FC-impairment measure of both areas constituting this connection are related to each other. Consequently, the relationship between all 444 FC-impairment values in the MS444 atlas is expressed by the FC-difference matrix FC_{MS}^{RL-LR} depicted in Figure 5 (top-left). Similar considerations hold for the relationship among control FC-impairment values, which is characterized by the control FC-difference matrix FC_{MS}^{s1-s2} (see Figure 5, top-right).

Taking all FC-impairment values of both dropout and control investigation together,

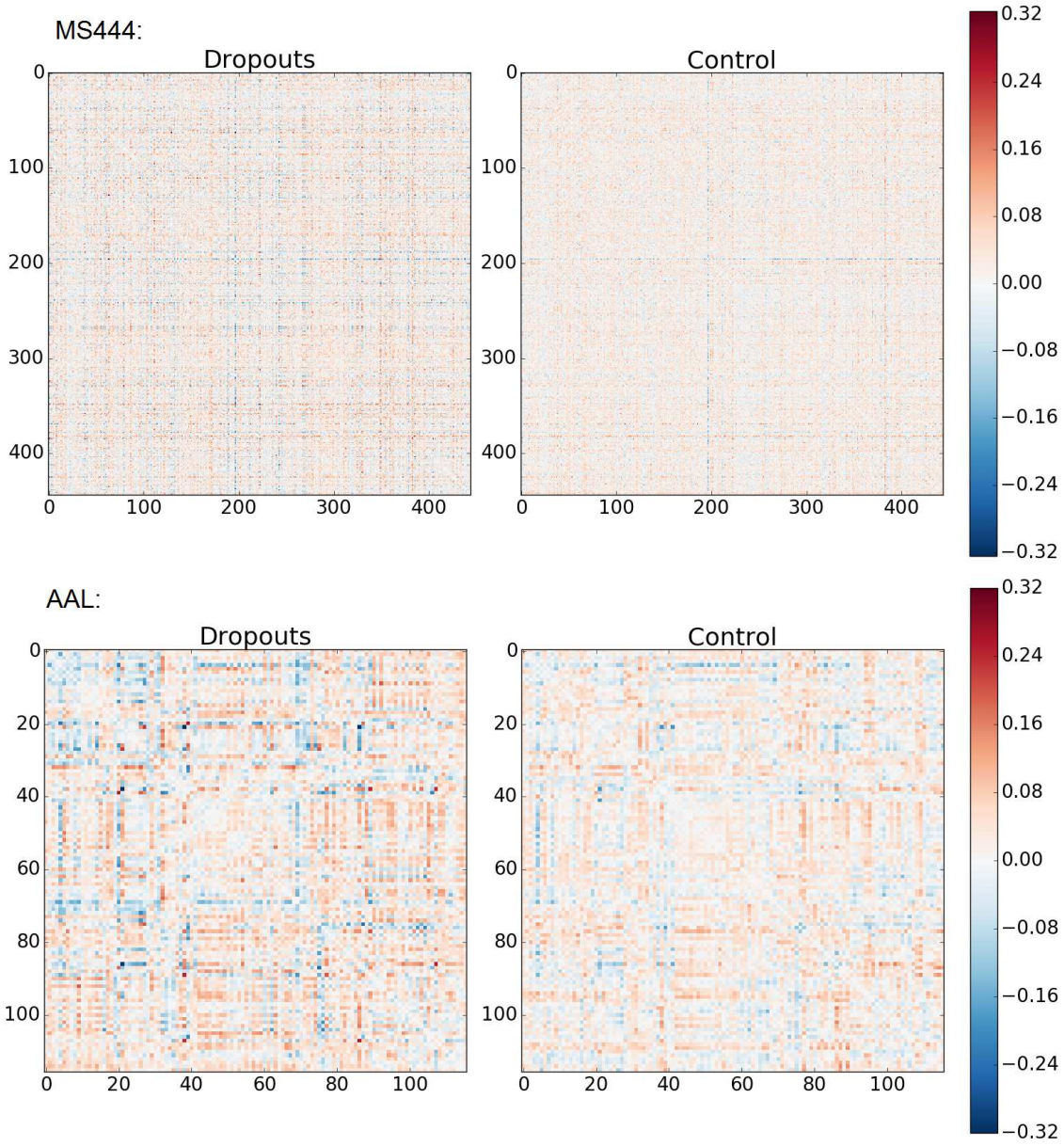

Figure 5: FC-difference matrices. Top-left: FC_{MS}^{RL-LR}. Top-right: FC_{MS}^{s1-s2}. Bottom-left: FC_{AAL}^{RL-LR}. Bottom-right: FC_{AAL}^{s1-s2}.

their relationship is expressed by the block diagonal matrix of both FC-difference matrices:

$$FC_{MS}^{\Delta} = \left[\begin{array}{c|c} FC_{MS}^{RL-LR} & \mathbf{0} \\ \hline \mathbf{0} & FC_{MS}^{s1-s2} \end{array}\right] \tag{11}$$

with $\mathbf{0}$ being a 444 x 444 zero-matrix.

The matrix being positive-definite and closest to FC_{MS}^{Δ} was computed by the algorithm of Higham [27] and can be used as correlation matrix $cor(f_{(r,m)}, f_{(r,m)'})$ with $(r,m) \neq (r,m)'$ for modelling the dependencies within the FC-impairment measure.

Accordingly, with the type of measure indicated by $m \in \{drop, cont\}$, the overall log-linear model is

$$log(f_{r,m}) = \beta_0 + l_{r,m}\beta_l + m\beta_m + ml_{r,m}\beta_{l,m} + \epsilon_{r,m} \tag{12}$$

where $\epsilon_{r,m}$ is the normal distributed error with mean 0, $cor(f_{(r,m)}, f_{(r,m)'})$ is the correlation matrix specified theoretically (as described above) and variance weights are induced by the exponential variance structure $Var(\epsilon_{r,m}) = \sigma^2 exp(2\delta_m l_{r,m})$ [31], in which δ_m is estimated to be $\hat{\delta}_{drop} = 3.66$ and $\hat{\delta}_{cont} = -15.51$. This relationship of signal loss to FC-impairment based on the MS444 pipeline can be seen in Figure 6 (top). All β-estimates are listed in Table 3 (top) and differ significantly from 0. As can be seen in the QQ-plot of the residuals (see Figure 7, top-left) and in the "Fitted vs. Residuals"-plot (see Figure 7, middle) the model fit was good.

Table 3: MS444 regression results.

All Data				
Coefficient	Value	Standard Error	t-value	p-value
β_0	-3.46	0.018	-193.46	< 0.0001
β_l	-4.06	1.111	-3.65	0.0003
β_m	0.17	0.025	6.86	< 0.0001
$\beta_{l,m}$	6.54	1.182	5.53	< 0.0001
β_0^d	-3.30	0.018	-185.53	< 0.0001
β_l^d	2.57	0.380	6.75	< 0.0001
Outliers Excluded				
Coefficient	Value	Standard Error	t-value	p-value
β_0	-3.46	0.018	-192.76	< 0.0001
β_l	-4.06	1.111	-3.64	0.0003
β_m	0.17	0.025	7.18	< 0.0001
$\beta_{l,m}$	5.97	1.169	5.11	< 0.0001
β_0^d	-3.29	0.018	-187.16	< 0.0001
β_l^d	2.01	0.338	5.96	< 0.0001

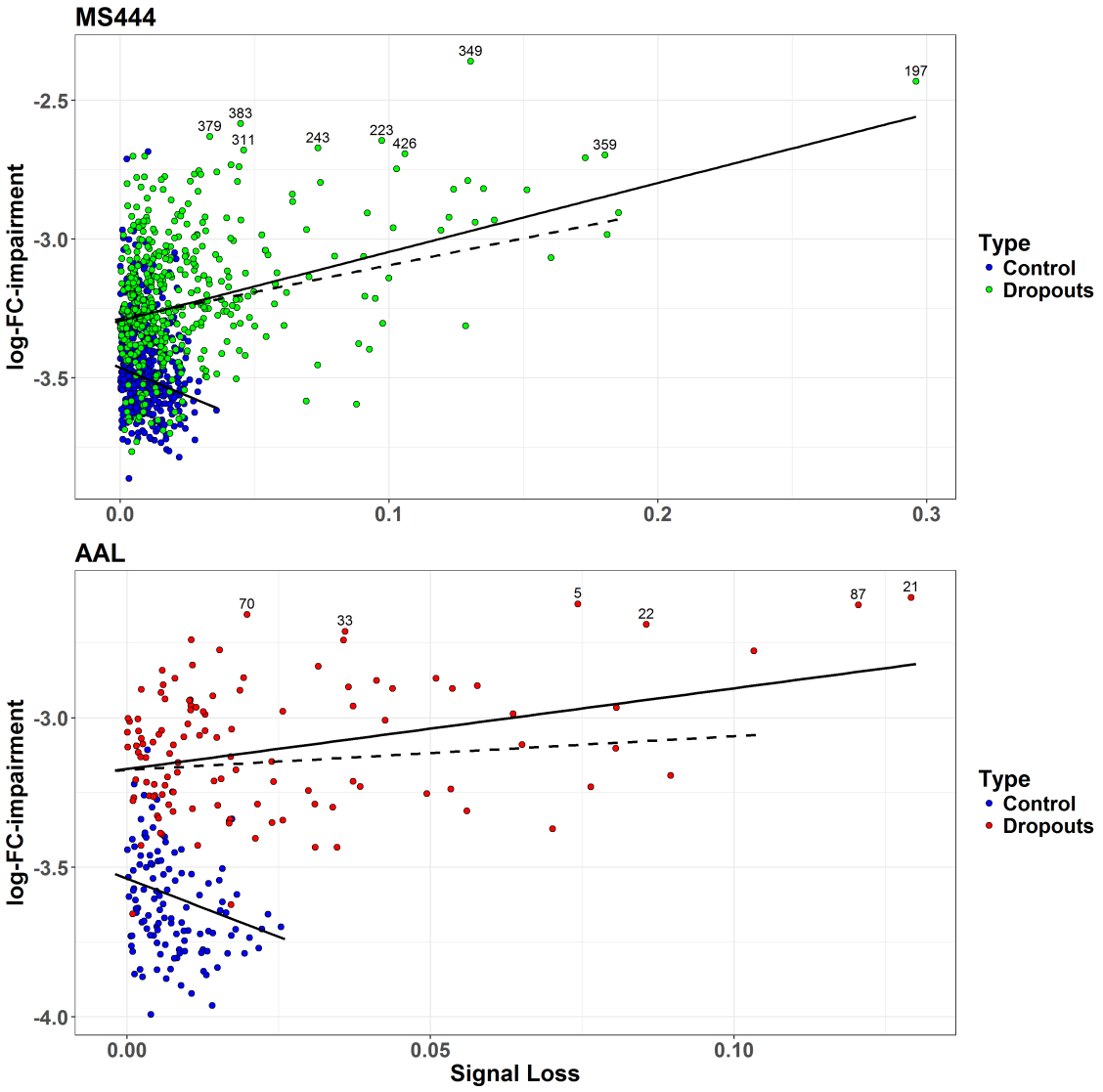

Figure 6: Influence of signal loss on FC-impairment. Scatterplots show the relationship between signal loss (x-axis) and the log-transformation of FC-impairment (y-axis) in the MS444 parcellation (top) and the ALL parcellation (bottom). Black solid lines indicate the prediction according to the weighted least squares regression using the model in equation 12. Labelled points refer to the regions defined as outlier regions listed in Tables 1 and 2. Dashed lines indicate the prediction of the weighted least squares regression, in which outlier regions (labelled) were excluded in the dropout-data. The dashed line in the AAL parcellation (bottom) is almost horizontal, reflecting a non-significant slope (see Table 4).

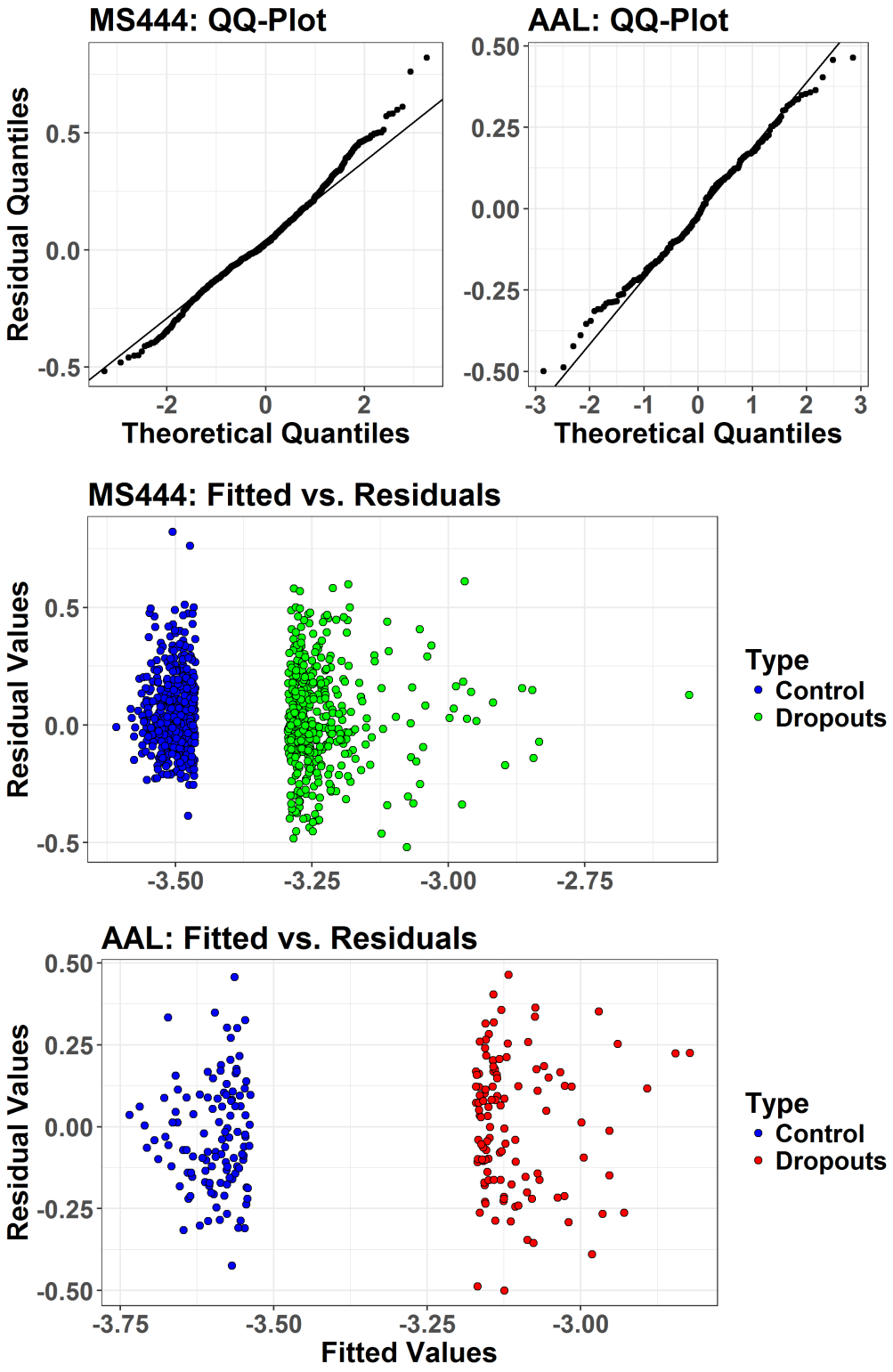

Figure 7: Model Fit: MS444 and AAL analyses. Top-left: QQ-plot in the MS444 analysis. Top-right: QQ-plot in the AAL analysis. Middle: "Fitted vs. Residuals" in the MS444 analysis. Bottom: "Fitted vs. Residuals" in the AAL analysis. All these plots refer to the models (MS444 or AAL) in equation 12.

Interpreting these results, the mean FC-impairment value is stated to be $exp(\hat{\beta_m}) = 1.185$ times the mean of the control FC-impairment values and the influence of signal loss on FC-impairment is different in both types of measures (dropouts vs control), as $\beta_{l,m}$ differs significantly from 0.

The nature of this effect is examined in a subsequent weighted least squares regression based on dropout data $m = drop$ only. Using the notation as above, the model is as follows:

$$log(f_{r,drop}) = \beta_0^d + l_{r,drop}\beta_l^d + \epsilon_{r,drop} \tag{13}$$

In this subsequent analysis, $\epsilon_{r,drop}$ is the normal distributed error with mean 0, $cor(f_{r,drop}, f_{r',drop})$ is the upper left part of $cor(f_{(r,m)}, f_{(r,m)'})$ (i.e. the positive-definite matrix being closest to FC_{MS}^{RL-LR}, see Figure 5, top-left) and the exponential variance structure is $Var(\epsilon_{r,drop}) = \sigma^2 exp(2\delta_{drop}^d l_{r,drop})$ [31], with the estimation $\hat{\delta}_{drop}^d = 2.50$.

Results of this subsequent analysis are listed in Table 3 (top). β_l^d constitutes a significant effect stating a multiplication of the FC-impairment value with $exp(\hat{\beta_l^d} \cdot 0.01) = 1.026$ for each signal loss increase of 0.01.

Typically, outliers have a strong influence on regression estimates, which might be responsible for a positive relationships detected in this regression analysis. In order to make sure that a positive relationship between signal loss and FC-impairment is valid with respect to all regions (in contrast to just outlier regions), this analysis is redone on data excluding outlier regions (see Table 1 and labels in Figure 6, top) in the "dropouts" comparison type. The correlation matrix was adapted accordingly. Weighted least squares regression results are depicted as dashed line in Figure 6 (top) and are listed in Table 3 (bottom), still showing significant effects.

As both β_l^d and $\beta_{l,m}$ represent significant effects, an impairment of FC due to signal dropouts in rs-data of the HCP is plausible, which does affect at least the Multiscale 444 brain parcellation [14].

Repeating these analyses on rs-data without spatial smoothing (MS444-no-smooth pipeline) yield a similar pattern of results (see Appendix C) and show a robustness of these findings against different kernel widths in spatial smoothing during preprocessing.

As discussed in section 3.1, different threshold criterions for defining outlier regions are possible. With the interquartile range criterion, there would have been 14 instead of 9 outlier regions in the MS444 analysis. These 5 extra regions have various different - not solely high - signal loss values (0.009, 0.005, 0.173, 0.041 and 0.044; see Figure

6, top), so that similar results in the analysis without these 14 outlier regions could be expected based on this more liberal threshold criterion for defining outliers.

AAL. The influence of signal loss on FC-impairment is also assessed for the AAL parcellation (see Figure 6, bottom). The analysis is analogue to the analysis of the MS444 atlas described above. Certainly, in the overall model the correlation structure of FC-values is now specified by the positive-definite matrix being closest [27] to the block-diagonal matrix of FC-difference matrices in the AAL parcellation (see Figure 5, left)

$$FC^{\Delta}_{AAL} = \left[\begin{array}{c|c} FC^{RL-LR}_{AAL} & \mathbf{0} \\ \hline \mathbf{0} & FC^{s1-s2}_{AAL} \end{array}\right] \tag{14}$$

with $\mathbf{0}$ being a 116 x 116 zero-matrix. Estimates in the exponential variance structure [31] were estimated to be $\hat{\delta}_{drop} = 3.66$ and $\hat{\delta}_{cont} = -26.35$ in the overall model and $\hat{\delta}^{d}_{drop} = 4.09$ in the model restricted to "dropout" data only. As can be seen in Figure 7 (top-right, bottom) the model fit in this log-linear weighted least squares regression was good as well. Without excluding outlier regions (as listed in Table 2 and labelled in Figure 6, bottom), all β-estimates are significant (see Table 4, top), showing an impairment of FC in AAL brain regions associated with signal loss. However, excluding these outlier regions (see Table 4, bottom), β^{d}_{l} is not significant ($p = 0.238$). Accordingly, the dashed line in Figure 6 (bottom) is close to horizontal.

Table 4: AAL regression results.

All Data				
Coefficient	Value	Standard Error	t-value	p-value
β_0	-3.54	0.029	-122.35	< 0.0001
β_l	-7.77	2.220	-3.50	0.0006
β_m	0.37	0.042	8.78	< 0.0001
$\beta_{l,m}$	10.47	2.375	4.41	< 0.0001
β^{d}_0	-3.17	0.030	-106.58	< 0.0001
β^{d}_l	2.66	0.851	3.12	0.0023
Outliers Excluded				
Coefficient	Value	Standard Error	t-value	p-value
β_0	-3.54	0.029	-123.26	< 0.0001
β_l	-7.80	2.225	-3.50	0.0006
β_m	0.36	0.042	8.70	< 0.0001
$\beta_{l,m}$	8.92	2.400	3.72	0.0003
β^{d}_0	-3.17	0.030	-107.00	< 0.0001
β^{d}_l	1.08	0.913	1.19	0.238

Taken together, results state that there are AAL regions with impaired FC-values (see Table 2), however, for the remaining AAL regions an influence of signal loss on

FC-impairment cannot be assumed.

With outlier regions in the AAL parcellation defined by the interquartile range criterion described in section 3.1, there would have been only one outlier area. A regression, based on the data set with this single area being excluded, would have had a significant β_l^d-value ($p = 0.0089$). Consequently, it would have been reasonable to assume regions with impaired FC due to signal loss among non-outlier regions, which - together with this single outlier area - could resemble the 6 outlier regions detected with the threshold criterion used in this thesis.

3.3. Coping Strategies

MS444. The HCP suggests different methods (averaging and concatenation) to cope with signal dropouts [6]. Systematic differences between two methods can be evaluated by examining the average difference of their FC matrices (over all subjects and sessions). If no systematic differences between two methods exist, this matrix is supposed to be a zero-matrix.

The comparison of the voxel-wise concatenation method with the averaging method, results in the FC-difference matrix[4]

$$FC_{MS}^{ave-concat(voxel)} = \left[\frac{1}{2N} \sum_{i,j} (fc_{i,j,r,r'}^{ave} - fc_{i,j,r,r'}^{concat(voxel)}) \right]_{r,r'} \tag{15}$$

as depicted in Figure 8 (top-left) with notation from section 2.3. This matrix has values between -1.308 and 0.891, which are extremely high, given that the data in both methods is the same. Accordingly, there is a big difference between the voxel-wise concatenation method and the averaging method.

The matrix describing the difference between the region-wise concatenation method and the averaging method is calculated as

$$FC_{MS}^{ave-concat(region)} = \left[\frac{1}{2N} \sum_{i,j} (fc_{i,j,r,r'}^{ave} - fc_{i,j,r,r'}^{concat(region)}) \right]_{r,r'} \tag{16}$$

and shown in Figure 8 (top-right) with notation from section 2.3. Values of this matrix range from -0.077 to 0.062 and are much smaller than values of $FC_{MS}^{ave-concat(voxel)}$. Therefore, the region-wise concatenation method and the aver-

[4]Note, that this FC-difference matrix is based on a comparison of different coping strategies, in contrast to the FC-difference matrix in section 3.2 which is based on the comparison of RL and LR runs.

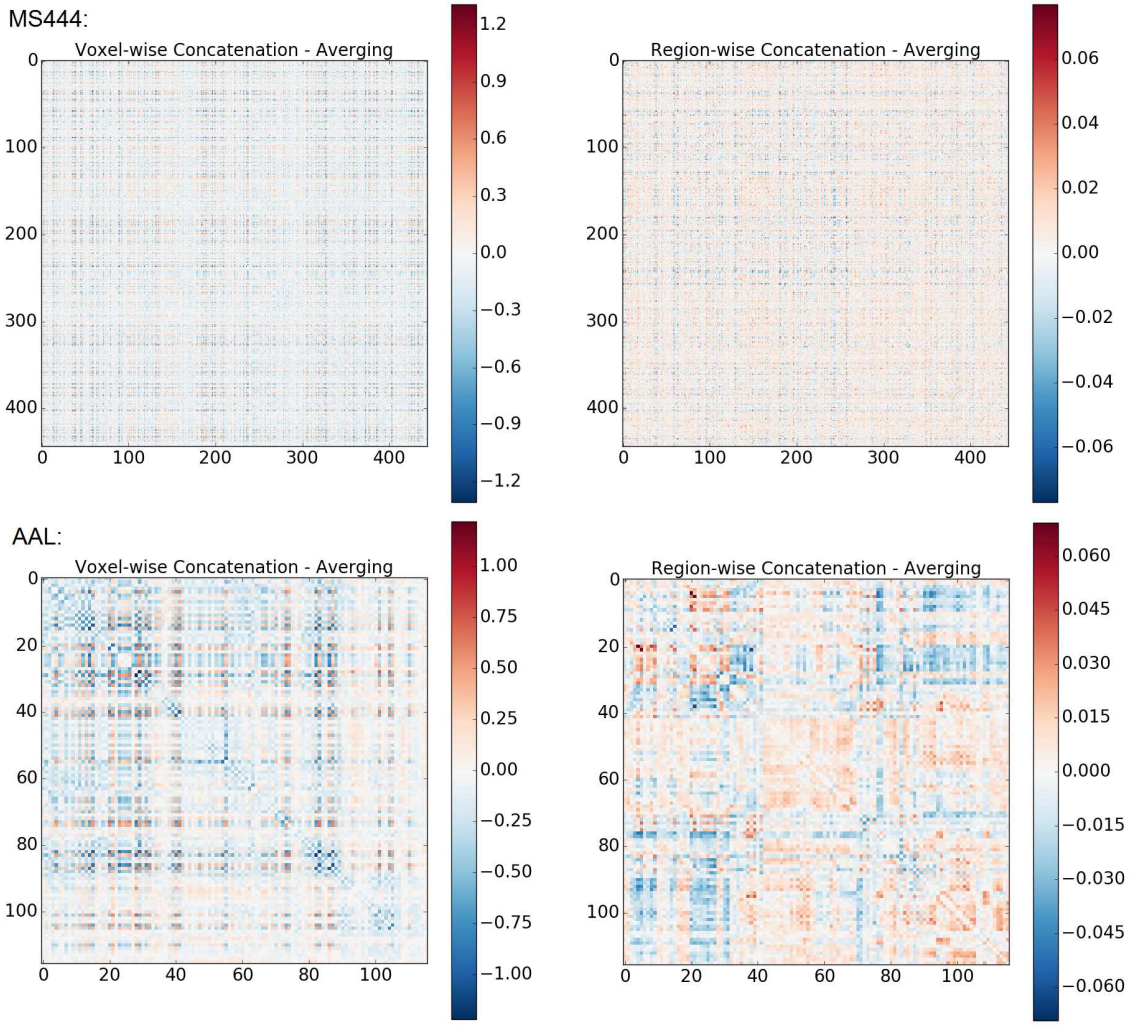

Figure 8: FC-difference matrices across different coping strategies. Top-left: $FC_{MS}^{ave-concat(voxel)}$ (see equation 15). Top-right: $FC_{MS}^{ave-concat(region)}$ (see equation 16). Bottom-left: $FC_{AAL}^{ave-concat(voxel)}$ (analogue to equation 15). Bottom-right: $FC_{AAL}^{ave-concat(region)}$ (analogue to equation 16).

aging method appear to be more similar to each other compared to the voxel-wise concatenation method. Nevertheless, there are difference between these (more similar) methods of up to 0.077, which are not negligible, given the context of FC-values, which are correlations and range between -1 and $+1$.

To investigate how these differences in methods are related to signal dropouts, the method-difference measure, as described in section 2.3, is evaluated for their relationship to the signal loss measure.

With respect to voxel-wise concatenation, this relationship (see green dots in Figure 9, top) was found to be non-linear, as method difference values seem to be bounded from above. This is plausible, because the method difference measure is constituted of differences of correlation values. In order to cope with this upper limit of method difference values, a logistic relationship is assumed.

In addition, considerations about dependencies within the method-difference measures (due to the way these measures are calculated) are analogue to those concerning the FC-impairment measure, so a correlation structure within a weighted least squares regression is specified by the positive-definite matrix which is closest [27] to $FC_{MS}^{ave-concat(voxel)}$ (see Figure 8, top-left).

To implement the logistic relationship into the context of the weighted least squares regression, signal loss l_r is transformed into a logistic signal loss measure

$$l_r^* = \frac{a}{a + exp(-b(l_r - c))} \tag{17}$$

where a, b and c are parameters for the maximum value, the slope and the point of inflection, respectively, which were estimated as $\hat{a} = 0.265$, $\hat{b} = 48.43$ and $\hat{c} = 0.019$ to model the logistic relationship between signal loss and the method difference using (non-linear) least squares. This logistic curve is shown in Figure 9 (top: black line).

With this transformed signal loss measures, the model under investigation is

$$d_r = \beta_{d,0} + l_r^* \beta_{d,l^*} + \epsilon_{d,r} \tag{18}$$

with $\epsilon_{d,r}$ being the normal distributed error with mean 0, $cor(d_r, d_{r'})$ the positive-definite matrix, which is closest [27] to $FC_{MS}^{ave-concat(voxel)}$ (see Figure 8, top-left), and variance weights being induced by the exponential variance structure $Var(\epsilon_{d,r}) = \sigma^2 exp(2\delta_d l_r^*)$ [31], in which δ_d is estimated to be $\hat{\delta_d} = 8.49$. The latter accounts for heteroscedasticity in the data, which can be seen in Figure 9 (middle). β-estimates are listed in Table 5 (top) and β_{d,l^*} is found to be significant. Accord-

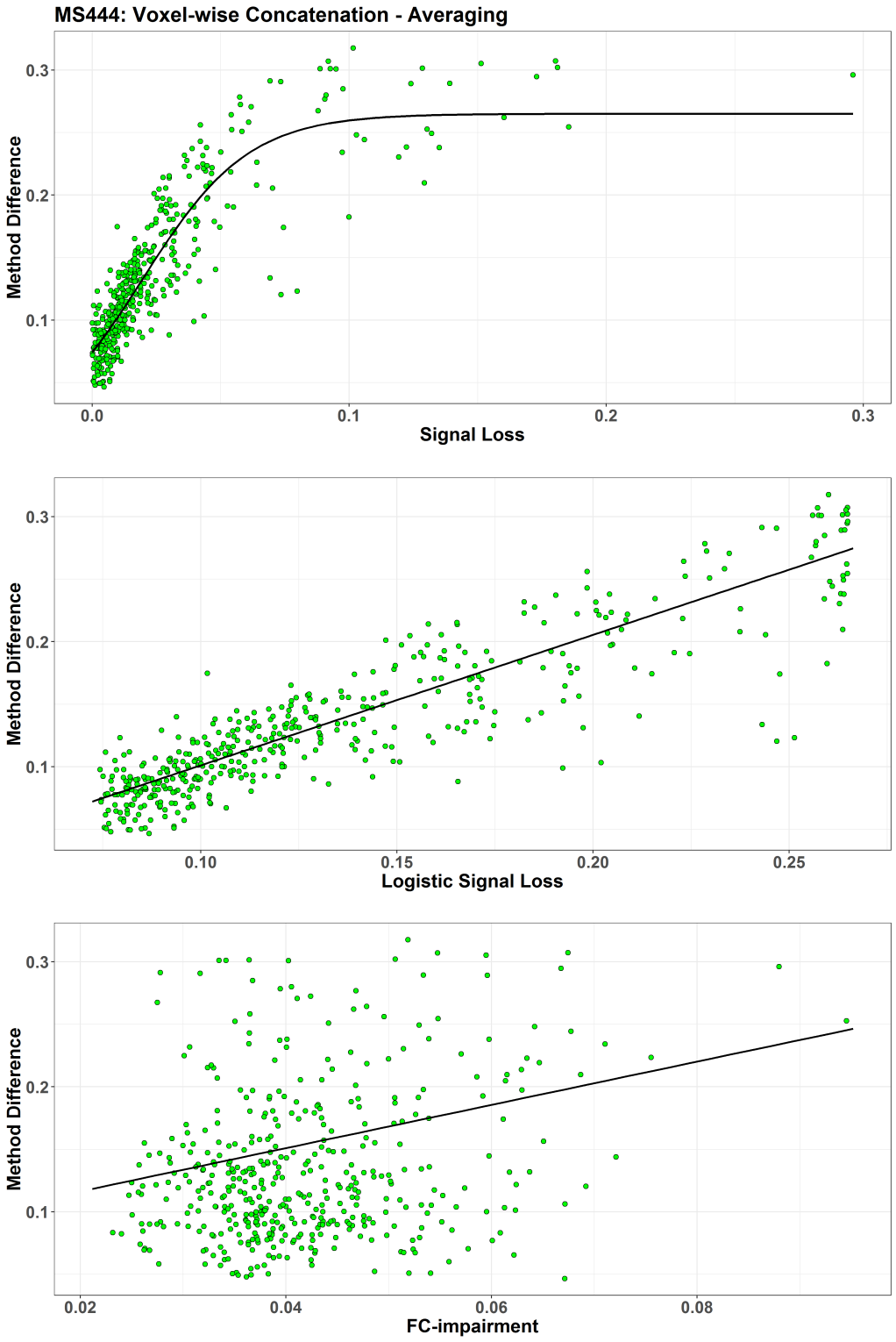

Figure 9: Regression on method differences (voxel) in the MS444 atlas. Scatterplots show the relationship of signal loss (x-axis, top), logistic signal loss (x-axis, middle, see equation 17) and FC-impairment (x-axis, bottom) to the method difference (voxel) measure (y-axis), respectively. Black lines display the estimated logistic curve with parameters $\hat{a} = 0.265$, $\hat{b} = 48.43$, $\hat{c} = 0.019$ (top), the prediction of the model in equation 18 (middle) and the prediction of the model in equation 19 (bottom).

ingly, the difference between the voxel-wise concatenation method and the averaging method seems to be influenced by signal dropouts.

Table 5: MS444 method comparison results.

Voxel-wise Concatenation - Averaging				
Coefficient	Value	Standard Error	t-value	p-value
$\beta_{d,0}$	-0.003	0.003	-0.90	0.370
β_{d,l^*}	1.043	0.030	34.43	< 0.0001
$\beta_{d2,0}$	0.082	0.011	7.50	< 0.0001
$\beta_{d2,f}$	1.730	0.243	7.13	< 0.0001
Region-wise Concatenation - Averaging				
Coefficient	Value	Standard Error	t-value	p-value
$\beta_{d,0}$	0.0085	0.0001	57.09	< 0.0001
$\beta_{d,l}$	0.0170	0.0029	5.95	< 0.0001
$\beta_{d2,0}$	0.0045	0.0004	10.76	< 0.0001
$\beta_{d2,f}$	0.1027	0.0093	11.06	< 0.0001

Furthermore, it is of interest, which of these methods is more strongly affected by signal dropouts. This can be assessed by the association between the FC-impairment measure and the method-difference measure: Given an association of the signal loss measure to both FC-impairment and method difference, no (or just a weak) correlation of the latter measures among each other, point to a stronger impact of the signal dropouts in the (voxel-wise) concatenation method (for a more detailed formulation of this reasoning see Appendix D).

This association between FC-impairment and method difference does not seem to be heteroscedastic (see Figure 9, bottom), so its model is

$$d_r = \beta_{d2,0} + f_r \beta_{d2,f} + \epsilon_{d2,r} \tag{19}$$

with $\epsilon_{d,r}$ being the normal distributed error with mean 0, the correlation matrix as above and without specifying variance weights. Both β-estimates are significant (see Table 5, top).

However, this association does not seem to be as strong as the one between signal loss and the method difference measure (compare variation in Figure 9). In order to investigate the explained variance in these three models with each other, adjusted R^2 values are calculated between the predicted and the actual dependent values: $R^2_{f,l} = 0.63$ in the association of signal loss with FC-impairment (using model from equation 12), $R^2_{d,l} = 0.80$ in the association of signal loss with the method difference measure and $R^2_{d,f} = 0.06$ in the association of FC-impairment with the

method difference measure. Certainly, the three underlying models contain a different number of parameters due to the non-linear modelling of the relationship between signal loss and method difference, so these values need to be interpreted with caution, especially as issues with R^2 values in the context of weigthe least squares were raised [32]. Nevertheless, without interpreting the exact value of these R^2 estimates, it can be said that the difference between $R^2_{d,f}$ to both $R^2_{d,l}$ and $R^2_{f,l}$ is very high, so it is reasonable to assume a low coefficient of determination in the former case and a high coefficient in the latter cases, even in a cautious interpretation.

Altogether, it seems that there is a difference between the voxel-wise concatenation method and the averaging method. This difference is related to signal dropouts, but can hardly be explained by the FC-impairment measure. As the latter measure is associated with the impairment of FC in the averaging method, it is plausible to assume another impact of signal dropouts in the voxel-wise concatenation method (see Appendix D).

These analyses are repeated for the comparison between the region-wise concatenation method and the averaging method. The relationship between signal loss and method difference measure (see Figure 10, top) does not seem to be non-linear as in the previous case, so the logistic transformation is not necessary. However, heteroscedasticity seems to be present, so variance weights are specified, resulting in the model

$$d_r = \beta_{d,0} + l_r\beta_{d,l} + \epsilon_{d,r} \tag{20}$$

with $\epsilon_{d,r}$ being the normal distributed error with mean 0, $cor(d_r, d_{r'})$ being the positive-definite matrix as above and variance weights induced by $Var(\epsilon_{d,r}) = \sigma^2 exp(2\delta_d l_r)$, in which δ_d is estimated as $\hat{\delta_d} = 1.034$. Concerning the association between FC-impairment and method differences, the same model as above was used (see Figure 10, bottom). Results of both weighted least squares regressions are listed in Table 5 (bottom) and all effects are significant.

Considering the strength of these associations in an analogous manner, adjusted R^2 values are $R^2_{d,l} = 0.09$ and $R^2_{d,f} = 0.24$ ($R^2_{f,l} = 0.63$ is the same as above). As $R^2_{d,l}$ is quite small (and it might be an optimistic estimate, see [32]), it cannot be assumed that the difference between the region-wise concatenation method and the averaging method is constituted by signal dropouts. Therefore, although these methods differ to some extent, with respect to signal dropouts there is no clear guidance, which method to prefer.

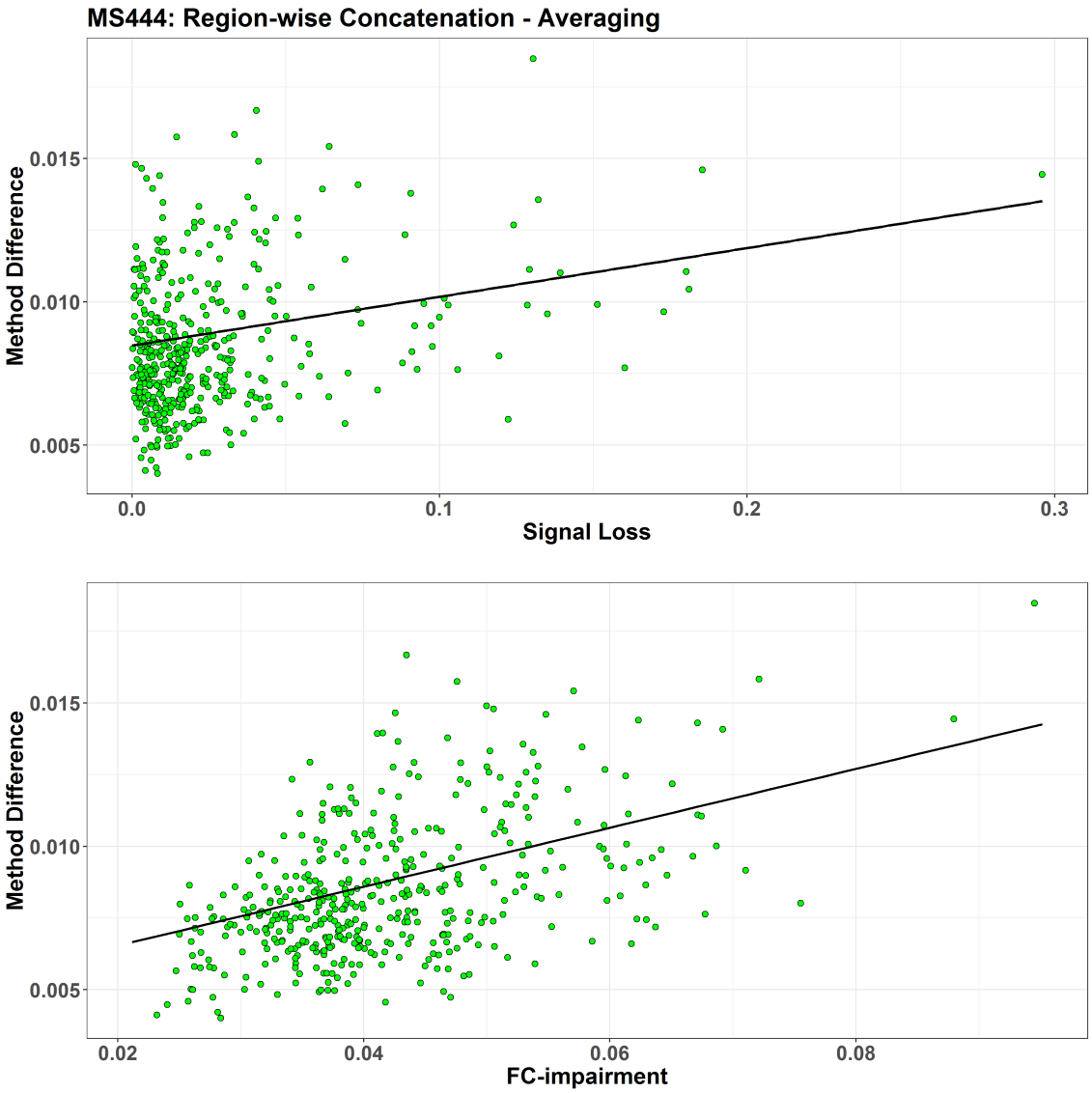

Figure 10: Regression on method differences (region) in the MS444 atlas. Scatterplots show the relationship of signal loss (x-axis, top) and FC-impairment (x-axis, bottom) to the method difference (region) measure (y-axis), respectively. Black lines display the prediction of the model in equation 20 (top) and the prediction of the weighted least squares regression of FC-impairment on the difference (region) measure which is analogue to equation 18 (bottom).

AAL. In order to evaluate the generalisability of the results in the previous section based on the MS444 atlas to other parcellations, analogue analyses are performed with the AAL parcellation. FC-difference matrices $FC_{AAL}^{ave-concat(voxel)}$ (see Figure 8, bottom-left) and $FC_{AAL}^{ave-concat(voxel)}$ (see Figure 8, bottom-right) are now 116 x 116 matrices and range from -1.220 to 0.711 and from -0.061 to 0.069, respectively. This is similar to the MS444 atlas and states that there is a big difference of the averaging method to the voxel-wise concatenation method and a small difference of the averaging method to the region-wise concatenation method.

The relationship between signal loss and the method difference measure $d_{r,voxel}$ is non-linear (see Figure 11, top) and modelled by a logistic curve (see equation 17) with estimated parameters $\hat{a} = 0.288$, $\hat{b} = 67.38$ and $\hat{c} = 0.016$. Applying the model in equation 18 in this context for the association between the transformed signal loss measure and the method difference measure (see Figure 11, middle) yields an estimate of $\hat{\delta_d} = 4.42$ for the variance parameter and β-estimates as listed in Table 6 (top), of which β_{d,l^*} is significant. Therefore, signal loss seems to influence the difference between the voxel-wise concatenation method and the averaging method in the AAL parcellation as well. The association between FC-impairment and the method difference measure (using the model in equation 19, see Figure 11, bottom) is significant (see Table 6, top). Respective adjusted R^2 values are $R^2_{f,l} = 0.71$, $R^2_{d,l} = 0.86$ and $R^2_{d,f} = 0.03$, expressing a similar pattern as in the MS444 analysis. Hence, an additional influence of signal dropouts in the voxel-wise concatenation method can be expected in the AAL parcellation as well.

Comparing the region-wise concatenation method to the averaging method, the influence of signal loss on their difference (see Figure 12, top) is modelled by equation 20, with $\hat{\delta_d} = -0.712$ and β-estimates are significant (see Table 6, bottom). Adjusted R^2 values are $R^2_{d,l} = 0.12$ and $R^2_{d,f} = 0.12$. Again, it is not reasonable to assume that signal dropouts constitute the difference between the region-wise concatenation method and the averaging method.

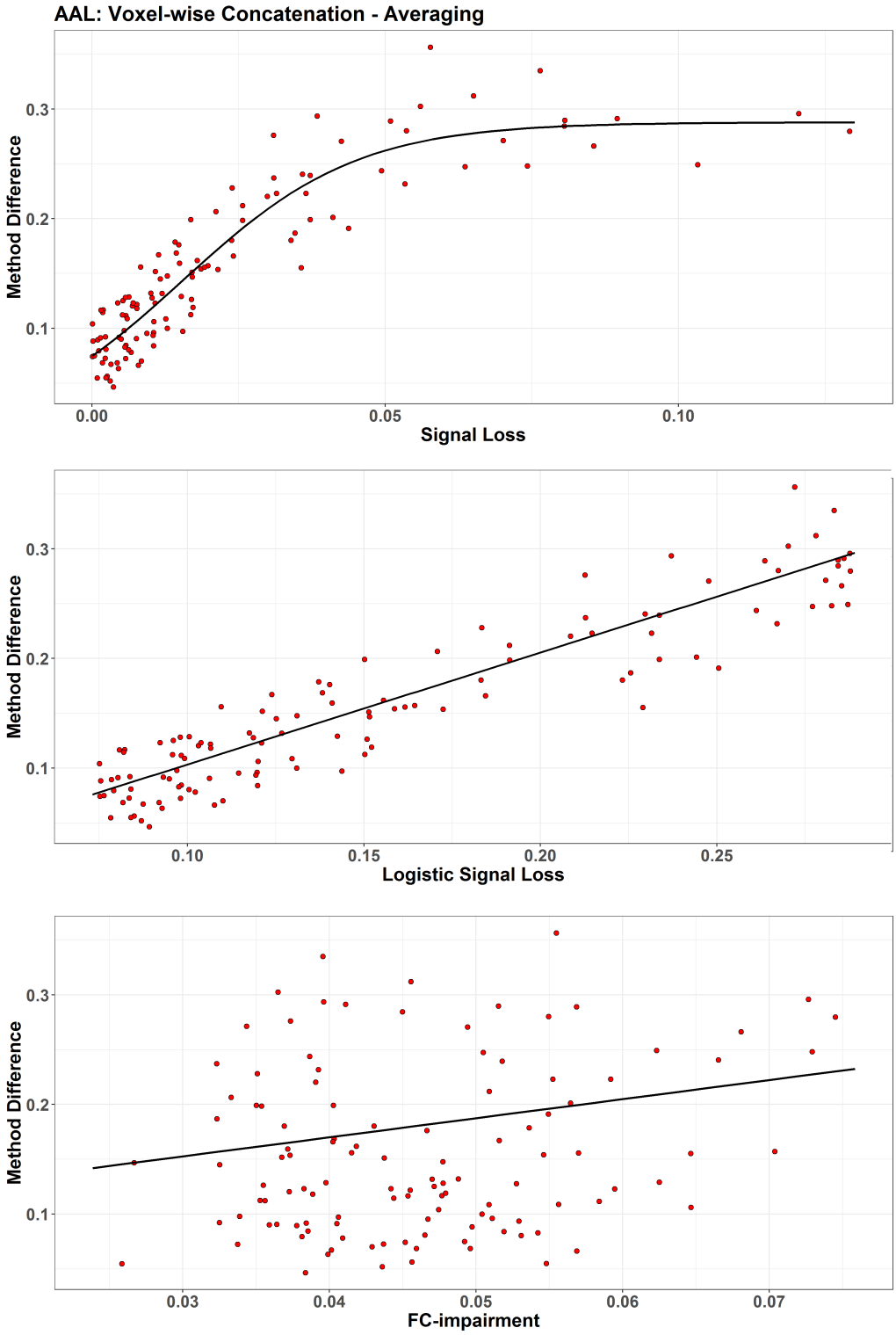

Figure 11: Regression on method differences (voxel) in the AAL atlas. Scatterplots show the relationship of signal loss (x-axis, top), logistic signal loss (x-axis, middle, see equation 17) and FC-impairment (x-axis, bottom) to the method difference (voxel) measure (y-axis), respectively. Black lines display the estimated logistic curve with parameters $\hat{a} = 0.288$, $\hat{b} = 67.38$, $\hat{c} = 0.016$ (top) and the predictions of the respective weighted least squares regressions (middle: analogue to equation 18; bottom: analogue to equation 19).

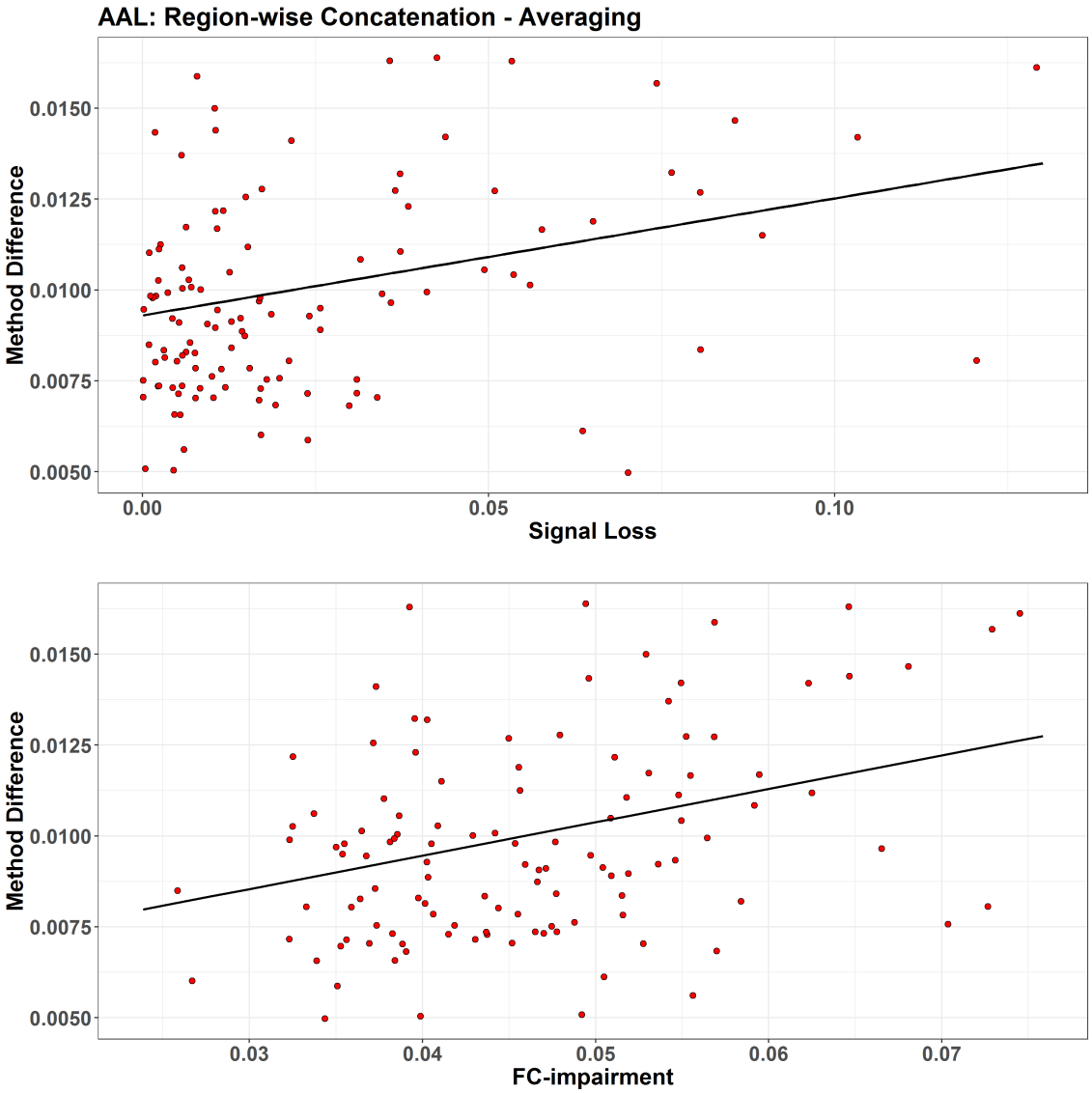

Figure 12: Regression on method differences (region) in the AAL atlas. Scatterplots show the relationship of signal loss (x-axis, top) and FC-impairment (x-axis, bottom) to the method difference (region) measure (y-axis), respectively. Black lines display the predictions of the respective weighted least squares regressions (top: analogue to equation 20; bottom: analogue to equation 18).

Table 6: AAL method comparison results.

Voxel-wise Concatenation - Averaging				
Coefficient	Value	Standard Error	t-value	p-value
$\beta_{d,0}$	0.001	0.006	0.19	0.852
β_{d,l^*}	1.021	0.043	23.50	< 0.0001
$\beta_{d2,0}$	0.100	0.032	3.16	0.0020
$\beta_{d2,f}$	1.741	0.661	2.63	0.0096
Region-wise Concatenation - Averaging				
Coefficient	Value	Standard Error	t-value	p-value
$\beta_{d,0}$	0.0093	0.0003	28.55	< 0.0001
$\beta_{d,l}$	0.0322	0.0090	3.59	0.0005
$\beta_{d2,0}$	0.0058	0.0011	5.42	< 0.0001
$\beta_{d2,f}$	0.0918	0.0224	4.10	0.0001

4. Discussion

4.1. Summary of Results

The fMRI acquisition protocol in the HCP utilizes an eight-factorial multi-band EPI sequence and a phase encoding direction along the left-to-right (or right-to-left) axis. These settings are accompanied by signal dropouts at specific locations within the brain (see Figure 1). HCP rs-fMRI data comprises runs with RL phase encoding and runs with LR phase encoding, which have different signal dropout locations. By comparing FC matrices between these runs, brain regions with an abnormal high systematic difference in FC values can be found. Apparently, these regions suffer the strongest impairment in FC between different phase encoding directions and are located in the ventral parts of the frontal cortex as well as in the left anterior temporal cortex (see Figures 3 and 4).

The (region-specific) measure of how strong a regions FC is altered across different phase encoding directions (equation 1) is significantly higher (see Tables 3 and 4, $\beta_{l,m}$) than a control measure, which is calculated across different acquisition sessions using the same phase encoding direction (equation 7). This suggests that - considering all brain regions in general - the change in phase encoding direction is characterized by a systematic change in FC-values, which does not resemble pure noise.

Concerning Multiscale 444 brain regions, a positive association between the strength of signal dropouts and the impairment in FC values was found, even though regions with the strongest FC-impairment were excluded. Using the AAL atlas, this association was only observable, if all 116 regions were examined.

As a method of dealing with these signal dropouts, the HCP suggested to either concatenate time series or average FC values of both runs [6]. With respect to the former, a voxel-wise concatenation before preprocessing and a region-wise concatenation after preprocessing of time-series was investigated in this thesis. A comparison between the averaging method and the voxel-wise concatenation method reveals strong differences in FC-values, which are highly associated with signal dropouts, but hardly related to the FC-impairment measure. Contrasting the averaging method with the region-wise concatenation method, only small differences are found, being barely associated with signal dropouts.

4.2. Guidelines

Based on these results, guidelines about handling signal dropouts in the HCP can be inferred.

In order to deal with signal dropouts, the voxel-wise concatenation method seems to be highly affected by those and should not be employed. Although a difference between the region-wise concatenation method and the averaging method exists, this difference is hardly related to signal loss, so either method might be utilized with respect to the dropouts.

For each brain connection there are two different FC values: one in the RL run and one in the LR run. Under the assumption that FC can only be impaired and not improved by signal dropouts, the higher of these values is closer to the true FC value of this connection. Therefore, by averaging LR and RL FC matrices (in the averaging method), the difference of the resulting FC value from the true FC value is at least half the difference of the RL and LR FC value. So the minimum FC-impairment in the averaging FC-matrix is $\frac{1}{2}|FC_a^{RL-LR}|$ (see Figure 5, left). As there is hardly any difference to the region-wise concatenation matrix, which is related to signal dropouts, a minimum FC-impairment of similar magnitude can be assumed for the region-wise concatenation FC-matrix.

However, if no coping method is applied and FC is calculated on a single (RL or LR) run, the impairment in FC is supposed to have a different structure: For some connections FC is less impaired and for other connections FC is more strongly impaired in the chosen run compared to the run with the other phase encoding direction. The minimal FC-impairment of the former is the FC-difference value and of the latter it is 0.

Therefore, by utilizing coping strategies the influence of FC-impairment is supposed to be more homogeneous across all connections, than by considering only a single run.

As a consequence of these findings, it is advised to be cautious in dealing with results based on rs FC in the HCP, especially, if regions with a high FC-impairment (listed in Table 1 for MS444 and Table 2 for AAL) are involved in generating these results.

In case strategies to cope with signal dropouts are utilized, the AAL atlas seems to exhibit a FC-impairment pattern for non-outlier regions, which is not associated with signal dropouts. FC-impairment in the MS444 parcellation, however, appears to be more prone to signal dropouts, so it is recommended to take dropout based signal loss into account when dealing with FC in the MS444 atlas. Furthermore, it

seems appropriate to do this for all brain parcellations being more detailed, or in all analyses, which involve FC based on time-series of small areas.

If no coping strategies are employed and only one single fMRI run is analysed, a considerable amount of heterogeneity of FC-impairment across regions is present. Consequently, a closer look at the regions being most relevant for results and a robustness evaluation using a run with the other phase encoding direction is advised in general (not only in detailed parcellations).

4.3. Studies being potentially affected by HCP Signal Dropouts

Among studies based on the HCP rs-fMRI data, two different types of impairment due to signal dropouts could found:

First, by defining regions or components based on FC or rs time-series, the spatial extent of these regions might not be accurate in dropout areas. Therefore, further analyses performed with these newly defined regions could provide results, being erroneous in dropout areas. In an investigation within the HCP rs-fMRI data, these results are likely to contain inaccuracies in the signal dropout areas anyway, however, if HCP rs-fMRI (FC) based atlases are applied in other data sets, results might be influenced by HCP signal dropouts. Examples of theses atlases are found in [33] and [34], which should be treated with caution, especially, as each study considers only one single run, i.e. one single phase encoding direction.

Second, FC results in the HCP rs-fRMI data might not be found equally well in all areas of the brain.

A study in the field of dynamic FC by [35] focussed on the temporal variability of the local functional connectivity density (lFCD) and reports a pronounced lFCD metric within most cortical regions, except for inferior ventral, orbito-frontal, anterior temporal and insular cortices. As explanation the authors attribute this to a reduced coil sensitivity of the utilized 32-channel head coil in deep brain regions and refer to [36]. This investigation, however, reports benefits of the 32-channel coil over a 12-channel coil, and mentions limited spatial resolutions and coil sensitivity only with respect to subcortical networks. As the regions of attenuated lFCD reported by [35] show a strong overlap with dropout regions (see Figure 1) as well as with other (mainly temporal) regions showing a high FC-impairment in this thesis (see Table 1 and 2), signal dropouts in the HCP would serve as a more plausible explanation for these lower lFDC values than a reduced coil sensitivity.

A FC examination of the habenula [37] based on HCP rs-fMRI data of concatenated

time-series (after main preprocessing, before temporal filtering) reports functional connections of the habenula to other brain areas, which are associated with subclinical depression. By using a seed-based approach, target areas to which FC was calculated are small and might be affected by signal dropouts. Accordingly, the possibility of missed connections to areas being affected by signal dropout cannot be excluded.

In addition to these types of impairment in HCP based rs-fMRI results, special caution should be applied to those studies, which utilize the concatenation method before preprocessing, as this might induces strong artifacts (see Figure 8, left). This might apply e.g. to [38] and possibly to [39]. In the latter study, no clear statement was given of when the concatenation step was applied during preprocessing.

4.4. Limitations

The ground truth of FC values is not known. Accordingly, the true impairment in FC due to signal dropouts cannot be assessed and results of this thesis represent only estimates of FC-impairment, which arise in the comparison between LR and RL runs. As can be seen in Figure 1 (from [6]), areas being affected by signal dropouts in one run, might contain signal loss in a run with the other phase encoding direction as well (see e.g. the orbitofrontal cortex). Therefore, it seems plausible that the actual strength of FC-impairment due to signal dropouts is stronger than indicated by estimates in this work.

On the one hand, the amount of FC-impairment of a region depends on its involvement in the brain connectivity. A region being hardly functionally connected to the rest of the brain will express a low value in the FC-impairment measure independent of its signal loss, which could be an explanation for the presence of brain regions with a high signal loss and a low FC-impairment value (see Figure 6, dots in the bottom-right of each plot). On the other hand, a region being strongly connected to the rest of the brain, might express a high FC-impairment even if the signal loss is small. Accordingly, FC-impairment due to signal dropouts might not be restricted to areas with a high signal loss value.

Furthermore, in addition to the signal dropout areas considered in this thesis, there are other areas with phase encoding dependent signal strengths (e.g. in the occipital region) [6], whose distortions were corrected by the HCP. As FC values seem to be impaired by signal dropouts, it is questionable, if these corrections are able to restore the FC structure within these regions properly. If not, further impairments in FC exist in regions, which could not be observed in this work.

The HCP decided to utilize a multi-band factor of $M = 8$ to increase acquisition

speed [6], which has a profound effect on the detection sensitivity of resting state networks [40, 41]. Altogether, this benefit in temporal resolution might be worth taking signal dropouts into account and for most parts of the brain these benefits could be crucial in generating novel insides, as e.g. with the detection of dynamic FC patterns [42]. With respect to the dropout regions, however, results might be erroneous and caution is advised.

In this investigation, only the impairment in FC values was considered. Certainly, the impact of signal dropouts on other rs measures, such as amplitude of low frequency fluctuations (ALFF), regional homogeneity (ReHo) (see [7]) or graph specific measures [43], in the HCP rs-fMRI data is not assessed and needs to be addressed in subsequent analyses.

4.5. Conclusion

Signal dropouts in the HCP rs-fMRI data due to the use of a multi-band EPI sequence, whose locations are dependent on the phase encoding direction, seem to affect FC values. Some brain regions appear to be more strongly affected be these dropouts and show a higher degree of FC impairment than others. These regions are predominantly located bilaterally in the orbito-frontal and unilaterally in the anterior parts of the left temporal cortex. This issue should be considered by all researchers working with the HCP rs-fMRI data. In order to reduce the spatial heterogeneity of the influence of these dropouts on FC results, either the averaging or the concatenation method could be applied. With respect to the latter, however, it is crucial to implement the concatenation of time-series after and not before (or during) preprocessing of the data.

References

1. Van Essen, D. C. *et al.* The WU-Minn human connectome project: an overview. *Neuroimage* **80,** 62–79 (2013).
2. Glasser, M. F. *et al.* The human connectome project's neuroimaging approach. *Nature Neuroscience* **19,** 1175–1187 (2016).
3. Ogawa, S., Lee, T.-M., Kay, A. R. & Tank, D. W. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences* **87,** 9868–9872 (1990).
4. Logothetis, N. K., Pauls, J., Augath, M., Trinath, T. & Oeltermann, A. Neurophysiological investigation of the basis of the fMRI signal. *Nature* **412,** 150–157 (2001).
5. Huettel, S. A., Song, A. W. & McCarthy, G. *Functional magnetic resonance imaging* (Sinauer Associates Sunderland, 2004).
6. Smith, S. M. *et al.* Resting-state fMRI in the human connectome project. *Neuroimage* **80,** 144–168 (2013).
7. Bijsterbosch, J., Smith, S. M. & Beckmann, C. F. *An Introduction to Resting State fMRI Functional Connectivity* (Oxford University Press, 2017).
8. Feinberg, D. A. & Setsompop, K. Ultra-fast MRI of the human brain with simultaneous multi-slice imaging. *Journal of magnetic resonance* **229,** 90–100 (2013).
9. Uğurbil, K. *et al.* Pushing spatial and temporal resolution for functional and diffusion MRI in the Human Connectome Project. *Neuroimage* **80,** 80–104 (2013).
10. Abraham, A. *et al.* Machine learning for neuroimaging with scikit-learn. *arXiv preprint arXiv:1412.3919* (2014).
11. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* **12,** 2825–2830 (2011).
12. Glasser, M. F. *et al.* The minimal preprocessing pipelines for the Human Connectome Project. *Neuroimage* **80,** 105–124 (2013).
13. Lindquist, M. A. The statistical analysis of fMRI data. *Statistical Science,* 439–464 (2008).
14. Bellec, P., Rosa-Neto, P., Lyttelton, O. C., Benali, H. & Evans, A. C. Multi-level bootstrap analysis of stable clusters in resting-state fMRI. *Neuroimage* **51,** 1126–1139 (2010).

15. Tzourio-Mazoyer, N. *et al.* Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* **15,** 273–289 (2002).

16. Friston, K. J., Williams, S., Howard, R., Frackowiak, R. S. & Turner, R. Movement-related effects in fMRI time-series. *Magnetic resonance in medicine* **35,** 346–355 (1996).

17. Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L. & Petersen, S. E. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage* **59,** 2142–2154 (2012).

18. Power, J. D. *et al.* Methods to detect, characterize, and remove motion artifact in resting state fMRI. *Neuroimage* **84,** 320–341 (2014).

19. Murphy, K., Birn, R. M. & Bandettini, P. A. Resting-state fMRI confounds and cleanup. *Neuroimage* **80,** 349–359 (2013).

20. Murphy, K., Birn, R. M., Handwerker, D. A., Jones, T. B. & Bandettini, P. A. The impact of global signal regression on resting state correlations: are anti-correlated networks introduced? *Neuroimage* **44,** 893–905 (2009).

21. Anderson, J. S. *et al.* Network anticorrelations, global regression, and phase-shifted soft tissue correction. *Human brain mapping* **32,** 919–934 (2011).

22. Jo, H. J., Saad, Z. S., Simmons, W. K., Milbury, L. A. & Cox, R. W. Mapping sources of correlation in resting state FMRI, with artifact detection and removal. *Neuroimage* **52,** 571–582 (2010).

23. He, B. J., Snyder, A. Z., Zempel, J. M., Smyth, M. D. & Raichle, M. E. Electrophysiological correlates of the brain's intrinsic large-scale functional architecture. *Proceedings of the National Academy of Sciences* **105,** 16039–16044 (2008).

24. Schölvinck, M. L., Maier, A., Frank, Q. Y., Duyn, J. H. & Leopold, D. A. Neural basis of global resting-state fMRI activity. *Proceedings of the National Academy of Sciences* **107,** 10238–10243 (2010).

25. Hallquist, M. N., Hwang, K. & Luna, B. The nuisance of nuisance regression: spectral misspecification in a common approach to resting-state fMRI preprocessing reintroduces noise and obscures functional connectivity. *Neuroimage* **82,** 208–225 (2013).

26. R Core Team. *R: A Language and Environment for Statistical Computing* R Foundation for Statistical Computing (Vienna, Austria, 2016). <`https://www.R-project.org/`>.

27. Higham, N. J. Computing a nearest symmetric positive semidefinite matrix. *Linear algebra and its applications* **103,** 103–118 (1988).

28. Tomasi, D. & Volkow, N. D. Laterality patterns of brain functional connectivity: gender effects. *Cerebral Cortex* **22,** 1455–1462 (2012).

29. Walfish, S. A review of statistical outlier methods. *Pharmaceutical technology* **30,** 82 (2006).

30. Benoit, K. Linear regression models with logarithmic transformations. *London School of Economics, London* **22,** 23–36 (2011).

31. Carey, V. & Wang, Y.-G. *Mixed-effects models in S and S-PLUS* 2001.

32. Willett, J. B. & Singer, J. D. Another cautionary note about R 2: Its use in weighted least-squares regression analysis. *The American Statistician* **42,** 236–238 (1988).

33. Lv, J. *et al.* Holistic atlases of functional networks and interactions reveal reciprocal organizational architecture of cortical function. *IEEE Transactions on Biomedical Engineering* **62,** 1120–1131 (2015).

34. Fan, L. *et al.* The human brainnetome atlas: a new brain atlas based on connectional architecture. *Cerebral Cortex* **26,** 3508–3526 (2016).

35. Tomasi, D, Shokri-Kojori, E & Volkow, N. Temporal changes in local functional connectivity density reflect the temporal variability of the amplitude of low frequency fluctuations in gray matter. *PloS one* **11,** e0154407 (2016).

36. Anteraper, S. A. *et al.* Exploring functional connectivity networks with multi-channel brain array coils. *Brain connectivity* **3,** 302–315 (2013).

37. Ely, B. A. *et al.* Resting-state functional connectivity of the human habenula in healthy individuals: Associations with subclinical depression. *Human brain mapping* **37,** 2369–2384 (2016).

38. Bolt, T., Nomi, J. S., Rubinov, M. & Uddin, L. Q. Correspondence between evoked and intrinsic functional brain network configurations. *Human brain mapping* **38,** 1992–2007 (2017).

39. Park, B.-y. & Park, H. Connectivity differences between adult male and female patients with attention deficit hyperactivity disorder according to resting-state functional MRI. *Neural regeneration research* **11,** 119 (2016).

40. Boubela, R. N., Kalcher, K., Nasel, C. & Moser, E. Scanning fast and slow: current limitations of 3 tesla functional MRI and future potential. *Frontiers in physics* **2** (2014).

41. Feinberg, D. A. *et al.* Multiplexed echo planar imaging for sub-second whole brain FMRI and fast diffusion imaging. *PloS one* **5,** e15710 (2010).

42. Zalesky, A., Fornito, A., Cocchi, L., Gollo, L. L. & Breakspear, M. Time-resolved resting-state brain networks. *Proceedings of the National Academy of Sciences* **111,** 10341–10346 (2014).

43. Bullmore, E. & Sporns, O. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience* **10,** 186–198 (2009).

List of Figures

1. Signal Dropouts . . . 3
2. Histograms of FC-impairment values: MS444 and AAL . . . 11
3. MS444 brain regions with strongest FC-impairment . . . 12
4. AAL brain regions with strongest FC-impairment . . . 14
5. FC-difference matrices . . . 15
6. Influence of signal loss on FC-impairment . . . 17
7. Model Fit: MS444 and AAL analyses . . . 18
8. FC-difference matrices across different coping strategies . . . 22
9. Regression on method differences (voxel) in the MS444 atlas . . . 24
10. Regression on method differences (region) in the MS444 atlas . . . 27
11. Regression on method differences (voxel) in the AAL atlas . . . 29
12. Regression on method differences (region) in the AAL atlas . . . 30
13. Model fit in preliminary assessment . . . 44
14. MS444-no-smooth: Influence of signal loss on FC-impairment . . . 45

List of Tables

1. MS444 brain regions with highest FC-impairment values . . . 10
2. AAL brain regions with highest FC-impairment values . . . 13
3. MS444 regression results . . . 16
4. AAL regression results . . . 20
5. MS444 method comparison results . . . 25
6. AAL method comparison results . . . 31
7. MS444-no-smooth regression results . . . 46

List of Abbreviations

AAL	Automated Anatomical Labeling
ALFF	Amplitude of Low Frequency Fluctuations
BOLD	Blood Oxygenation Level Dependent
CSF	Cerebro Spinal Fluid
EPI	Echo-Planar-Imaging
FC	Functional Connectivity
fMRI	functionl Magnetic Resonance Imaging
FWHM	Full Width at Half Maximum
HCP	Human Connectome Project
lFCD	local Functional Connectivity Density
LR	Left-to-Right
MNI	Montreal Neurological Institute
MS	Multiscale
ReHo	Regional Homogeneity
RF	Radio Frequency
RL	Right-to-Left
rs	resting-state
TR	Repetition Time
WM	White Matter

A. Subject IDs

Analyses in this thesis are based on 100 subjects of the HCP with the following IDs:

100307, 100408, 101006, 101107, 101309, 101915, 102008, 102311, 102816, 103111, 103414, 103515, 103818, 104820, 105014, 105115, 105216, 106016, 106319, 106521, 107321, 107422, 108121, 108323, 108525, 108828, 109123, 109325, 110411, 111312, 111413, 111716, 113215, 113619, 113922, 114419, 114924, 115320, 116524, 117122, 117324, 118528, 118730, 118932, 120111, 120212, 120515, 121618, 122317, 122620, 123117, 123420, 123925, 124220, 124422, 124826, 125525, 126325, 126628, 127630, 127933, 128127, 128632, 129028, 130013, 130316, 130922, 131217, 131722, 131924, 132118, 133019, 133625, 133827, 133928, 134324, 135225, 135528, 135932, 136227, 136833, 137027, 137128, 137633, 137936, 138231, 138534, 139233, 139637, 140117, 140824, 140925, 141422, 141826, 142828, 143325, 144226, 144832, 146432, 147030.

B. Preliminary Assessment

In a preliminary analysis data characteristics were evaluated with respect to their fit of the model. Using the terminology from section 3.2, a simple first model M_1 was

$$M_1 : f_{r,m} = \beta_0 + l_{r,m}\beta_l + m\beta_m + ml_{r,m}\beta_{l,m} + \epsilon_{r,m} \tag{21}$$

without specifying either correlation structure nor variance weights. A QQ-plot of the residuals (see Figure 13, top-left) reveals that their distribution is tailed and rather right-skewed. This behaviour can be stabilised by a log-transformation of the dependent variable. This was evaluated in a second preliminary log-linear model M_2 (without modelling correlation structure and variance weights)

$$M_2 : log(f_{r,m}) = \beta_0 + l_{r,m}\beta_l + m\beta_m + ml_{r,m}\beta_{l,m} + \epsilon_{r,m} \tag{22}$$

yielding a QQ-plot as depicted in Figure 13 (top-right), which shows a better model fit with respect to skewness. In addition, plotting residuals against fitted values in M_2 (see Figure 13, bottom) displays heteroscedasticity of the FC-impairment measure with respect to signal loss.

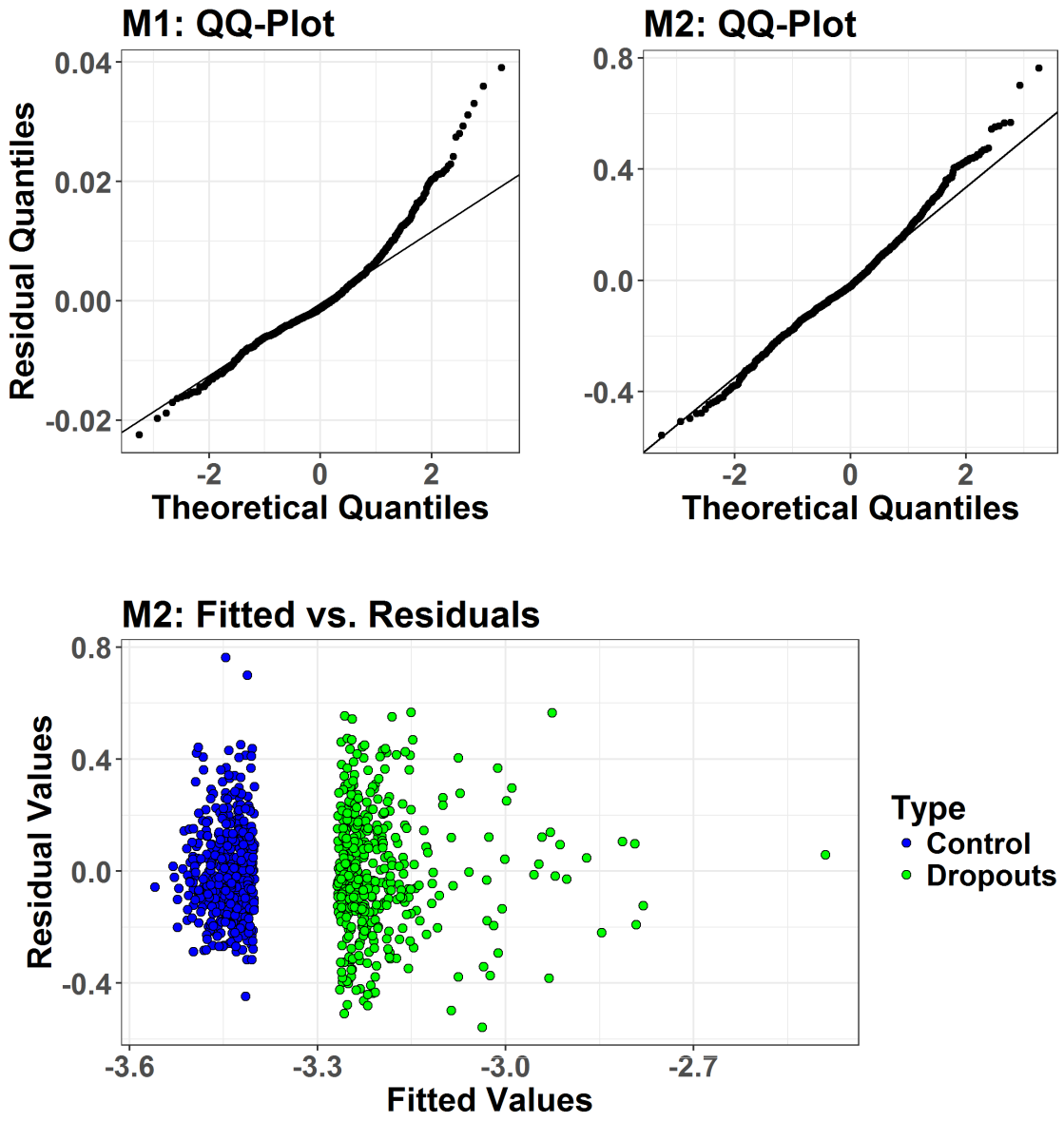

Figure 13: Model fit in preliminary assessment. Top-left: The QQ-plot of a first preliminary model M_1 (see equation 21) shows a right-skewness with a pronounced right-sided tail. Top-right: The QQ-plot of the second preliminary model M_2 (see equation 22) displays a less pronounced right-sided tail. Accordingly, the log-transformation of the dependent variable copes with the right-skewness in the data. Bottom: The "Fitted vs. Residuals"-plot of model M_2 reveals heteroscedasticity.

C. Investigation using the MS444-no-smooth preprocessing pipeline

The weighted least squares analyses as described in section 3.2 for MS444 data, are repeated on data without spatial smoothing during preprocessing (MS444-no-smooth pipeline). The relationship between signal loss and FC-impairment is depicted in Figure 14 and β-estimates are listed in Table 7. Again, all effects are significant, yielding a similar pattern of results, which expresses a robustness of these analyses against the choice of kernel width for spatial smoothing.

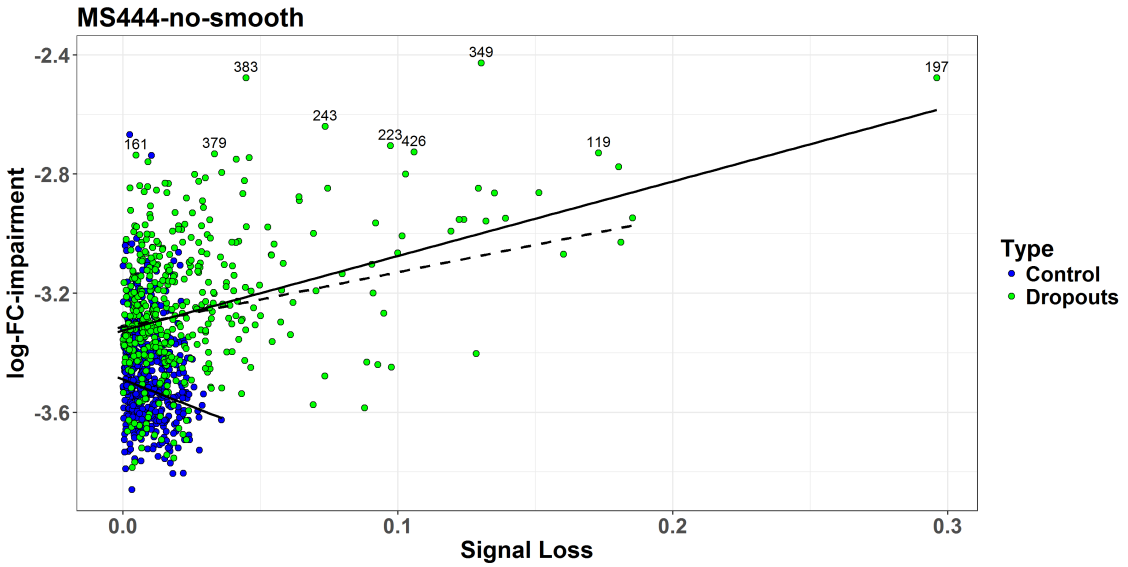

Figure 14: MS444-no-smooth: Influence of signal loss on FC-impairment. The scatterplot shows the relationship between signal loss (x-axis) and the log-transformation of FC-impairment (y-axis) in the MS444 parcellation using pipeline MS444-no-smooth. Black solid lines indicate the prediction according to the weighted least squares regression. Labelled points refer to the regions defined as outlier regions, which are 349, 197, 383, 379, 223, 243, 426, 119 and 161 using the MS444-no-smooth pipeline. The dashed line indicates the prediction of the weighted least squares regression, in which outlier regions (labelled) were excluded in the dropout-data.

Table 7: MS444-no-smooth regression results.

All Data				
Coefficient	Value	Standard Error	t-value	p-value
β_0	-3.49	0.017	-201.09	< 0.0001
β_l	-3.61	1.095	-3.30	0.001
β_m	0.16	0.024	6.90	< 0.0001
$\beta_{l,m}$	6.11	1.167	5.24	< 0.0001
β_0^d	-3.33	0.017	-194.44	< 0.0001
β_l^d	2.56	0.380	6.73	< 0.0001
Outliers Excluded				
Coefficient	Value	Standard Error	t-value	p-value
β_0	-3.49	0.017	-201.03	< 0.0001
β_l	-3.61	1.095	-3.29	0.001
β_m	0.18	0.024	7.48	< 0.0001
$\beta_{l,m}$	5.45	1.149	4.74	< 0.0001
β_0^d	-3.31	0.017	-198.41	< 0.0001
β_l^d	1.88	0.329	5.73	< 0.0001

D. Considerations on Coping Strategies

Apparently, there is a difference between the FC-matrices of the averaging method and the voxel-wise concatenation method that reflect signal dropouts (see Figures 9 and 11 in section 3.3). Regarding of how these methods are affected by dropouts, there are possible options:

1 *Averaging-method not impaired - Concatenation-method impaired.*

2 *Averaging-method impaired - Concatenation-method more strongly impaired.*

3 *Averaging-method impaired - Concatenation-method not impaired.*

4 *Averaging-method more strongly impaired - Concatenation-method impaired.*

The FC-matrix in the averaging-method is constituted by both RL FC-matrix and LR FC-matrix. Under the assumption that signal dropouts do not systematically improve FC strengths, for each connection (i.e. matrix entry) the higher absolute FC-value of both matrices can be expected to be (on average) closer to the true FC-value. Utilizing the averaging-method, its FC-values are further apart from the one FC-value of a separate run, which is considered to be closer to the true FC-value. For connection (r, r') this distance of the FC-values (between averaging-method and

the better of separate runs) is estimated as

$$\frac{1}{2}\left|\frac{1}{2N}\sum_{i,j}(fc^{RL}_{i,j,r,r'} - fc^{LR}_{i,j,r,r'})\right| \tag{23}$$

using the notation of section 2.3.

Accordingly, for quantifying how far the FC-values of a region r are apart from the best FC-values of the separate runs, the following measure can be used:

$$\frac{1}{2}\frac{1}{R-1}\sum_{r',r'\neq r}\left|\frac{1}{2N}\sum_{i,j}(fc^{RL}_{i,j,r,r'} - fc^{LR}_{i,j,r,r'})\right| \tag{24}$$

Except for the factor 1/2, these values are the same as the measure of FC-impairment. As these values differ from those of the control comparison (see Tables 3 and 4, β_m), one can assume that there is an impairment in FC-values in the averaging-method. So, option 1 is not plausible.

It is not possible, to compare FC-values of the concatenation-method to the true FC-values, as latter are not known. Thus, option 3 is not assessable. However, the recommendation of this option would be to use the concatenation-method, which is the same as in option 4. So, it remains to assess, which of these methods is accompanied with a stronger dropout based impairment of FC-values.

The difference measure between these two methods, represent the region-specific absolute difference in FC-values between these methods averaged over all connections. Apparently, this difference is closely related to the signal dropouts (see section 3.3). So, the impact of signal dropouts on FC-values is stronger in one of the methods compared to the other.

Based on the theoretical considerations described above, the FC-impairment measure can be used as an estimate for the FC-impairment in the averaging-method.

If this impairment in FC-values in the averaging-method is to account for the difference in FC-values between both methods, then this impairment is not present in the concatenation-method, else there would not be such a pronounced difference between these methods. In this case there would be a clear relationship between the FC-impairment measure and the difference measure (option 4).

If this impairment in FC-values in the averaging-method is not to account for the differences in FC-values between both methods, there must be another effect of signal dropouts on FC-values for only one of the methods, which results in the clear difference of these methods.

If there would be an additional effect of signal dropouts on FC-values in the averaging-method, it would have affected the FC-impairment measure. However, it is not plausible to assume another impact of signal dropouts that does affect the difference of the averaged FC-matrix to the concatenated FC-matrix, but does not affect the RL and LR FC-matrix, which constitute the averaged FC-matrix.

If this additional effect of signal dropouts on FC-values is present in the concatenation method, this effect would not be related to the FC-impairment in the averaging method. Accordingly, there would not be a clear relationship between the FC-impairment measure and the difference measure (option 2).

The latter is found in this investigation, so it is reasonable to assume a stronger effect of signal dropouts in the voxel-wise concatenation method than in the averaging method.

Disclaimer

The author declares that the work presented in this Master thesis is original and that no sources other than those mentioned in the text and its references have been used in creating the Master thesis.

Patrick Schwaferts
Munich, 24. July 2017