



Contents lists available at ScienceDirect

NeuroImage

journal homepage: [www.elsevier.com/locate/neuroimage](http://www.elsevier.com/locate/neuroimage)

# The effect of global signal regression on DCM estimates of noise and effective connectivity from resting state fMRI

Hannes Almgren<sup>a,\*</sup>, Frederik Van de Steen<sup>a</sup>, Adeel Razi<sup>b,c,d</sup>, Karl Friston<sup>c</sup>, Daniele Marinazzo<sup>a</sup>

<sup>a</sup> Department of Data Analysis, Faculty of Psychology and Educational Sciences, Ghent University, Henri Dunantlaan 2, 9000, Gent, Belgium

<sup>b</sup> Turner Institute for Brain and Mental Health, Monash University, 770 Blackburn Road, Building 220, Monash University, Clayton, VIC, 3800, Australia

<sup>c</sup> The Wellcome Trust Centre for Neuroimaging, University College London, 12 Queen Square, London, WC1N 3AR, United Kingdom

<sup>d</sup> Department of Electronic Engineering, NED University of Engineering and Technology, Karachi, University Road, Karachi, 75270, Pakistan

## ABSTRACT

The influence of global BOLD fluctuations on resting state functional connectivity in fMRI data remains a topic of debate, with little consensus. In this study, we assessed the effects of global signal regression (GSR) on effective connectivity within and between resting state networks (RSNs) - as estimated with dynamic causal modelling (DCM) for resting state fMRI (rsfMRI). DCM incorporates a forward (generative) model that quantifies the contribution of different types of noise (including global measurement noise), effective connectivity, and (neuro)vascular processes to functional connectivity measurements. DCM analyses were applied to two different designs; namely, longitudinal and cross-sectional designs. In the modelling of longitudinal designs, we considered four extensive longitudinal resting state fMRI datasets with a total number of 20 subjects. In the analysis of cross-sectional designs, we used rsfMRI data from 361 subjects from the Human Connectome Project. We hypothesized that (1) GSR would have no discernible impact on effective connectivity estimated with DCM, and (2) GSR would be reflected in the parameters representing global measurement noise. Additionally, we performed comparative analyses of information gain with and without GSR. Our results showed negligible to small effects of GSR on effective connectivity within small (separately estimated) RSNs. However, although the effect sizes were small, there was substantial to conclusive evidence for an effect of GSR on connectivity parameters. For between-network connectivity, we found two important effects: the effect of GSR on between-network effective connectivity (averaged over all connections) was negligible to small, while the effect of GSR on individual connections was non-negligible. In the cross-sectional (but not in the longitudinal) data, some connections showed substantial to conclusive evidence for an effect of GSR. Contrary to our expectations, we found either no effect (in the longitudinal designs) or a non-specific (cross-sectional design) effect of GSR on parameters characterising (global) measurement noise. Data without GSR were found to be more informative than data with GSR; however, in small resting state networks the precision of posterior estimates was greater after GSR. In conclusion, GSR is a minor concern in DCM studies; however, quantitative interpretation of between-network connections (as opposed to average between-network connectivity) and noise parameters should be treated with some caution. The Kullback-Leibler divergence of the posterior from the prior (i.e., information gain) - together with the precision of posterior estimates - might offer a useful measure to assess the appropriateness of GSR in resting state fMRI.

## 1. Introduction

The fMRI signal is corrupted by noise from several sources; for example, motion-induced noise, background (thermal) noise, and non-neural physiological noise that arises from cardiac and respiratory processes (Liu, 2016). This is a particular problem in resting state fMRI research, which usually aims to quantify low-frequency fluctuations in the absence of explicit perturbations. Extensive research has focused on developing and applying methods to de-noise the (resting-state) fMRI signal (e.g., Erdoğan et al., 2016; Kasper et al., 2017; Rummel et al., 2013). Some studies focus on modelling explicit (external) sources of noise (e.g., by including motion and cardiac signal as regressors in a GLM;

see, e.g., Kasper et al., 2017), others focus on identifying signal and noise directly from fMRI signals (e.g., using ICA; see, e.g., Rummel et al., 2013). A widely applied (group of) method(s) – used to de-noise fMRI signals – is to correct resting-state fMRI time-series for fluctuations in the global signal (GS), which is the average signal across all voxels of the entire MRI volume. Different types of corrections for GS fluctuations have been developed and applied, including GS regression (GSR), GS normalization, and GS subtraction (Liu et al., 2017).

The use (or omission) of GSR in fMRI connectivity studies has been a hot topic of debate (Murphy and Fox, 2016). GSR is usually applied to account for multiple non-neural sources of noise (e.g., motion and cardiac-induced signal). Some studies have shown that GSR increases the

\* Corresponding author.

E-mail addresses: [Hannes.Almgren@ugent.be](mailto:Hannes.Almgren@ugent.be) (H. Almgren), [Frederik.VandeSteen@UGent.be](mailto:Frederik.VandeSteen@UGent.be) (F. Van de Steen), [Adeel.Razi@monash.edu](mailto:Adeel.Razi@monash.edu) (A. Razi), [k.friston@ucl.ac.uk](mailto:k.friston@ucl.ac.uk) (K. Friston), [daniele.marinazzo@gmail.com](mailto:daniele.marinazzo@gmail.com) (D. Marinazzo).

<https://doi.org/10.1016/j.neuroimage.2019.116435>

Received 27 May 2019; Received in revised form 30 September 2019; Accepted 2 December 2019

Available online 6 December 2019

1053-8119/© 2019 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

efficiency of detecting significant functional connectivity (Liu et al., 2017). Fox et al. (2009), for example, showed that global signal regression (GSR) enhanced both the spatial specificity of positive correlations and detection of anti-correlation between networks (e.g., between default mode and dorsal attention networks). Similarly, Yan et al. (2013) showed that GSR decreased the correlation between motion and functional connectivity. On the other hand, other studies have shown that global signal regression introduces spurious connectivity, and leads to complex region-dependent biases (Anderson et al., 2011; Murphy et al., 2009; Saad et al., 2012). Murphy et al. (2009), for example, showed both analytically and through simulations that global signal regression causes spurious negative connectivity. Saad et al. (2012) and Gotts et al. (2013) showed that group differences can become biased after GSR. With this background, some authors have argued for new perspectives on how to study GSR (see, e.g., Power et al., 2017; Uddin, 2017).

Typically, studies investigating the impact of global signal corrections have focused on measures of connectivity that do not include a biologically plausible (forward) model; for example, correlation and independent component analysis (ICA). Therefore, connectivity estimates in these studies are not separated from estimates of hemodynamic processes or measurement noise. Dynamic causal modelling (DCM; Friston et al., 2003) is a method that allows such separation, by explicitly incorporating parameters representing state and measurement noise, effective connectivity and (neuro)vascular processes. Parameters representing noise are separated into three components: neural fluctuations that drive the system (i.e., state noise), observation or measurement noise (both local and global; e.g., caused by changes in scanner temperature), and sampling error (caused by imperfect sampling). Parameters representing (neuro)vascular processes include parameters modelling vasodilatory signal decay (related to neurovascular coupling), mean blood transit time (the average time it takes for blood to pass the veins), and the ratio of intravascular to extravascular contributions to the measured fMRI signal (Stephan et al., 2007).

In the present study, we assessed the effects of global signal regression on effective connectivity and noise parameters as estimated by (spectral) DCM for resting state fMRI (Friston et al., 2014; Razi et al., 2015). We expected that (1) global signal regression would not have a substantial impact on effective connectivity, and (2) its main effect would be reflected in the parameters representing global observation noise. In addition, we investigated whether data with GSR affords a greater information gain compared to data without GSR – as well as increasing the precision of posterior connectivity estimates. To address these questions, we decomposed (negative) free energy – a lower bound on the log model evidence – into an accuracy and a complexity term. The latter reflects the information gain afforded by the data.

Four resting state networks were analysed: three small-scale resting state networks (RSNs; namely somatomotor, saliency, and default mode network) and an additional larger network comprising all three networks. The smaller networks allowed us to assess the effect of GSR on within-network connectivity, while the combined network allowed us to assess its effects on between-network connectivity. To ensure generalizability – and a comprehensive examination of how GSR affects connectivity – we performed analyses using two different designs; namely, longitudinal and cross-sectional designs. For the former, we used four longitudinal datasets (see, Almgren et al., 2018), to which a hierarchical approach was applied (i.e., from sessions to subjects, and from subjects to group). The benefits of such a hierarchical analysis are that: (a) the total variance is captured by multiple variance components (e.g., between-session variance, between-subject variance), hence rendering parameter estimates potentially more precise, (b) within-subject effects (e.g., fluctuations in the amount of noise) are mitigated, and (c) our conclusions are not limited to one acquisition protocol or dataset. The second design was cross-sectional, for which we used the human connectome project's dataset. The benefits of this dataset are that (a) data are preprocessed by a standardized HCP pipeline, (b) the sample size is large (361 unrelated subjects), hence mitigating subject-specific effects.

## 2. Methods

### 2.1. Datasets and subjects

The longitudinal datasets were acquired by four different research institutions (see, Choe et al., 2015; Filevich et al., 2017; Gordon et al., 2017; Laumann et al., 2015). Together they comprised 20 subjects (11 females, mean and standard deviation of age at onset study:  $30.1 \pm 5.2$ ) and contained a total of 653 rsfMRI sessions (at the least 10 for each subject). For a further description of the longitudinal datasets, see Almgren et al. (2018).

In addition, the Human Connectome Project's 900 subject release (HCP; Van Essen et al., 2012) was analysed as cross-sectional dataset. To avoid issues arising with data from related subjects, we only modelled data from 361 unrelated subjects (194 females; mean and standard deviation age:  $28.7 \pm 3.7$ ; see link in the 'software availability' note). Only the first session of all subjects in the HCP dataset was analysed.

### 2.2. Data analyses

All analyses were performed using the SPM12 software package (revision 6906; Wellcome Centre for Human Neuroimaging; [www.fil.ion.ucl.ac.uk/spm/software/spm12](http://www.fil.ion.ucl.ac.uk/spm/software/spm12)), including DCM for resting state fMRI (DCM12; revision 6801), and parametric empirical Bayes (PEB; revision 6778). For details concerning data analysis we refer the reader to the link in 'software availability'.

#### 2.2.1. Preprocessing

Concerning the longitudinal datasets, the same preprocessing and time series extraction steps were used as in Almgren et al. (2018). In short, the initial five images for each resting state fMRI session were discarded, then rsfMRI data were corrected for differences in slice time (using the central slice as a reference), realigned to the first volume of each session, coregistered to an anatomical image (anatomical image prior to first functional scan session), normalized to MNI space and smoothed using a Gaussian kernel (FWHM = 6 mm). Three sessions (across all subjects and datasets) were discarded because of insufficient quality. Concerning the cross sectional data from HCP dataset, the minimally preprocessed rsfMRI data of the first scanning session of each subject were used (see, Glasser et al., 2013). These data were additionally smoothed using a Gaussian kernel of 6 mm FWHM.

#### 2.2.2. Time-series extraction

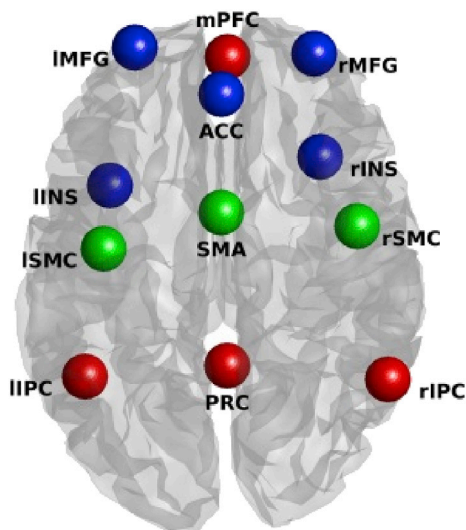
Extraction of regional time-series was the same for both types of datasets – and closely resembles the time-series extraction described in detail in Almgren et al. (2018). In short, voxels showing low frequency fluctuations were identified using a GLM with a discrete cosine basis set as regressor of interest (0.0078–0.1Hz; where the number of components was a function of the number of scans and TR), 24 motion regressors (6 regressors representing instantaneous motion, 6 regressors representing motion on the previous timepoint, and the squares of both; this motion set is commonly referred to as Friston-24; Friston et al., 1996), and two nuisance regressors (CSF signal from 5 mm ROI in circulatory system, WM signal from 7 mm ROI in brainstem). For the analyses with GSR the average time-series across the whole brain (using `spm_global.m`) were included as a nuisance regressor (see paragraph 2.2.5). SPMs including these regressors were subsequently evaluated (creating separate SPMs for regression with and without GSR). The results of an F-contrast (here, an identity matrix) across all DCT components were masked with spherical ROIs (10 mm radius), centred on coordinates extracted from template ICA maps (Smith et al., 2009). Table 1 reports the exact centre coordinates for each region and Fig. 1 shows the regions superimposed on a template brain. Time-series were summarised as the principal eigenvariate of voxels (with a corresponding F-value exceeding an alpha-threshold of 0.05) centred on the peak voxel (within the aforementioned spherical ROIs) of the SPM (sphere radius = 8 mm). This

**Table 1**

ROI regions and coordinates. Coordinates were adopted from template ICA images (Smith et al., 2009).

Network	Region	Coordinates
<b>Somatomotor</b>	Supplementary motor area	0 -12 50
	Right precentral gyrus	44 -16 48
	Left postcentral gyrus	-38 -26 56
<b>Salience</b>	Anterior cingulate cortex	0 36 22
	Left middle frontal gyrus	-28 54 14
	Right middle frontal gyrus	30 52 14
	Left insula	-36 12 -2
	Right insula	32 20 4
<b>Default mode</b>	Precuneus	2 -58 30
	Medial prefrontal cortex <sup>a</sup>	2 56 -4
	Left inferior parietal cortex	-44 -60 24
	Right inferior parietal cortex	54 -62 28

<sup>a</sup> Not included in combined network.



**Fig. 1.** Location of regions included in the present study. Colors of spheres represent the network from which they were taken (blue represents salience network, red represents default mode network, and green represents somatomotor network). Anatomical labels: PRC = precuneus; l/rIPC = left/right inferior parietal cortex; mPFC = medial prefrontal cortex; SMA = supplementary motor area; l/rSMC = left/right somatomotor region; l/rINS = left/right insula; ACC = anterior cingulate cortex; l/rMFG = left/right middle frontal gyrus.

procedure allowed for subject-specific peak locations of low-frequency fluctuations within the boundaries of the (subject-independent) template ROIs. In summary, the final ROIs were the result of a conjunction of general spherical ROIs (centred on coordinates extracted from ICA maps produced by Smith et al., 2009) and ROIs centred on the voxel with the highest resulting F-values within the aforementioned ROIs. Within this conjunction, voxels were only included if they exceeded an (uncorrected) F-value threshold ( $\alpha = 0.05$ ). The somatomotor network (SMR), salience network (SAL), and default mode network (DMN), comprised respectively three, five, and four regions. The same time-series were also included in the combined network (excluding mPFC because of its proximity to ACC).

### 2.2.3. Dynamic causal modelling for resting state fMRI

Fully-connected DCMs without exogenous inputs were created and inverted for each session and subject separately (DCM12; DCM for resting state fMRI revision 6801). Four DCMs were estimated for each session: a DCM for each network separately and a DCM for the combined network. Default shrinkage priors – not informed by functional connectivity – were

used for all networks. For each network (both separate and combined) sessions were excluded from further analyses if they did not meet all acceptance criteria (see also, Almgren et al., 2018) for both analysis types (i.e., with and without GSR). The acceptance criteria were as follows: explained variance above 60%; at least one (extrinsic) connection strength greater than  $1/8\text{Hz}$ ; at least one effectively estimated parameter, and a maximum alpha-threshold of 5% (i.e., type-I error rate; uncorrected for multiple comparisons) for which significant voxels were found in all regions of the specific network. In addition, the stringent motion exclusion criteria described in Parkes et al. (2018) were used: sessions were excluded if their mean framewise displacement (FD; as described in Jenkinson et al., 2002) was above 0.25 mm, if any FD in that session exceeded 5 mm, and if more than 20 percentage of that session's FDs exceeded 0.2 mm. For the longitudinal datasets; subjects were rejected if they had less than 8 sessions after diagnostic checks. This criterion excluded three subjects (i.e., S13, S18, S20) for all networks, S14 was excluded for SAL and the combined network, S19 was excluded for SAL and the combined network, and S4 was excluded for SMR, DMN, and the combined network. Additionally, 149, 108, 80, and 115 sessions were excluded (across subjects and datasets) for the SMR, SAL, DMN, and combined networks, respectively. For the HCP dataset, a total of 51, 42, 41, and 36 subjects were excluded because of failure to reach acceptance criteria for the SMR, SAL, DMN, and combined networks, respectively. Table 2 gives an overview of the sessions and subjects that were included for each dataset type.

### 2.2.4. Parametric empirical bayes (PEB)

Connectivity at the subject or group level was modelled using a Parametric Empirical Bayesian (PEB; Friston et al., 2016) model, with a single regressor to model average connectivity across sessions or subjects. For the longitudinal datasets, two-level hierarchical PEB models were constructed. First, connectivity over sessions was estimated at the subject-level (i.e., average connectivity over sessions was computed for each subject separately). Then, subject-level PEB models were subsequently included in a group-level model; modelling average connectivity across subjects. Default PEB settings were used for estimation at the subject and group level (see, Almgren et al., 2018). In the cross-sectional HCP dataset, connectivity was only estimated at the group level (i.e., across subjects), since only a single session was considered for each subject. For both datasets, PEB models – equipped with a single between-session (or between-subject) precision component – were specified and estimated separately for connectivity and (spectral) noise. Spectral noise parameters represented three types of noise: global state noise (i.e., across-region neural fluctuations), global observation noise, and local (i.e., region-specific) observation noise.

### 2.2.5. Global signal regression (GSR)

Dynamic causal modelling was performed with and without GSR. GSR was performed by adding a regressor representing the average signal intensity across the whole brain to the GLM model used to extract time-

**Table 2**

Number of included sessions and subjects for each network and dataset type.

Dataset Type	Number sessions and/or subjects				
	Total sample	SMR	SAL	DMN	Combined
Longitudinal	653 sessions	460 sessions	492 sessions	529 sessions	474 sessions
	(20 subj)	(70%; 16 subj)	(75%; 15 subj)	(81%; 16 subj)	(73%; 14 subj)
HCP	361 subj	310 subj	319 subj	320 subj	325 subj
		(86%)	(88%)	(89%)	(90%)

Abbreviations: HCP = Human Connectome Project, SMR = somatomotor network, SAL = salience network, DMN = default mode network; combined = combined network, subj = subjects.

series (in addition to regressors representing motion, low-frequency fluctuations, white matter, and CSF signals). Time-series were extracted from preprocessed images and were corrected for the effect of all regressors except the effect of low-frequency fluctuations. Analyses without global signal regression were performed in the same way, but excluding the regressor representing the average global signal.

### 2.2.6. Statistical tests

To compare results with and without GSR we performed a paired-sample Bayesian *t*-test on the session and subject-specific estimates of the cross-sectional and longitudinal designs, respectively. This Bayesian *t*-test has been extensively described in Roudier et al. (2009). In short, we contrasted the null-hypothesis of no effect of GSR on average connectivity with the alternative hypothesis that there is an effect of GSR on average connectivity. The null-hypothesis was represented by a standardized effect size equal to zero, while the alternative hypothesis was represented by a standardized effect size following a Cauchy distribution with the scale parameter set to  $\frac{\sqrt{2}}{2}$ . The latter distribution represents medium (standardized) effect sizes. The strength of the evidence in the Bayes factor was interpreted according to the suggestions of Kass and Raftery (1995). Practically this means we label Bayes factors ranging from 1 to 3.2 (log-scale: 0 to 1.16) as ‘negligible’ (in the words of Kass and Raftery, 1995: ‘not worth more than a bare mention’), 3.2 to 10 (log-scale: 1.16 to 2.3) as ‘substantial’, 10 to 100 (log-scale: 2.3 to 4.6) as ‘strong’, and above 100 (log-scale: more than 4.6) as ‘decisive’. Note that these Bayesian *t*-tests were performed using subject and session-specific estimates, and not directly on the first- and second-level PEB results for the cross-sectional and longitudinal data, respectively.

### 2.2.7. Assessing the quality of data features

In this work we wanted to compare the quality of *different data features* (i.e., time-series extracted with or without GSR), under the *same* model. For this, we adopted the Bayesian data comparison (BDC) approach described in Zeidman et al. (2019). This assesses the relative usefulness of data features (e.g., multiband factor of an fMRI acquisition scheme) in terms of making inferences about parameters and models. The approach can be summarised as follows: usually, the approximate log model evidence (i.e., variational free energy), is used to assess the evidence for *different models* of the *same data*. The log model evidence can always be decomposed into *accuracy* minus *complexity* (Penny et al., 2004). The complexity term represents the Kullback-Leibler divergence of the posterior density from the prior density. In other words, it scores the relative entropy or information gain afforded by the data. This is usually construed in terms of the number of degrees of freedom or parameters used to explain the data. However, it also reflects the reduction in uncertainty about model parameters after having seen the data. We therefore can use the decomposition of (negative) free energy into accuracy and complexity to ask whether different data features are more or less salient. In other words, we can quantify the reduction in uncertainty about model parameters afforded by one sort of data feature, relative to another – by simply looking at the differences in complexity. Generally speaking, we would consider data that were more informative (i.e., have a greater complexity) to be better than uninformative data. We therefore evaluated the complexity with and without GSR at the first level (i.e. after estimating DCMs at the session-level). To pool evidence across subjects, we assumed that the relative informative value of the data would be similar for all sessions and subjects. Therefore a fixed-effects approach was conducted to pool evidence across sessions (for the longitudinal datasets) and subjects (for the human connectome data). Practically, we computed the sum of the complexities over sessions for both models for each subject separately. The difference between the present analyses and the empirical part of Zeidman et al. (2019) are the data features (i.e., data with or without GSR) and the level at which usefulness of data was assessed (1st level DCM vs group-level). Additionally, we also compared the certainty of posterior estimates for both analysis types.

**Table 3**  
Summary of the results.

Influence of GSR on:	Longitudinal analysis	Cross-sectional analysis (HCP dataset)
<b>Within-network connectivity</b> (small RSNs)	<ul style="list-style-type: none"> <li>- <b>Network-level:</b></li> <li>→ RMSD<sup>1</sup>: small (<math>\leq 0.1</math>Hz)</li> <li>→ HA<sup>2</sup>: no difference in hemispheric dominance</li> <li>- <b>Connection-level:</b></li> <li>→ Most connections showed negligible effects of GSR</li> <li>→ Minority of connections small effects of GSR</li> </ul>	<ul style="list-style-type: none"> <li>- <b>Network-level:</b></li> <li>→ RMSD<sup>1</sup>: small (<math>\leq 0.1</math>Hz), except SMR (0.17Hz)</li> <li>→ HA<sup>2</sup>: no difference in hemispheric dominance</li> <li>- <b>Connection-level:</b></li> <li>→ Some extrinsic connections showed decisive evidence for change in connectivity after GSR; however, most effect sizes were small (<math>&lt; 0.1</math>Hz)</li> </ul>
<b>Between-network (BN) connectivity</b>	<ul style="list-style-type: none"> <li>- <b>Network-level:</b></li> <li>→ RMSD<sup>1</sup>: small (0.04Hz)</li> <li>→ Negligible effect size concerning influence on BN<sup>3</sup> connectivity (maximum difference 0.04Hz)</li> <li>→ Substantial to strong evidence for a very small decrease (0.02Hz and 0.04Hz) in average influence of DMN<sup>5</sup> on SAL<sup>6</sup> and SMR<sup>4</sup></li> <li>- <b>Connection-level:</b></li> <li>→ Small decreases in connectivity were found after GSR</li> </ul>	<ul style="list-style-type: none"> <li>- <b>Network-level:</b></li> <li>→ RMSD<sup>1</sup>: small (0.09Hz)</li> <li>→ Decisive evidence for influence of GSR on average BN<sup>3</sup> connectivity, but effects were small (average 0.05Hz)</li> <li>- <b>Connection-level:</b></li> <li>→ Excitatory influences became inhibitory after GSR (from SMR<sup>4</sup> and DMN<sup>5</sup> on SAL<sup>6</sup>)</li> <li>→ Additional inhibitory influences after GSR (from SAL<sup>6</sup> and SMR<sup>4</sup> on DMN<sup>5</sup>)</li> <li>→ Inhibitory effect SMA (part of SMR) on all connections in SAL disappeared after GSR</li> </ul>
<b>Global observation noise parameters</b>	- No influence	- Non-negligible influences, but not specific to observation noise parameters
<b>Optimal data features</b>	<ul style="list-style-type: none"> <li>- Small networks, separate estimates: Data <i>without</i> GSR more informative, but less precise estimates</li> <li>- Combined network: Data <i>without</i> GSR more informative, and more precise estimates</li> </ul>	<ul style="list-style-type: none"> <li>- Small networks, separate estimates: Data <i>without</i> GSR more informative, but less precise estimates</li> <li>- Combined network: Data <i>without</i> GSR more informative, and more precise estimates</li> </ul>

Abbreviations: <sup>1</sup>Root mean squared error, <sup>2</sup>Hemispheric Asymmetry, <sup>3</sup>Between-network, <sup>4</sup>somatomotor network, <sup>5</sup>default mode network, <sup>6</sup>salience network.

## 3. Results

Detailed results for the longitudinal and cross-sectional datasets are outlined (separately) in [subsections 3.1 and 3.2](#), respectively. [Table 3](#) gives a summary of the main results. Results for simulations with and without GSR are shown in [Supplementary Fig. 1](#).

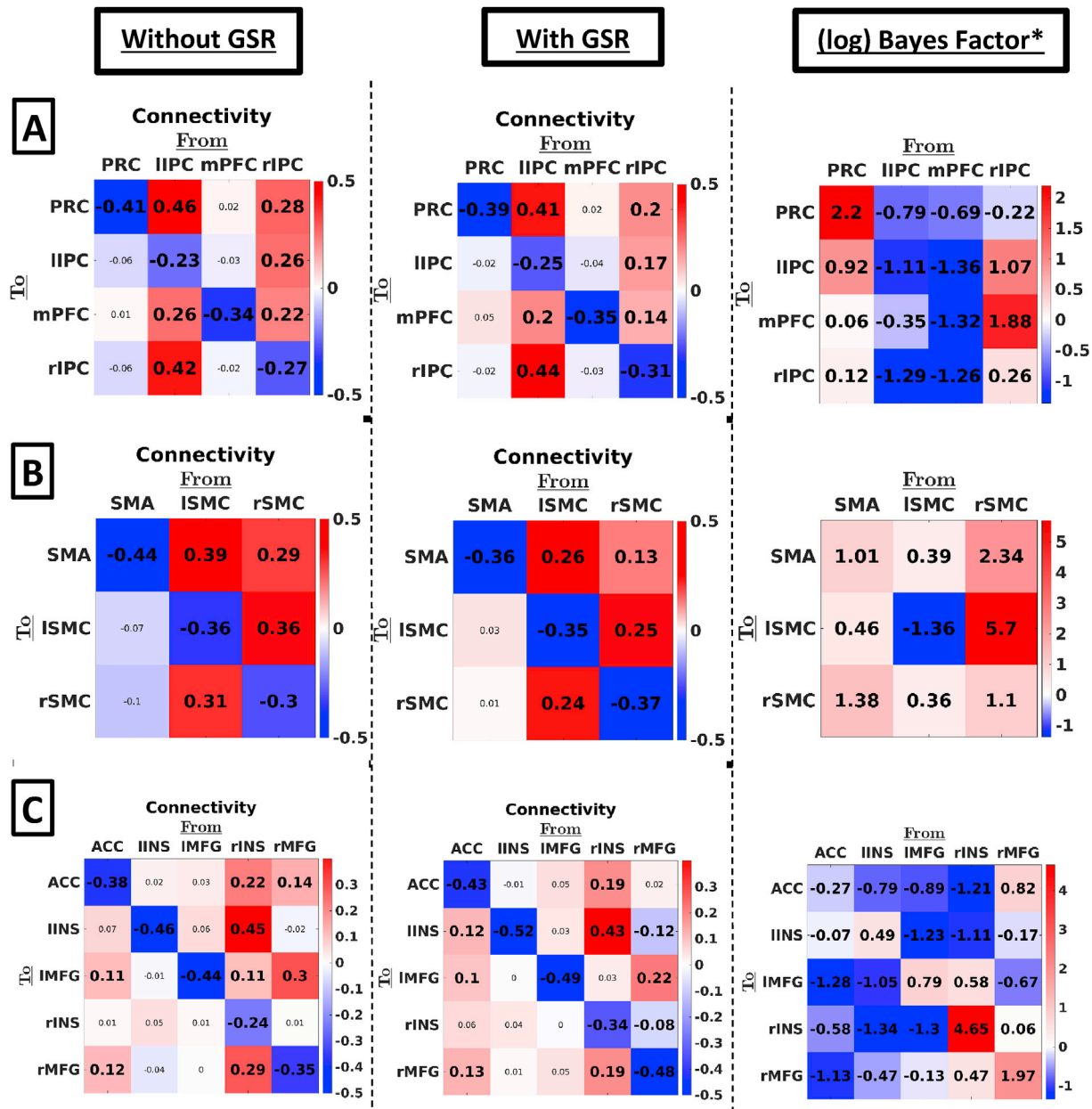
### 3.1. Longitudinal datasets

This section describes the results concerning the analysis using longitudinal datasets.

#### 3.1.1. The influence of GSR on within-network connectivity

[Fig. 2](#) shows the effect of global signal regression on effective connectivity within the three networks separately (i.e., DMN, SMR, and SAL). The differences between connectivity for analyses using data with and without GSR (middle and left panel, respectively) were remarkably small. *Across connections*, the average root mean squared difference (RMSD) between connectivity with and without GSR (self-connections





**Fig. 2.** First and second column represent effective connectivity without and with GSR – using the longitudinal datasets – within three resting state networks: (A) DMN, (B) SMR, and (C) SAL. The color of the squares indicates the MAP connection strength for the respective connection (all in linear scale). Thresholding for inference used a posterior probability of 90%. Values in smaller fonts represent connections with a posterior probability smaller than 90%. Right columns show the (natural) log Bayes factors of the alternative hypothesis of an effect of GSR versus the null hypothesis of no effect (see methods for a more detailed description). Anatomical Abbreviations: SMA = supplementary motor area, l/r SMR = left/right somatomotor region, ACC = anterior cingulate gyrus, l/r INS: left/right insula, l/r MFG: left/right middle frontal gyrus, PRC = precuneus, l/r IPC = inferior parietal cortex, mPFC = medial prefrontal cortex. \*In this and the next figure, the Bayes factors are based on the first level (subject-specific) PEB estimates of connectivity – not on the second level (between subject) PEB estimates of group means.

converted to Hz) was 0.10, 0.06, and 0.05Hz for SMR, SAL, and DMN networks, respectively, which is smaller than or equal to the heuristic threshold of 0.1Hz usually used in DCM studies (see, e.g., Razi et al., 2015). On average, connectivity was slightly closer to the prior mean after GSR, however the effect was small (average difference in deviation from the prior mean was 0.07, 0.02, and 0.03Hz for SMR, SAL, and DMN, respectively).

Global signal regression also had little effect on patterns of hemispheric asymmetry: the salience network showed higher outgoing influence from regions located in the right compared to the left hemisphere (without GSR: asymmetry = -0.20Hz, SD = 0.03Hz, posterior probability (PP) < 0.01 and with GSR: asymmetry = -0.11Hz, SD = 0.03Hz,

PP < 0.01), while no hemispheric asymmetry concerning outgoing connectivity was found within the SMR in both cases (without GSR: asymmetry = -0.04Hz, SD = 0.06Hz, PP = 0.26; with GSR: asymmetry = 0.01Hz, SD = 0.04Hz, PP = 0.60), and higher outgoing influence from the left compared to the right hemisphere was found both without and with GSR (without GSR: asymmetry = 0.11Hz, SD = 0.05, PP = 0.99; with GSR: asymmetry = 0.15Hz, SD = 0.05Hz, PP > 0.99).

At the connection-level, seven extrinsic connections (18% of a total of 38 extrinsic connections) showed a practically relevant change after GSR (i.e., difference > 0.1Hz; see also, Razi et al., 2015), which were all either part of SMR or SAL. Additionally, in the salience network two excitatory connections disappeared, and two inhibitory plus an excitatory

connection emerged after GSR (using a threshold of PP = 0.90).

To statistically test the difference between connectivity using data with versus without GSR we performed a Bayesian *t*-test on each connection (see paragraph 2.2.6 in methods-section). The resulting Bayes factors (alternative over null hypothesis) are shown in the right column of Fig. 2. Focusing on extrinsic (i.e., between-network) connections, results showed that 11% showed substantial to decisive evidence (i.e., log Bayes factor greater than 1.16) for the alternative hypothesis modelling a medium effect size. Focusing on intrinsic connectivity (i.e., estimates of self-inhibition) we found that about 25% of connections showed such evidence for the alternative hypothesis.

3.1.2. The influence of GSR on between-network connectivity

Fig. 3 shows the effect of GSR on connectivity in the large combined network, which includes within- and between-network connectivity. Row A shows the connection-specific estimates, Row B shows the average connectivity between networks. Note the high similarity in within-network connectivity for DMN and SAL in the combined network compared to networks that were estimated separately (see, Fig. 2). Subject-specific connectivity patterns are shown in Supplementary Fig. 2.

Across between-network connections, the effect of GSR was small: The root mean squared difference between connectivity with and without GSR was 0.04Hz across all between-network connections. The mean influence networks had on each other (panel B) was negligible (<0.05Hz) both with and without GSR. Results from statistical testing showed that two cases of average between-network connectivity showed substantial to strong evidence for a medium effect of GSR (here: decrease in connectivity after GSR). However, changes in connectivity were only 0.02

and 0.04Hz.

For individual connections, without GSR we found small excitatory reciprocal connectivity between all three networks, with weak inhibitory connectivity from rINS to PRC. After GSR outgoing connectivity from ISMC (SMR network) disappeared (using a threshold of PP = 0.90), and weak inhibitory influence was found from rINS (salience network) on bilateral IPC (regions in DMN). The ACC had an excitatory influence on the DMN both with and without GSR. Statistical tests showed that only one between-network connection showed substantial evidence for a change in connectivity (i.e., a decrease in connectivity from ACC to lIPC).

3.1.3. The influence of GSR on noise parameters

To have sufficient spatially distributed information to allow robust estimates of (global) spectral noise components – and their average across sessions or subjects – we restricted our analysis to the combined network. In DCM, the spectral density of noise is modelled with a power law distribution having two parameters; namely, its amplitude and an exponent. Fig. 4 shows the effect of GSR on parameters representing (spectral) noise, which include endogenous fluctuations that drive the system (i.e., state-noise) and global measurement (or observation) noise. The first parameter of each noise-component represents the amplitude of fluctuations, while the second parameter represents the shape (exponent) of the noise spectrum. It is evident that, GSR did not have an effect on any of the (global) noise components.

3.1.4. Identifying data features

To identify the most informative data features, we decomposed (negative) free energy into accuracy and complexity, where accuracy is

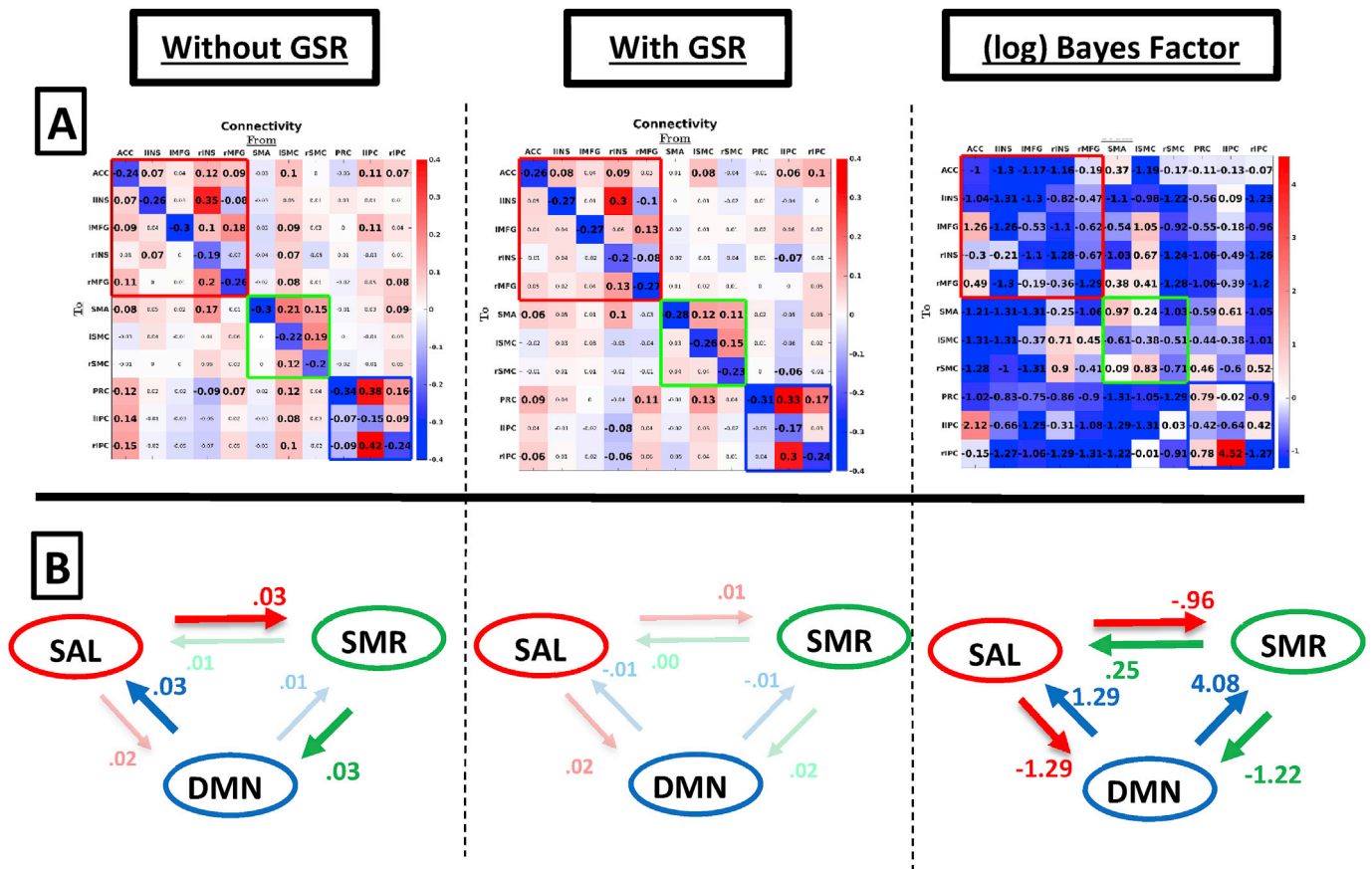
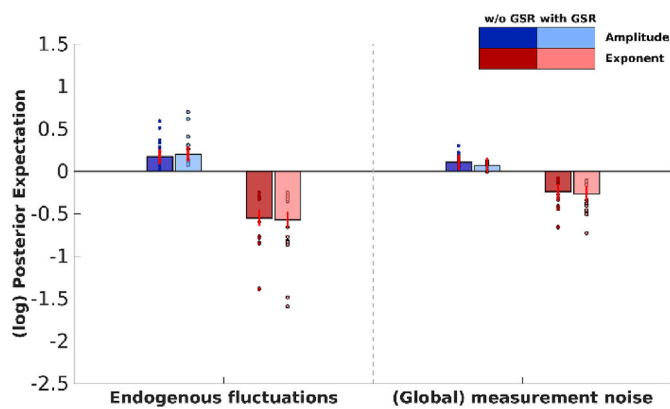


Fig. 3. Group-level between-network connectivity without GSR (left column), with GSR (middle column), and (natural) log Bayes factors (right column) testing for a difference in connectivity after GSR versus the null hypothesis of no difference in connectivity after GSR. Row A shows estimates for individual connections. Red, green, and blue empty squares indicate connectivity within SAL, SMR, and DMN, respectively. Values in smaller fonts represent connections with a posterior probability smaller than 90%. Row B shows average connectivity between networks (across connections). Dimmed arrows and values depict connections with a posterior probability <90%.



**Fig. 4.** The effect of GSR on (spectral) noise parameters within the combined network. Bar heights depict the group-level maximum *a posteriori* (MAP) estimates for the respective parameter, small circles depict subject-specific estimates (not empirically optimized). The color of the bars depict the parameter type (i.e., amplitude and frequency). Lightness of the bars and markers depicts connectivity without or with GSR (darker and lighter, respectively). Red lines depict 90% credibility intervals for the posterior estimates (PEB.Cp). Parameters are shown in log-scale and are relative to the prior mean.

the expected likelihood of the data under posterior beliefs and complexity is the Kullback-Leibler divergence between the posterior and prior densities. Complexity can thus be considered as a measure of the informative value of the data (Zeidman et al., 2019). We compared the estimated complexity with and without GSR at the session-level, which is directly dependent on the data. We used a fixed-effect approach to pool complexity over sessions for each subject separately (see, e.g., Stephan et al., 2009). DCM estimation *without* GSR was found to have the highest complexity in 87.5% of subjects for the SMR network, in 100% of subjects for the SAL network, 75% of participants for the DMN, and for 100% of subjects in the combined network, with strong evidence in all subjects. However, the certainty of posterior estimates (computed as the negative entropy of the posterior distribution) was greater for the data with GSR in 81.25%, 86.7%, and 87.5% of subjects for SMR, SAL, and DMN, respectively. For the combined network, the opposite pattern was observed: 100% of subjects showed more precise estimates without GSR.

### 3.2. Cross-sectional dataset

This section outlines the results concerning the HCP dataset.

#### 3.2.1. The influence of GSR on within-network connectivity

Fig. 5 shows the effect of GSR in the three resting-state networks, both with and without GSR. *Across all connections*, the average root mean squared difference (RMSD) between connectivity with and without GSR was 0.17, 0.10, and 0.07Hz for SMR, SAL, and DMN networks, respectively. The average decrease in deviation from the prior mean with GSR (compared to without GSR) was 0.09, 0.04, 0.04Hz for SMR, SAL, and DMN, respectively. Global signal regression had little effect on patterns of hemispheric asymmetry: the salience network showed higher outgoing connectivity from regions located in the left compared to the right hemisphere (without GSR: asymmetry = 0.10Hz, SD = 0.02Hz, PP > 0.99; with GSR: asymmetry = 0.05Hz, SD = 0.01Hz, PP > 0.99), right hemispheric dominance was found within the SMR for both cases (without GSR: asymmetry = -0.46Hz, SD = 0.04, PP < 0.01; with GSR: asymmetry = -0.22Hz, SD = 0.03Hz, PP < 0.01), and left dominance was found for the DMN with and without GSR (without GSR: asymmetry = 0.14Hz, SD = 0.03Hz, PP > 0.99; with GSR: asymmetry = 0.08Hz, SD = 0.02Hz, PP > 0.99). Interestingly, asymmetry decreased for all networks after GSR.

*At the connection-level*, nine extrinsic connections (24% of extrinsic connections) showed a practically significant change after GSR (i.e.,

difference > 0.1Hz): two connections in SMR, four connections in SAL, and three connections in DMN. With the exception of two connections, these were efferent connections from the (network-specific) dominant region. Five extrinsic connections (13% of total number of extrinsic connections) emerged or disappeared (threshold PP = 0.90) after GSR (although they were all small in magnitude). Additionally, one extrinsic connection changed sign (from excitatory to inhibitory) after GSR (i.e., connections from left INS to right MFG in the salience network). Statistical tests showed that around 39% of extrinsic connections showed substantial to conclusive evidence in favour of the alternative hypothesis of a difference in connectivity after GSR. Concerning intrinsic connections (i.e., inhibitory self-connections) 83% showed substantial to conclusive evidence in favour of the alternative hypothesis.

#### 3.2.2. The influence of GSR on between-network connectivity

Fig. 6 shows the between- and within network connectivity for the combined network. Row A shows the separate connectivity estimates, Row B shows the average connectivity between networks.

*Across connections*, the root mean squared difference between connectivity with and without GSR was 0.09Hz across all between-network connections. Note again the similarity in connectivity compared to the separately estimated networks (see, Fig. 5). The average influence networks had on each other (Fig. 6; Row B) was negligible irrespective of processing method, except for a small positive influence from DMN on SAL which became negligible after GSR. Statistical tests showed that all connections had conclusive evidence (log Bayes factor above 4.6) for a change in connectivity after GSR, which is probably related to low levels of between-subject variability in *averaged* between-network connectivity.

*At the level of individual connections*, we mainly observe excitatory influence from SMR and DMN on SAL using data without GSR, which becomes mainly inhibitory after GSR. Additionally, after GSR we observe inhibitory connectivity from SAL and SMR on DMN, and from rSMC on all DMN regions, which was not found without GSR. Statistical tests showed that 49% of between-network connections showed substantial to conclusive evidence in favour of the alternative that GSR had an effect on effective connectivity.

#### 3.2.3. The influence of GSR on noise parameters

Fig. 7 shows the effect of GSR on parameters representing (spectral) noise, which include endogenous fluctuations that drive the system (i.e., state-noise) and global measurement (or observation) noise. Clearly, GSR had an effect on both estimated endogenous fluctuations and global measurement (observation) noise, except for the exponent of endogenous fluctuations, showing that the effect of GSR is not specific to the parameter representing global measurement noise. Note that the exponent components (i.e., parameters that model the shape of the spectrum) of group-effects yielded in some cases unintuitive results (i.e., group estimates at lower end of subject-specific estimates), which most likely arise from complex covariance structures between these components (see, Kasess et al., 2010 for similar unintuitive results using Bayesian parameter averaging in DCM).

#### 3.2.4. Identifying data features

A fixed-effect approach was used to pool complexity over subjects (see, e.g., Stephan et al., 2009). For all networks, DCM estimation without GSR was found to have greater complexity – or information gain – compared to estimation with GSR, with strong evidence in all cases. However, again the certainty of the estimates (i.e., negative entropy) was greater for the data with GSR for SMR, DMN, SAL, respectively, but the opposite pattern was found for the combined network.

## 4. Discussion

In this study we investigated the effect of GSR on effective connectivity and parameters representing global observation noise, estimated with spectral DCM. Additionally, we investigated which data (i.e., data



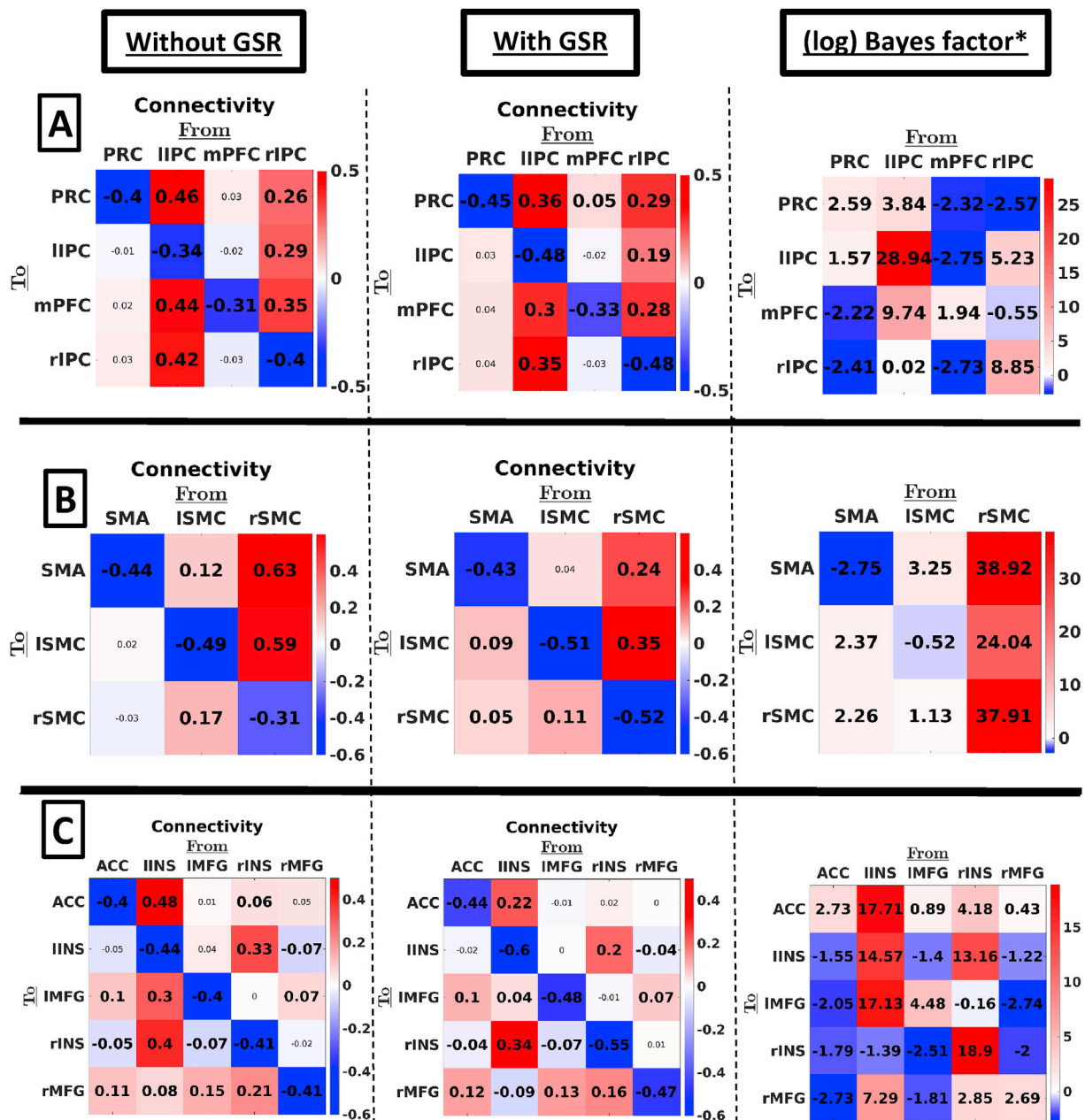


Fig. 5. First and second column represent effective connectivity without and with GSR – using the HCP dataset – within three resting state networks: (A) DMN, (B) SMR, and (C) SAL. The color of the squares indicates the MAP connection strength for the respective connection (all in linear scale). Thresholding for inference was done using a posterior probability of 90%. Values in smaller fonts represent connections with a posterior probability smaller than 90%. Right columns show the (natural) log Bayes factors of the alternative hypothesis of an effect of GSR versus the null hypothesis of no effect (see methods for a more detailed description). Anatomical Abbreviations: SMA = supplementary motor area, l/r SMR = left/right somatomotor region, ACC = anterior cingulate gyrus, l/r IINS: left/right insula, l/r MFG: left/right middle frontal gyrus, PRC = precuneus, l/r IPC = inferior parietal cortex, mPFC = medial prefrontal cortex. \*In this and the next figure, the Bayes factors are based on the (subject-specific) DCM estimates of connectivity – not on the (between subject) PEB estimates of group means.

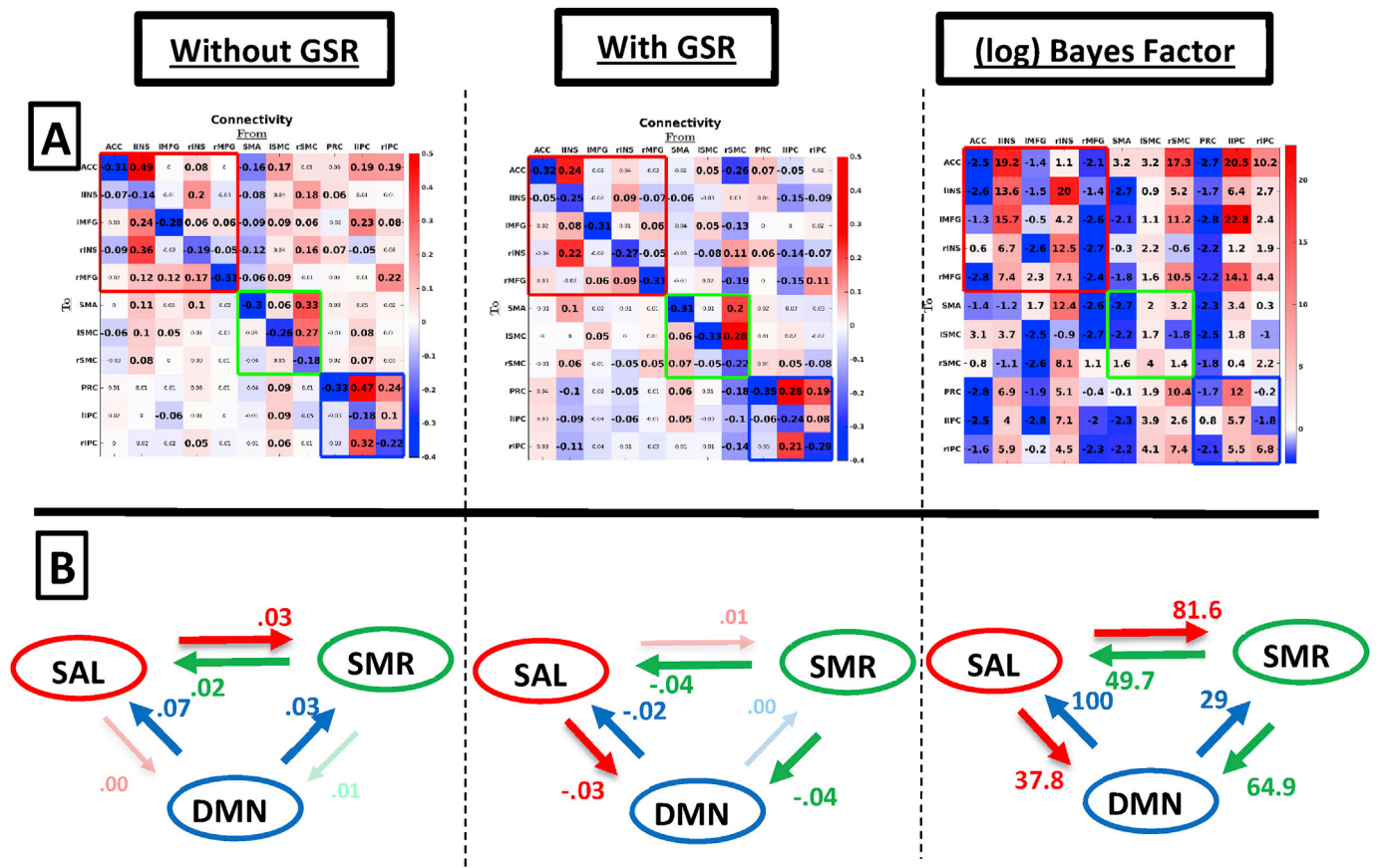
extraction with or without GSR) were most informative (i.e., have the largest complexity) for estimation of DCM parameters. We focused on both within- and between-network connectivity in two different designs (longitudinal and cross-sectional). The use of different designs allowed us to investigate the generalizability of our conclusions, and to discover divergent effects of GSR in different designs.

In general, we found negligible to small effects of GSR on connectivity within small (separately estimated) RSNs, with little evidence for a change in effective connectivity after GSR. At the network-level the effect sizes attributable to GSR were very small, both in terms of root-mean-squared deviation as well as hemispheric asymmetry. Additionally, the majority of individual connections did not change sign. These results

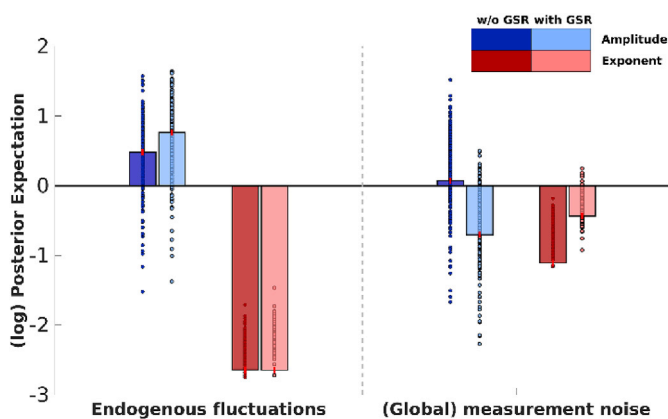
show that GSR in small RSNs should not constitute a major concern in future research using resting state DCM. Our results agree with studies of functional connectivity, showing that connections with *functionally related* areas mostly remain significantly positive after GSR (e.g., Chang & Glover, 2009; Weissenbacher et al., 2009).

Concerning between-network connectivity, we found two important effects: the effect of GSR on the between-network connectivity *across* connections (i.e., net influence between networks and root-mean-squared difference) was negligible to small, while the effect of GSR on *individual* connections was non-negligible. Results of (Bayesian) statistical tests showed that most of the latter connections showed evidence for the null-hypothesis in the longitudinal dataset, while many effective





**Fig. 6.** Group-level (HCP dataset) between-network connectivity without GSR (left column), with GSR (middle column), and (natural) log Bayes factors (right column) of the alternative of a difference in connectivity after GSR versus the null hypothesis of no difference in connectivity after GSR. Row A shows estimates for each connection separately. Red, green, and blue empty squares indicate connectivity within SAL, SMR, and DMN, respectively. Values in smaller fonts represent connections with a posterior probability smaller than 90%. Row B shows average connectivity between networks. Dimmed arrows and values depict connections with a posterior probability <90%.



**Fig. 7.** The effect of GSR on (spectral) noise parameters within the combined network. Bar heights depict the group-level maximum *a posteriori* estimates for the respective parameter, small circles depict subject-specific estimates. The color of the bars depicts the parameter type (i.e., amplitude and exponent). Lightness of the bars and markers depicts connectivity without or with GSR (darker and lighter, respectively). Red lines depict 90% credible intervals for the posterior estimates (PEB.Cp). Parameters are shown in log-scale and are relative to the prior mean.

connections showed substantial to conclusive evidence for a change after GSR in the cross-sectional dataset. At the level of individual connections many connections between DMN and SAL were excitatory when GSR was not applied, most prominently for the HCP dataset (unidirectional

influence from DMN in case of HCP), while many connections were inhibitory (and slightly more bidirectional) *after* GSR. Furthermore, we found both excitatory and inhibitory influence from SMR on SAL, which became inhibitory or non-existent after GSR. The general decrease in excitatory (and increase in inhibitory) between-network connectivity after GSR has also been found in studies focusing on functional connectivity (see, [Murphy and Fox, 2017](#)). In conclusion, the robustness of network-level effects, in contrast to connection-level effects, with respect to GSR shows that DCM connections are best interpreted jointly (which has been argued in other studies, e.g., [Almgren et al., 2018](#)).

The between-network results are somewhat in contrast to the DCM study of [Zhou et al. \(2018\)](#), which found inhibitory influence from the salience network on the core DMN, but not *vice versa*, *without* GSR (which we found in the opposite direction after GSR in the HCP dataset). There are two important differences in results between [Zhou et al. \(2018\)](#) and the result we obtained in the present study that should be explained. First, in the HCP data we found mainly inhibitory connectivity (after GSR) from DMN to SAL, while [Zhou et al. \(2018\)](#) found inhibitory connectivity in the opposite direction. This might be explained by a difference in subject characteristics: Subjects in our study were mainly young adults (mean age was 30.1 and 28.7 years respectively, for longitudinal and cross-sectional design), while the age-group included in [Zhou et al. \(2018\)](#) was on average younger (mean age = 17.4 years). Indeed, some studies suggest that functional connectivity between DMN and higher-order cognitive networks are affected by age (e.g., [Zonneveld et al., 2019](#)). Second, [Zhou et al. \(2018\)](#) found important between-network connectivity when averaged across connections, while we only found effects at the connection-level (see, e.g., [Fig. 6](#)). This might

be explained by their use of functional connectivity to inform priors: between-network effects might be more coherent when connectivity parameters are informed by FC, hence increasing detection of between-network effects. In our study we did not use FC-informed priors for two reasons: (1) Bayesian model comparison showed strong evidence in favour of the use of standard (i.e., shrinkage) priors, (2) the use of group-specific (across-subject) FC-priors at higher levels in our hierarchical design might have obscured results. Additionally, the difference in sample size and specific ROI specification method might also have had an effect on the divergent results.

On average, connectivity was closer to the prior mean after GSR. Moreover, hemispheric asymmetry (difference between left and right outgoing connectivity) also decreased after GSR. For extrinsic connectivity, this is in line with multiple studies showing that the connectivity decreases after GSR (e.g., Chang & Glover, 2009; Fox et al., 2009; Weissenbacher et al., 2009). The deviation from the prior mean also has an influence on the informative value of the data, which is discussed further below. Notably, hemispheric asymmetry also decreased after GSR. Hemispheric asymmetry was here defined as the (network-specific) efferent connectivity from regions in one hemisphere relative to the outgoing connectivity from regions in the other hemisphere. Hence, a change in hemispheric asymmetry would mean that the magnitude of the decrease in connectivity is different for regions in both hemispheres. Possibly the extent to which certain regions contribute to the global signal differed between hemispheres. This explanation is in line with McAvoy et al. (2016) who found an asymmetric contribution of regions to the global signal. However, a direct comparison with the functional connectivity literature concerning the size of the effect of GSR is difficult given the differences in models, difference in statistical framework (Bayesian vs frequentist), and specific networks under investigation. It is also important to note that our conceptualization of hemispheric asymmetry differs from that in functional connectivity studies, since DCM estimates directed influences.

Remarkably, the *within-network* connectivity was often very similar when networks were estimated separately compared to networks as a whole (compare e.g., Figs. 5 and 6). This was also found by Ushakov et al. (2016), who showed that connectivity within the core DMN was very similar when either left or right hippocampus was included in the network. This shows that group-level resting-state DCM results are quite robust against addition of extra networks and regions.

Contrary to our expectations, we either found no effect (longitudinal designs) or an unspecific effect (cross sectional design) on measurement noise parameters. Several explanations can be put forward to account for these counter-intuitive results, and the explanations might be different for both designs. Concerning the absence of effect in the longitudinal designs, it might be that the effect of GSR was captured by the between-subject components of the hierarchical Bayesian model. This is possible, since the datasets differed in terms of scanner type, pulse sequence parameters, and subject characteristics, which might cause divergent effects on noise parameters. Therefore, it is possible that the effect of GSR was cancelled out at the group-level, while captured by parameters representing between-subject variance. Indeed, we found an important effect (i.e., increase vs decrease in parameters) of GSR on noise parameters for some individual subjects, but not for others. Concerning the cross-sectional HCP dataset, the effect was captured by multiple noise parameters. This unspecific effect is possible, since resting state fMRI has no explicit exogenous input (e.g., an experimental design) that could inform more precise estimation. Noise parameters that are shared across regions (e.g., global state and measurement noise) might therefore be all sensitive to the global signal. Note that we did not model the sources of global signal regression, but its effects on the parameter representing global signal noise.

Additionally, we asked whether data was more informative for 1st level parameter estimation (i.e., had the highest complexity) with or without GSR. Connectivity estimation using data without GSR was found to yield the greatest complexity for all datasets and networks. The

complexity term is the Kullback-Leibler divergence between the prior and posterior density, which can be interpreted as the increase in information after updating the prior with the data. Complexity essentially depends on (1) divergence of the mean of the posterior from the mean of the prior density, (2) precision of the posterior density compared to the precision of the prior, and (3) change in covariance patterns between posterior and prior densities of parameters. The observations of higher complexity without GSR in most cases probably is for a part attributable to a greater divergence from the posterior mean. The few (subject-specific) observations of higher complexity with GSR are at odds with this explanation, since after GSR a smaller deviation from the prior mean was observed. In general, data without GSR provide more information to estimate effective connectivity compared to data after GSR, which might slightly encourage the use of data without GSR in DCM studies (however, more measures are needed to fully licence this conclusion). However, the precision of the parameter estimates also has an influence on the KL-divergence. Therefore, we compared the precision of estimations (computed as the negative entropy) for data with versus without GSR. Here we found that for *small networks*, data with GSR was found to yield greatest precision. However, for the combined network, data *without* GSR was found to yield most precise estimates. This might be associated with the (potentially spurious) inhibitory connectivity that was found for estimation with GSR in the combined network.

In general, we can thus conclude that data without GSR are more informative for estimation of effective connectivity. One should note that this complexity was computed at the session-level, and that this level is directly dependent on the data. For the other levels (i.e., subject- and group-level), the complexity term includes the Kullback-Leibler divergence of second-level parameters (e.g., between-session variance; group-level effective connectivity) from their prior density, which depends on the estimated parameters at lower levels, and not directly on the data. Possibly, between-session (between-subject) differences and subject (group) level estimates are less informed by the estimates based on data without GSR.

In this work we assessed the (practical) effects of GSR on parameters estimated with DCM for resting state fMRI. However, multiple alternative methods exist to eliminate (global) noise from fMRI data, including temporal ICA (Glasser et al., 2018), dynamic global signal regression based on blood arrival time (dGSR; Erdoğan, Tong, Hocke, Lindsey, & deB Frederick, 2016), and corrections for respiratory and cardiac signals (Chang et al., 2009). Although these alternatives are less extensively studied in fMRI research compared to ‘canonical’ GSR, they have certain features that *might* make them superior to GSR. These methods are outside of the scope of the present study, which was to assess the effect of the (widely studied) ‘canonical’ GSR on DCM parameters. However, we also checked the consistency of our results compared to the effects of global signal *normalization* (dividing timeseries by the global signal instead of regressing out), and we found quite similar results. We would also like to stress that it is important to not generalize the present results – which concern effective connectivity estimated with DCM – to studies of functional connectivity (statistical dependencies among brain signals; Friston, 2011). Effective connectivity with DCM includes a biophysically plausible forward model, is embedded in a Bayesian framework, and estimates directed, causal influences among neural populations. In contrast, most studies on functional connectivity describe undirected statistical relations between brain regions, seldom include a biophysical forward model, and are mostly embedded in a frequentist framework. Effects of GSR on effective and functional connectivity can therefore diverge.

In conclusion, GSR is a minor concern in DCM studies. However, individual between-network connections (as opposed to average between-network connectivity) should be interpreted with some caution. Additionally, we suggest the use of the complexity term – in combination with the certainty of estimation – to assess the relative informative value of the data with versus without GSR.

## Data and code availability statement

### Software availability

Code to reproduce analyses, results and figures is available at: [https://github.com/halmgren/Pipeline\\_effect\\_GSR\\_effective\\_connectivity\\_rsFMRI](https://github.com/halmgren/Pipeline_effect_GSR_effective_connectivity_rsFMRI).

### Data availability

The following longitudinal datasets were used (and are accessible through):

- (1) 'Myconnectome': Downloaded from <https://openneuro.org/ds000031>
- (2) 'The Midnight Scan Club': Downloaded from <https://openneuro.org/ds000224>
- (3) 'Kirby Weekly': Downloaded from <https://www.nitrc.org/project/s/kirbyweekly>
- (4) 'Day2day': For availability, see [Filevich et al. \(2017\)](#); section 'Availability of data and materials' (available upon request).

The 900 subjects release of the Human Connectome Project was used for the cross-sectional dataset. They are downloadable from <https://www.humanconnectome.org>. For more information on the specific release, see <https://www.humanconnectome.org/study/hcp-young-adult/document/900-subjects-data-release>.

### Acknowledgements

This work has been funded by the Special Research Fund of Ghent University (awarded to HA; grant No. BOF16/DOC/282; <https://www.ugent.be/>), a grant for long stay abroad from the Fund for Scientific Research-Flanders (FWO-V; awarded to HA; <http://www.fwo.be>), the Fund for Scientific Research-Flanders (FWO-V; grant No. FWO14/ASP/255 awarded to FvdS; <http://www.fwo.be>), the Australian Research Council DECRA Fellowship (awarded to AR; Ref: DE170100128), and funding from the Wellcome Trust (KF: 088130/Z/09/Z; AR). The authors would like to thank the reviewers for their help in reporting this work.

### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.neuroimage.2019.116435>.

### References

- Almgren, H., Van de Steen, F., Kühn, S., Razi, A., Friston, K., Marinazzo, D., 2018. Variability and reliability of effective connectivity within the core default mode network: a multi-site longitudinal spectral DCM study. *Neuroimage* 183, 757–768. <https://doi.org/10.1016/j.neuroimage.2018.08.053>.
- Anderson, J.S., Druzgal, T.J., Lopez-Larson, M., Jeong, E.-K., Desai, K., Yurgelun-Todd, D., 2011. Network anticorrelations, global regression, and phase-shifted soft tissue correction. *Hum. Brain Mapp.* 32 (6), 919–934. <https://doi.org/10.1002/hbm.21079>.
- Chang, C., Cunningham, J.P., Glover, G.H., 2009. Influence of heart rate on the BOLD signal: the cardiac response function. *Neuroimage* 44 (3), 857–869. <https://doi.org/10.1016/j.neuroimage.2008.09.029>.
- Choe, A.S., Jones, C.K., Joel, S.E., Muschelli, J., Belegu, V., Caffo, B.S., et al., 2015. Reproducibility and temporal structure in weekly resting-state fMRI over a period of 3.5 years. *PLoS One* 10 (10), 1–29. <https://doi.org/10.1371/journal.pone.0140134>.
- Erdogan, S.B., Tong, Y., Hocke, L.M., Lindsey, K.P., deB Frederick, B., 2016. Correcting for blood arrival time in global mean regression enhances functional connectivity analysis of resting state fMRI-BOLD signals. *Front. Hum. Neurosci.* 10 (June), 1–22. <https://doi.org/10.3389/fnhum.2016.00311>.
- Filevich, E., Lisofsky, N., Becker, M., Butler, O., Lochstet, M., Martensson, J., et al., 2017. Day2day: investigating daily variability of magnetic resonance imaging measures over half a year. *BMC Neurosci.* 18 (1), 1–8. <https://doi.org/10.1186/s12868-017-0383-y>.
- Fox, M.D., Zhang, D., Snyder, A.Z., Raichle, M.E., 2009. The global signal and observed anticorrelated resting state brain networks. *J. Neurophysiol.* 101 (6), 3270–3283. <https://doi.org/10.1152/jn.90777.2008>.

- Friston, K.J., Harrison, L., Penny, W., 2003. Dynamic causal modelling. *Neuroimage* 19 (4), 1273–1302. [https://doi.org/10.1016/S1053-8119\(03\)00202-7](https://doi.org/10.1016/S1053-8119(03)00202-7).
- Friston, Karl J., 2011. Functional and effective connectivity: a review. *Brain Connect.* 1 (1), 13–36. <https://doi.org/10.1089/brain.2011.0008>.
- Friston, Karl J., Kahan, J., Biswal, B., Razi, A., 2014. A DCM for resting state fMRI. *Neuroimage* 94, 396–407. <https://doi.org/10.1016/j.neuroimage.2013.12.009>.
- Friston, K.J., Litvak, V., Oswal, A., Razi, A., Stephan, K.E., van Wijk, B.C.M., Ziegler, G., Zeidman, P., 2016. Bayesian model reduction and empirical Bayes for group (DCM) studies. *Neuroimage* 128, 413–431. <https://doi.org/10.1016/j.neuroimage.2015.11.015>.
- Friston, Karl J., Williams, S., Howard, R., Frackowiak, R.S.J., Turner, R., 1996. Movement-Related effects in fMRI time-series. *Magn. Reson. Med.* 35 (3), 346–355. <https://doi.org/10.1002/mrm.1910350312>.
- Glasser, M.F., Coalson, T.S., Bijsterbosch, J.D., Harrison, S.J., Harms, M.P., Anticevic, A., et al., 2018. Using temporal ICA to selectively remove global noise while preserving global signal in functional MRI data. *Neuroimage* 181, 692–717. <https://doi.org/10.1016/j.neuroimage.2018.04.076>.
- Glasser, M.F., Sotiropoulos, S.N., Wilson, J.A., Coalson, T.S., Fischl, B., Andersson, J.L., et al., 2013. The minimal preprocessing pipelines for the Human Connectome Project. *Neuroimage* 80, 105–124. <https://doi.org/10.1016/j.neuroimage.2013.04.127>.
- Gordon, E.M., Laumann, T.O., Gilmore, A.W., Newbold, D.J., Greene, D.J., Berg, J.J., et al., 2017. Precision functional mapping of individual human brains. *Neuron* 95 (4), 791–807. <https://doi.org/10.1016/j.neuron.2017.07.011> e7.
- Gotts, S.J., Saad, Z.S., Jo, H.J., Wallace, G.L., Cox, R.W., Martin, A., 2013. The perils of global signal regression for group comparisons: a case study of Autism Spectrum Disorders. *Front. Hum. Neurosci.* 7, 356. <https://doi.org/10.3389/fnhum.2013.00356>.
- Jenkinson, M., Bannister, P., Brady, M., Smith, S., 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* 17 (2), 825–841. <https://doi.org/10.1006/nimg.2002.1132>.
- Kasess, C.H., Stephan, K.E., Weissenbacher, A., Pezawas, L., Moser, E., Windischberger, C., 2010. Multi-subject analyses with dynamic causal modeling. *Neuroimage* 49 (4), 3065–3074. <https://doi.org/10.1016/j.neuroimage.2009.11.037>.
- Kasper, L., Bollmann, S., Diaconescu, A.O., Hutton, C., Heinze, J., Iglesias, S., et al., 2017. The PhysIO toolbox for modeling physiological noise in fMRI data. *J. Neurosci. Methods* 276, 56–72. <https://doi.org/10.1016/j.jneumeth.2016.10.019>.
- Kass, R.E., Raftery, A.E., 1995. Bayes factors. *J. Am. Stat. Assoc.* 90 (430), 773–795. <https://doi.org/10.1080/01621459.1995.10476572>.
- Laumann, T.O., Gordon, E.M., Adeyemo, B., Snyder, A.Z., Joo, S. J. un, Chen, M.Y., et al., 2015. Functional system and areal organization of a highly sampled individual human brain. *Neuron* 87 (3), 657–670. <https://doi.org/10.1016/j.neuron.2015.06.037>.
- Liu, T.T., 2016. Noise contributions to the fMRI signal: an overview. *Neuroimage* 143, 141–151. <https://doi.org/10.1016/j.neuroimage.2016.09.008>.
- Liu, T.T., Nalci, A., Falahpour, M., 2017. The global signal in fMRI: nuisance or Information? *Neuroimage* 150, 213–229. <https://doi.org/10.1016/j.neuroimage.2017.02.036>.
- McAvoy, M., Mitra, A., Coalson, R.S., d'Avossa, G., Keidel, J.L., Petersen, S.E., Raichle, M.E., 2016. Unmasking language lateralization in human brain intrinsic activity. *Cerebr. Cortex* 26 (4), 1733–1746. <https://doi.org/10.1093/cercor/bhw007>.
- Murphy, K., Fox, M.D., 2016. Towards a consensus regarding global signal regression for resting state functional connectivity MRI. *Neuroimage* (November). <https://doi.org/10.1016/j.neuroimage.2016.11.052>, 0–1.
- Murphy, Kevin, Birn, R.M., Handwerker, D.A., Jones, T.B., Bandettini, P.A., 2009. The impact of global signal regression on resting state correlations: are anti-correlated networks introduced? *Neuroimage* 44 (3), 893–905. <https://doi.org/10.1016/j.neuroimage.2008.09.036>.
- Murphy, K., Fox, M.D., 2017. Towards a consensus regarding global signal regression for resting state functional connectivity MRI. *Neuroimage* 154, 169–173. <https://doi.org/10.1016/j.neuroimage.2016.11.052>.
- Parke, L., Fulcher, B., Yücel, M., Fornito, A., 2018. An evaluation of the efficacy, reliability, and sensitivity of motion correction strategies for resting-state functional MRI. *Neuroimage* 171, 415–436. <https://doi.org/10.1016/j.neuroimage.2017.12.073>.
- Penny, W.D., Stephan, K.E., Mechelli, A., Friston, K.J., 2004. Comparing dynamic causal models. *Neuroimage* 22 (3), 1157–1172. <https://doi.org/10.1016/j.neuroimage.2004.03.026>.
- Power, J.D., Laumann, T.O., Plitt, M., Martin, A., Petersen, S.E., 2017. On global fMRI signals and simulations. *Trends Cogn. Sci.* 21 (12), 911–913. <https://doi.org/10.1016/j.tics.2017.09.002>.
- Razi, A., Kahan, J., Rees, G., Friston, K.J., 2015. Construct validation of a DCM for resting state fMRI. *Neuroimage* 106, 1–14. <https://doi.org/10.1016/j.neuroimage.2014.11.027>.
- Rouder, J.N., Speckman, P.L., Sun, D., Morey, R.D., Iverson, G., 2009. Bayesian t tests for accepting and rejecting the null hypothesis. *Psychon. Bull. Rev.* 16 (2), 225–237. <https://doi.org/10.3758/PBR.16.2.225>.
- Rummel, C., Verma, R.K., Schöpf, V., Abela, E., Hauf, M., Berruocos, J.F.Z., Wiest, R., 2013. Time course based artifact identification for independent components of resting-state fMRI. *Front. Hum. Neurosci.* 7, 214. <https://doi.org/10.3389/fnhum.2013.00214>.
- Saad, Z.S., Gotts, S.J., Murphy, K., Chen, G., Jo, H.J., Martin, A., Cox, R.W., 2012. Trouble at rest: how correlation patterns and group differences become distorted after global signal regression. *Brain Connect.* 2 (1), 25–32. <https://doi.org/10.1089/brain.2012.0080>.



- Smith, S.M., Fox, P.T., Miller, K.L., Glahn, D.C., Fox, P.M., Mackay, C.E., et al., 2009. Correspondence of the brain's functional architecture during activation and rest. *Proc. Natl. Acad. Sci. U.S.A.* 106 (31), 13040–13045. <https://doi.org/10.1073/pnas.0905267106>.
- Stephan, K.E., Penny, W.D., Daunizeau, J., Moran, R.J., Friston, K.J., 2009. Bayesian model selection for group studies. *Neuroimage* 46 (4), 1004–1017. <https://doi.org/10.1016/j.neuroimage.2009.03.025>.
- Stephan, K.E., Weiskopf, N., Drysdale, P.M., Robinson, P.A., Friston, K.J., 2007. Comparing hemodynamic models with DCM. *Neuroimage* 38 (3), 387–401. <https://doi.org/10.1016/j.neuroimage.2007.07.040>.
- Uddin, L.Q., 2017. Mixed signals: on separating brain signal from noise. *Trends Cogn. Sci.* 21 (6), 405–406. <https://doi.org/10.1016/j.tics.2017.04.002>.
- Ushakov, V., Sharaev, M.G., Kartashov, S.I., Zavyalova, V.V., Verkhlyutov, V.M., Velichkovsky, B.M., 2016. Dynamic causal modeling of hippocampal links within the human default mode network: lateralization and computational stability of effective connections. *Front. Hum. Neurosci.* 10 (October), 1–14. <https://doi.org/10.3389/fnhum.2016.00528>.
- Van Essen, D.C., Ugurbil, K., Auerbach, E., Barch, D., Behrens, T.E.J., Bucholz, R., et al., 2012. The Human Connectome Project: a data acquisition perspective. *Neuroimage* 62 (4), 2222–2231. <https://doi.org/10.1016/j.neuroimage.2012.02.018>.
- Weissenbacher, A., Kasess, C., Gerstl, F., Lanzenberger, R., Moser, E., Windischberger, C., 2009. Correlations and anticorrelations in resting-state functional connectivity MRI: a quantitative comparison of preprocessing strategies. *Neuroimage* 47 (4), 1408–1416. <https://doi.org/10.1016/j.neuroimage.2009.05.005>.
- Yan, C.-G., Cheung, B., Kelly, C., Colcombe, S., Craddock, R.C., Di Martino, A., et al., 2013. A comprehensive assessment of regional variation in the impact of head micromovements on functional connectomics. *Neuroimage* 76, 183–201. <https://doi.org/10.1016/j.neuroimage.2013.03.004>.
- Zeidman, P., Kazan, S.M., Todd, N., Weiskopf, N., Friston, K.J., Callaghan, M.F., 2019. Optimizing data for modeling neuronal responses. *Front. Neurosci.* 12, 986. <https://doi.org/10.3389/fnins.2018.00986>.
- Zhou, Y., Friston, K.J., Zeidman, P., Chen, J., Li, S., Razi, A., 2018. The hierarchical organization of the default, dorsal attention and salience networks in adolescents and young adults. *Cerebr. Cortex* 28, 726–737. <https://doi.org/10.1093/cercor/bhx307>.
- Zonneveld, H.I., Pruim, R.H., Bos, D., Vrooman, H.A., Muetzel, R.L., Hofman, A., et al., 2019. Patterns of functional connectivity in an aging population: the Rotterdam Study. *Neuroimage* 189, 432–444. <https://doi.org/10.1016/j.neuroimage.2019.01.041>.