Insight and Inference for DVARS

Soroosh Afyouni[a,b], Thomas E. Nichols[c,b,*]

[a] *Institute for Advanced Studies, University of Warwick, Coventry, CV4 7AL, UK*
[b] *Institute for Digital Healthcare, WMG, University of Warwick, Coventry, CV4 7AL, UK*
[c] *Department of Statistics, University of Warwick, Coventry, CV4 7AL, UK*

Abstract

Estimates of functional connectivity using resting state functional Magnetic Resonance Imaging (rs-fMRI) are acutely sensitive to artifacts and large scale nuisance variation. As a result much effort is dedicated to preprocessing rs-fMRI data and using diagnostic measure to identify bad scans. One such diagnostic measure is DVARS, the spatial standard deviation of the data after temporal differencing. A limitation of DVARS however is the lack of concrete interpretation of the absolute values of DVARS, and finding a threshold to distinguish bad scans from good. In this work we describe a variance decomposition of the entire 4D dataset that shows DVARS to be just one of three sources of variation we refer to as D-var (closely linked to DVARS), S-var and E-var. D-var and S-var partition the average variance between adjacent time points, while E-var accounts for edge effects, and each can be used to make spatial and temporal summary diagnostic measures. Extending the partitioning to global (and non-global) signal leads to a rs-fMRI DSE ANOVA table, which decomposes the total and global variance into fast (D-var), slow (S-var) and edge (E-var) components. We find expected values for each variance component under nominal models, showing how D-var (and thus DVARS) scales with overall variance and is diminished by temporal autocorrelation. Finally we propose a sampling distribution for squared DVARS (a multiple of D-var) and robust methods to estimate this null model, allowing computations of DVARS p-values. We propose that these diagnostic time series, images, p-values and ANOVA table will provide a succinct summary of the quality of a rs-fMRI dataset that will support comparisons of datasets over preprocessing steps and between subjects.

Keywords: DVARS, Mean Square of Successive Differences, Variance Decomposition, fMRI, Resting-State

1. Introduction

Functional connectivity obtained with resting state functional magnetic resonance imaging (rs-fMRI) is typically computed by correlation coefficients between different brain regions, or with a multivariate decomposition like Independent Components Analysis (Cole et al., 2010). Both approaches can be corrupted

*Corresponding author
Email addresses: `s.afyouni.1@warwick.ac.uk` (Soroosh Afyouni), `t.e.nichols@warwick.ac.uk` (Thomas E. Nichols)

by artifacts due to head motion or physiological effects, and much effort is dedicated to preprocessing rs-fMRI data and using diagnostic measure to identify bad scans.

Smyser et al. (2011) proposed and Power et al. (2012) popularized a measure to characterize the quality of fMRI data, an image-wide summary that produces a time series that can detect problem scans. They called their measure DVARS, defined as the spatial standard deviation of successive difference images. In fact, DVARS can be linked to old statistical methods developed to estimate noise variance in the presence of drift (see Appendix A for DVARS history).

While DVARS appears to perform well at the task of detecting bad scans — bad pairs of scans — it does not have any absolute units nor a reference null distribution from which to obtain p-values. In particular, the typical "good" values of DVARS varies over sites and protocols. The purpose of this work is to provide a formal description of DVARS as part of a variance decomposition of all data variance, present a more interpretable variants of DVARS, and compute DVARS p-values for a null hypothesis of homogeneity. By combining information from p-values and meaningful diagnostic plots, bad scans will be able to be more confidently identified and actioned.

The remainder of this work is organized as follows. We first describe the variance decomposition for the 4D data and how this relates to traditional DVARS, and other new diagnostic measures it suggests. We then describe a sampling distribution for DVARS under the null hypothesis, and mechanisms for estimating the parameters of this null distribution. We conduct some basic simulations to validate this sampling distribution and demonstrate the method on representative datasets.

2. Theory

Here we state our results concisely relegating full derivations to Appendices.

2.1. Notation

For T time-points and I voxels, let the original raw rs-fMRI data at voxel i and t be Y_{it}^R. Denote the mean at voxel i as $M_i^R = \frac{1}{T}\sum_t Y_{it}^R$, and by m^R some type of typical mean value (e.g. mean or median of the mean image $\{M_i^R\}$). We take as our starting point for all calculations the centered and scaled data:

$$Y_{it} = \frac{Y_{it}^R - M_i^R}{m^R} 100. \tag{1}$$

The scaling ensures that typical brain values before centering are around 100 and are comparable across datasets, and centering simplifies variance calculations.

2.2. DSE Variance Decomposition

Let the total ("all") variance at scan t be

$$A_t = \frac{1}{I}\sum_{i=1}^{I} Y_{it}^2 \tag{2}$$

and define two variance terms, one for fast ("differenced") variance

$$D_t = \frac{1}{I}\sum_{i=1}^{I}\left(\frac{Y_{it} - Y_{i,t+1}}{2}\right)^2, \tag{3}$$

the half difference between time t and $t+1$ at each voxel, squared and averaged over space, and one for slow variance

$$S_t = \frac{1}{I}\sum_{i=1}^{I}\left(\frac{Y_{it} + Y_{i,t+1}}{2}\right)^2, \tag{4}$$

the average between t and $t+1$ at each voxel, squared and averaged over space.

We then have the following decomposition of the average variance at time points t and $t+1$, $A_{t,t+1} = (A_t + A_{t+1})/2$

$$A_{t,t+1} = D_t + S_t, \tag{5}$$

for $t = 1, \ldots, T-1$. This has a particularly intuitive graphical interpretation: If we plot D_t and S_t at $t+1/2$, they sum to the midpoint between variances A_t and A_{t+1} found at $t+1/2$ (see Figure 1). Noting that the usual DVARS measure is

$$\mathrm{DVARS}_t = 2\sqrt{D_t}, \tag{6}$$

this shows that DVARS has a concrete interpretation, with $\mathrm{DVARS}_t^2/4$ being the "fast" variance component in the average variance at t and $t+1$.

This also leads to a decomposition of the total average variance. If we define the temporal averages

$$\begin{aligned} A =& \frac{1}{T}\sum_{t=1}^{T} A_t, \\ D =& \frac{1}{T}\sum_{t=1}^{T-1} D_t, \\ S =& \frac{1}{T}\sum_{t=1}^{T-1} S_t, \end{aligned} \tag{7}$$

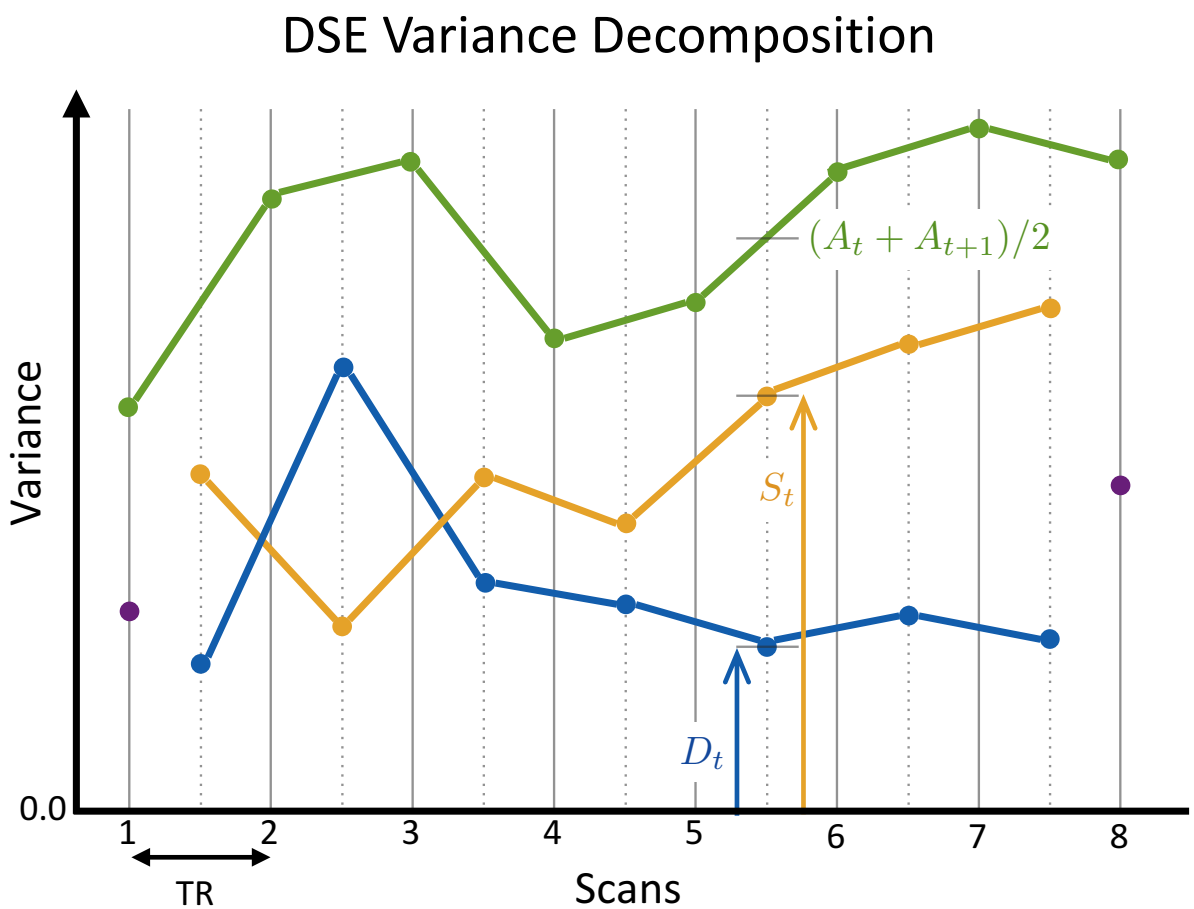
and lastly an "edge" variance term

$$\begin{aligned} E_t =& \frac{1}{I}\sum_{i=1}^{I} Y_{it}^2/2, \\ E =& \frac{1}{T}\left(E_1 + E_T\right), \end{aligned} \tag{8}$$

we have the following "DSE" decomposition

$$A = D + S + E. \tag{9}$$

That is, the total variance ("A-var") in the 4D dataset is the sum of variance terms attributable to fast variance ("D-var"), slow variance ("S-var") and edge variance ("E-var").



Decomposition at *t*,*t*+1

$$(A_t + A_{t+1})/2 \;=\; D_t + S_t$$

Overall Decomposition

$$\sum_{t=1}^{T} A_t = \sum_{t=1}^{T-1} D_t + \sum_{t=1}^{T-1} S_t + \sum_{t=1,T} E_t$$

Figure 1: Illustration of the DSE variance decomposition, where A-var (green) is the total variance at each scan, D-var (blue) is the variance of the half difference of adjacent scans, S-var (yellow) is the variance of the average of adjacent scans, and E-var is the edge variance at times 1 and T. D-var and S-var for index t (D_t and S_t) sums to A-var between t and $t+1$ ($(A_t+A_{t+1})/2$). Note how the S-var and D-var time series allow insight to the behavior of the total variance: The excursion of A-var around $t = 2$ and $t = 3$ arise from fast variance while the rise for $t \geq 6$ is due to slow variance. For perfectly clean, i.e. independent data, D-var and S-var will converge and each explain approximately half of A-var.

We can further extend this decomposition into global and non-global variance at time point t

$$A_t = A_{\mathrm{G}t} + A_{\mathrm{N}t}, \tag{10}$$

where

$$\begin{aligned} A_{\mathrm{G}t} =& \bar{Y}_t^2, \\ A_{\mathrm{N}t} =& \frac{1}{I}\sum_{i=1}^{I}\left(Y_{it} - \bar{Y}_t\right)^2, \end{aligned} \tag{11}$$

and $\bar{Y}_t = \frac{1}{I}\sum_{i=1}^{I} Y_{it}$ is the global intensity for time t. Analogously, D_t can be decomposed into $D_{\mathrm{G}t}$ and $A_{\mathrm{G}t}$ (as noted in Burgess et al. (2016)), and likewise for S_t and E_t. Creating temporal averages as in Eqn. (7), this likewise extends to a decomposition of global variance into fast, slow and edge components

$$A_{\mathrm{G}} = D_{\mathrm{G}} + S_{\mathrm{G}} + E_{\mathrm{G}}. \tag{12}$$

Table 1 provide the full list of values that make up this decomposition, and indicate how they can be plotted. See Appendix B for options on plotting the global variance decomposition.

This framework also leads to diagnostics in image form. Just as a variance image with voxels $A_i = \sum_t Y_{it}^2/T$ is useful, so could a D-var image, $D_i = \sum_t (Y_{it} - Y_{i,t+1})^2/(4T)$ and a S-var image, $S_i = \sum_t (Y_{it} + Y_{i,t+1})^2/(4T)$ offer more information on the noise structure.

2.3. DSE ANOVA Table & Reference Values

This DSE decomposition can be usefully assembled into a variant of an Analysis of Variance (ANOVA) table that summaries contributions from fast, slow, end, global and non-global components to the total variance in a 4D dataset. Traditionally ANOVA tables use sum-of-squares to partition variance, but we instead focus on root mean squared (RMS) or mean squared (MS) values to leverage intuition on typical noise standard deviation (or variance) of resting state fMRI data. Table 2 shows the values that make up what we call the DSE ANOVA table.

To understand this decomposition we define reference values for "good", artifact-free data using a null model. In Appendix D we detail the most arbitrary version of this model, based only on time-constant spatial covariance, Σ^S, and find expected values for each element of the DSE ANOVA table. More interpretable expected values, however, come from a slightly restrictive model with time-space-separable correlation. This separable noise model assumes data with arbitrary spatial covariance Σ^S but a common temporal autocorrelation for all voxels with a constant lag-1 autocorrelation ρ. While this is less restrictive than an AR(1) model, in real data temporal autocorrelation varies widely over space, and we only consider this as a tractable working model to understand the DSE ANOVA table. (Our null model for DVARS p-values, below, is more realistic). We also consider the idealized model of "perfect" data with completely independent and identically distributed (IID) 4D noise.

Table 3 shows three sets of reference values for the DSE ANOVA table. (Going forward we drop the third row of the DSE ANOVA table showing non-global variance, since in practice the global explains so little variance that the first and third rows are essentially the same; see e.g. Table 6 entries' for A_{G}). The first pair of rows shows the expected value of the MS of each component for the separable noise model. This shows that all variance components scale with the average voxel-wise variance ($\mathrm{tr}(\Sigma^S)/I$, where $\mathrm{tr}(\cdot)$ is the trace), and as temporal autocorrelation increases D-var shrinks and S-var grows. The global components are seen to depend on $\mathbf{1}^\top \Sigma^S \mathbf{1}/I$, the average summed spatial covariance, where $\mathbf{1}$ is a vector of ones. This indicates, intuitively, that the greater the spatial structure in the data the more variance that is explained by the global.

The next pair of rows in Table 3 show the expected MS values normalized to the expected A-var term. The A-var-normalized D-var and S-var diverge from 1/2 exactly depending on ρ, and normalized E-var is $1/T$. The global terms here depend on the balance between average spatial covariance and average variance, $\mathbf{1}^\top \Sigma^S \mathbf{1} / \mathrm{tr}(\Sigma^S)$.

Finally, the final pair of rows shows expected values under the most restrictive case of IID noise. Here D-var and S-var are exactly equal, about 1/2, and we see that the global variance explained should be tiny, $1/I$. This suggests that normalized global variance relative to the nominal IID value, i.e. $(A_G/A)/(1/I)$, is an estimate of $\mathbf{1}^\top \Sigma^S \mathbf{1} / \mathrm{tr}(\Sigma^S)$, a unitless index of the strength of spatial structure in the data. (This particular result doesn't depend on the separable model; see Appendix D).

These reference models provide a means to provide DSE values in three useful forms. For each A-var, D-var, S-var and E-var term we present:

1. RMS, the square root of the mean squared variance quantity,
2. %A-var, a variance as a percentage of total variance A, and
3. Relative IID, A-var-normalized values in ratio to nominal IID values.

For example, for A-var we have (1) RMS is $\sqrt{A}$, (2) %A-var is 100% and (3) relative IID is 1.0. For D-var, (1) RMS is $\sqrt{D}$, (2) %A-var is $D/A \times 100$ and (3) relative IID is

$$\frac{D}{A} \Big/ \frac{1}{2}\frac{T-1}{T} \,. \tag{13}$$

For D_G-var, (2) RMS is $\sqrt{D_G}$, (2) %A-var is $D_G/A \times 100$ and (3) relative IID is

$$\frac{D_G}{A} \Big/ \frac{1}{2}\frac{1}{I}\frac{T-1}{T} \,, \tag{14}$$

noting that we normalize to A and not A_G.

2.4. Inference for DVARS

We seek a significance test for the null hypothesis

$$H_0 : \mathbb{E}(\mathrm{DVARS}_t^2) = \mu_0, \tag{15}$$

where μ_0 is the mean under artifact-free conditions. Note this is equivalent to a null of homogeneity for DVARS_t or D_t. If we further assume that the null data are normally distributed, we can create a χ^2 test statistic

$$X(\mathrm{DVARS}_t) = \frac{2\hat{\mu}_0}{\hat{\sigma}_0^2}\,\mathrm{DVARS}_t^2, \tag{16}$$

approximately following a χ^2_ν distribution with $\nu = 2\hat{\mu}_0^2/\hat{\sigma}_0^2$ degrees of freedom, where σ_0^2 is the null variance (see Appendix E).

What remains is finding estimates of μ_0 and σ_0^2. The null mean of DVARS_t is the average differenced data variance,

$$\mu_0 = \frac{1}{I}\sum_i \sigma_{Di}^2, \tag{17}$$

where σ_{Di}^2 is the variance of the differenced time series at voxel i. To avoid sensitivity to outliers, we robustly estimate each σ_{Di}^2 via the interquartile range (IQR) of the differenced data,

$$\hat{\sigma}_{Di}^2 = \frac{\mathrm{IQR}\left(\{Y_{it} - Y_{i,t+1}\}_{t=1,\ldots,T-1}\right)}{\mathrm{IQR}_0}, \tag{18}$$

where $\mathrm{IQR}_0 = (\Phi^{-1}(0.75) - \Phi^{-1}(0.25)) \approx 1.349$ is the IQR of a standard normal, and $\Phi^{-1}(\cdot)$ is the inverse cumulative distribution function of the standard normal. Below we evaluate alternate estimates of μ_0, including the median of $\{\hat{\sigma}_{Di}^2\}$ and directly as the median of $\{\mathrm{DVARS}_t^2\}$.

The variance of DVARS_t^2 unfortunately depends on the full spatial covariance, and thus we're left to robustly estimating sample variance of $\{\mathrm{DVARS}_t^2\}$ directly. We consider several estimates based on IQR and evaluate each with simulations below. Since the IQR-to-standard deviation depends on a normality assumption, and we consider various power transformations before IQR-based variance estimation (see Appendix F for details). We also consider a "half IQR" estimate of variance

$$\mathrm{hIQR}\left(\{\mathrm{DVARS}_t^2\}_t\right) / \,\mathrm{hIQR}_0, \tag{19}$$

where hIQR is the difference between the median and first quartile, and $\mathrm{hIQR}_0 = \mathrm{IQR}_0\,/2$. This provide additional robustness when more than just the upper quartile of DVARS_t^2 values are corrupted.

Finally, the $X(\mathrm{DVARS}_t)$ values can be converted to p-values $P(\mathrm{DVARS}_t)$ with reference to a χ^2_ν distribution, and subsequently converted into equivalent Z scores,

$$Z(\mathrm{DVARS}_t) = \Phi^{-1}(1 - \mathrm{P}(\mathrm{DVARS}_t)). \tag{20}$$

Note that for extremely large values of DVARS_t numerical underflow will result in p-values of zero; in such cases an approximate Z score can be obtained directly as $Z(\mathrm{DVARS}_t) = (\mathrm{DVARS}_t^2 - \mu_0)/\sigma_0$.

Also note that under complete spatial independence the degrees of freedom will equal the number of voxels I, and so ν can be thought of an effective number of spatial elements; large scale structure will decrease ν while larger ν should be found with cleaner data. Though we caution that estimates of ν will be very sensitive to the particular estimators used for μ_0 and σ_0^2.

2.5. Standardized DVARS

We propose that our D-var time series, $D_t = \mathrm{DVARS}_t^2/4$, is a more interpretable variant of DVARS, as it represents a particular "fast" portion of noise variance, and when added to "slow" variance S_t gives the average variance $A_{t,t+1}$. However, there are various transformations that may be considered better for plotting or reporting (see Table 4 and Figure 4).

In addition to the original DVARS_t and our D_t, we might also consider the percent D-var variance explained at a time point. Eqn. (5) could be used to find, in sums-of-squares units, the percent variance attributable to D-var at t, $t+1$:

$$\frac{I \times D_t}{I \times A_{t,t+1}} 100. \tag{21}$$

However, problem scans can inflate A_t and could mask problem time points. Hence we instead propose to replace $A_{t,t+1}$ with its average A and compute percent D-var as

$$\%D\text{-var}: \ \frac{D_t}{A} 100. \tag{22}$$

This has the merit of being interpretable across datasets, regardless of total variance. As shown in Table 3, IID data have D around half of A, i.e. yield $\%D$-var of 50%.

While $\%D$-var can be more interpretable than unnormalized D-var, its overall mean is still influenced by the temporal autocorrelation. For example, if $\%D$-var is overall around 30% and at one point there is a spike up to 50%, what is interesting is the 20 percentage point change, not 30% or 50% individually. Hence another useful alternative is change in percent D-var

$$\Delta\%D\text{-var}: \ \frac{D_t - \mu_0/4}{A} 100, \tag{23}$$

interpretable as the excess fast variance as a percentage of average variance.

We previously have proposed scaling DVARS relative to its null mean (Nichols, 2013),

$$\mathrm{RDVARS} \ \ \mathrm{DVARS}_t / \sqrt{\mu_0}. \tag{24}$$

(While we had called this "Standardized DVARS", a better label is "Relative DVARS.") This gives a positive quantity that is near 1 for good scans and substantially larger than one for bad ones. However, there is no special interpretation "how large" as the units (multiples of $\mu_0^{-1/2}$) are arbitrary; as noted above, DVARS falls with increased temporal correlation, making the comparison of these values between datasets difficult.

Finally the Z-score $Z(\mathrm{DVARS}_t)$ or $-\log_{10} P(\mathrm{DVARS}_t)$ may be useful summaries of evidence for anomalies.

3. Methods

3.1. Simulations

To validate our null distribution and p-values for DVARS we simulate 4D data as completely independent 4D normally distributed noise

$$Y_{it} \sim \mathcal{N}(0, \sigma_i^2), i = 1, \ldots, I,\ t = 1, \ldots, T, \tag{25}$$

for σ_i drawn uniformly between $\sigma_{\min}$ and $\sigma_{\max}$ for each i, $I = 90,000$.

We manipulate two aspects in our simulations, time series length and heterogeneity of variance over voxels. We consider T of 100, 200, 600 and 1200 data-points, reflecting typical lengths as well as those in the Human Connectome Project. We use three variance scenarios, homogeneous with $\sigma_{\min} = \sigma_{\max} = 200$, low heterogeneity $\sigma_{\min} = 200$ and $\sigma_{\max} = 250$, and high heterogeneity $\sigma_{\min} = 200$ and $\sigma_{\max} = 500$.

We consider four estimates of μ_0. First is the very non-robust sample mean of $\{\text{DVARS}_t^2\}$, denoted $\hat{\mu}_0^{\text{DVARS}}$, considered for comparative purposes. The next two are based on the IQR-based estimate of voxel-wise variance of the differenced data, Eqn. (18), considering the mean $\hat{\mu}_0^D$ and median $\tilde{\mu}_0^D$ of the robust variances $\hat{\sigma}_{Di}^2$. Finally we also consider the empirical median of $\{\text{DVARS}_t^2\}$, $\tilde{\mu}_0^{\text{DVARS}}$. For σ_0^2 all estimates were based directly on $\{\text{DVARS}_t^2\}$; for comparative purposes we considered the (non-robust) sample variance of $\{\text{DVARS}_t^2\}$, $\hat{\sigma}_0^2$, and IQR-based and hIQR-based estimates of variance with power transformations d of 1, 1/2, 1/3 and 1/4, denoted generically $\tilde{\sigma}_0^2$; note $d = 1/3$ is theoretically optimal for χ^2 (see Appendix F).

For p-value evaluations, we only evaluate the most promising null moment estimators, $\tilde{\mu}_0^D$ and $\tilde{\mu}_0^{\text{DVARS}}$ for μ_0, and $\tilde{\sigma}_0^2$ with hIQR, $d = 1$ and hIQR, $d = 1/3$. We measure the bias our estimators in percentage terms, as $(\hat{\mu}_0 - \mu_0)/\mu_0 \times 100$ and $(\hat{\sigma}_0^2 - \sigma_0^2)/\sigma_0^2 \times 100$, where the true value are $\mu_0 = 2\sum_i \sigma_i^2/I$ and $\sigma_0^2 = 8\sum_i \sigma_i^4/I^2$ (as per Appendix E when $\Sigma^S = I$).

For each setting we use 1,000 realisations. We obtain P-values from each method and validate them via log P-P plots (probability-probability plots) and histograms of approximate Z-scores.

3.2. Real Data

We use two publicly available data-sets to demonstrate the results of methods proposed in this paper on real-data. First, we drew 20 healthy subjects at random from the Human Connectome Project (HCP,S900 release). We chose this dataset due to the high quality and long sessions of the data (Smith et al., 2013; Glasser et al., 2013). Second, we used first 25 healthy subjects from the New York University (NYU) cohort of the Autism Brain Imaging Data Exchange (ABIDE) consortium via Preprocessed Connectome Project (PCP) (Craddock et al., 2013). We selected this cohort for its high signal-to-noise ratio and the more typical (shorter) time series length (Di Martino et al., 2014).

3.2.1. Human Connectome Project Data

For full details see (Van Essen et al., 2013; Glasser et al., 2013); in brief, 15 minute eyes-open resting acquisitions were taken on a Siemens customized Connectome 3T scanner with a gradient-echo EPI sequence, TR=720ms, TE=33.1 ms, flip angle=52° and 2 mm isotropic voxels. For each subject, we used the first session, left to right phase encoding direction (See Supplementary Table S1 for full details of subjects). We considered each subject's data in three states of pre-processing: unprocessed, minimally pre-processed and ICA-FIXed processed. Unprocessed refers to the raw data as acquired from the machine without any pre-processing step performed, useful as a reference to see how the variance components change with preprocessing steps. Minimally pre-processed data have undergone a range of conventional pre-processing steps such as correction of gradient-nonlinearity-induced distortion, realignment aiming to correct the head movements, registration of the scans to the structural (T1w) images and finally transformation of the images to the MNI standard space.

Finally, an ICA-based clean up (Salimi-Khorshidi et al., 2014) is applied, where artifactual ICA components, such as movement, physiological noises of the heart beat and respiration, are regressed out the data. Due to extent of the FIX denoising and an ongoing debate regarding the nature of the global signal, we did not consider global signal regression with the HCP data. From now on, we call this stage 'fully pre-processed' to be consistent with the ABIDE-NYU cohort we describe in the following.

3.2.2. ABIDE - New York University Data

For full details visit PCP website `http://preprocessed-connectomes-project.org/`; in brief, 6 minute eyes-closed resting acquisitions were taken on an Allegra 3T scanner with a gradient echo EPI sequence, TR=2000ms, TE=15ms, flip angle=90°, and 3 mm isotropic voxels (See Supplementary Table S2 for full details of subjects). In this study, each subject was analyzed using Configurable Pipeline for the Analysis of Connectomes (C-PAC) pipeline, in three stages; unprocessed, minimally pre-processed and fully pre-processed. The unprocessed data are raw except for brain extraction with FSL's BET. Minimally pre-processed data were only corrected for slice timing, motion by realignment and then the data were transformed into a template with 3mm isotropic voxels. Fully pre-processed data additionally had residualisation with respect to 24-motion-parameters, signals from white matter (WM) and cerebrospinal fluid (CSF), and linear and quadratic low-frequency drifts. Conventionally this pipeline deletes the first three volumes to account for T1 equilibration effects, but we examine the impact of omitting this step for the raw data.

4. Results

4.1. Simulations

Figure 2 shows the percentage bias for the null expected value μ_0 (left panel) and variance σ_0^2 (right panel) for different levels of variance heterogeneity and time series length. In general, the bias for both

parameters is modest, increasing with greater heterogeneity and decreasing with growing T.

The direct estimates of the mean based on the DVARS_t^2 time series perform best on this clean, artifact-free data, while mean estimated based on voxel-wise median difference variances $\tilde{\mu}_0^D$ degrades the most with increasing heterogeneity. The estimates of variance have relatively less bias but it is difficult to identify one particular best method, save for IQR often (but not always) having less bias than hIQR, and lower d generally associated with less bias.

On balance, given the generally equivocal results and concerns about robustness, for further consideration we focused on $\tilde{\mu}_0^{\mathrm{DVARS}}$ (median of $\{\mathrm{DVARS}_t^2\}$) and $\tilde{\mu}_0^D$ (median of $\hat{\sigma}_{Di}^2$) as promising candidates for μ_0, and hIQR with $d = 1$ and hIQR with $d = 1/3$ for σ_0^2.

Figure 3 shows log P-P plots for χ^2 p-values and histograms of approximate Z scores, $(\mathrm{DVARS}_t^2 - \mu_0)/\sigma_0$; values above the identity in the P-P plot correspond to valid behavior. While all methods have good performance under homogeneous data, $\tilde{\mu}_0^D$ (panels A & C) is not robust to variance heterogeneity and results in inflated significance. In contrast, $\tilde{\mu}_0^{\mathrm{DVARS}}$ (panels B & D) has good performance over all, for variance estimated with either $d = 1$ or $d = 1/3$ (top and bottom panels, respectively), and also yields good approximate Z-scores.

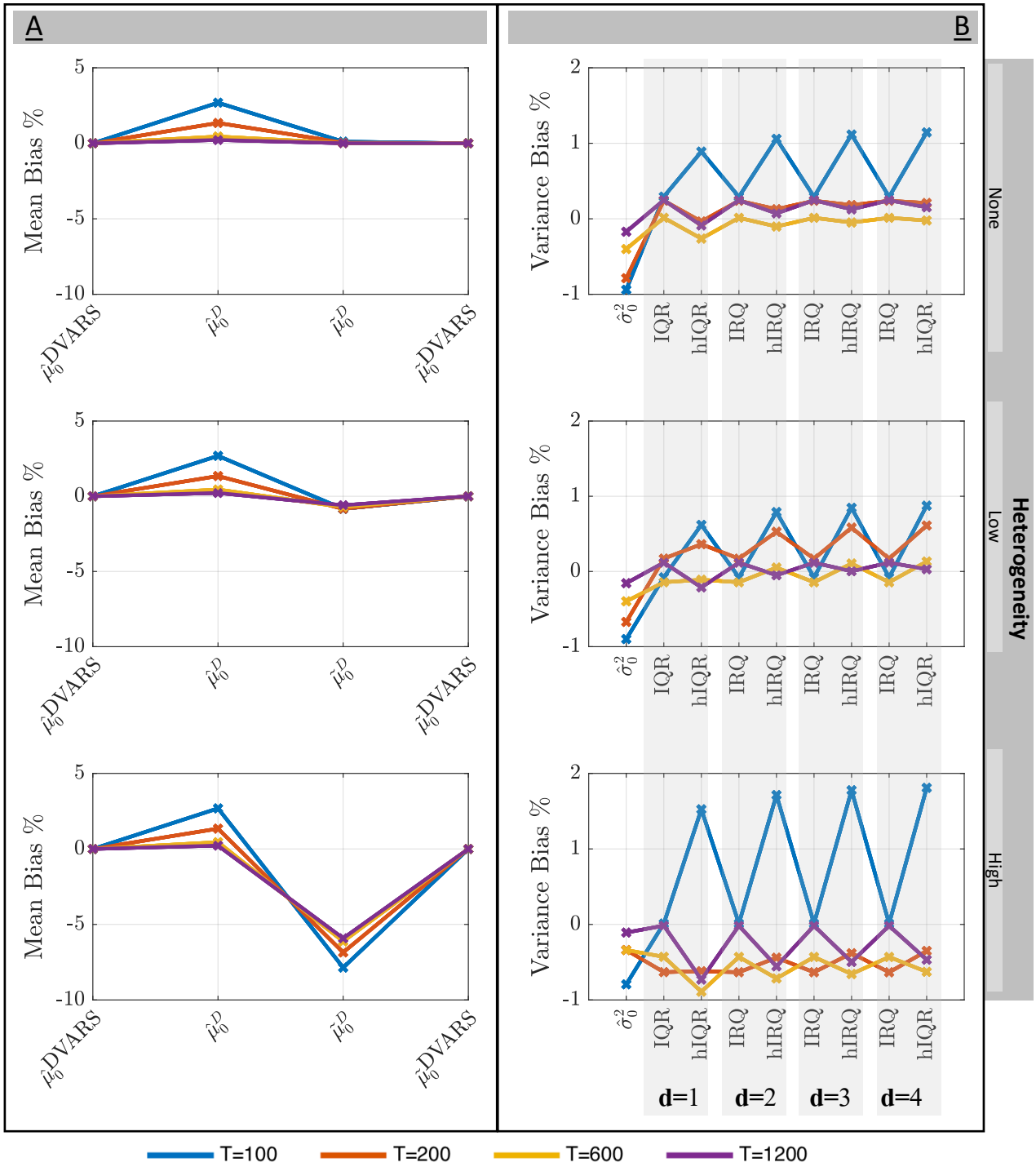


Figure 2: Simulation results for estimation of DVARS^2 null mean μ_0 (left panel) and variance σ_0^2 (right panel), for no, low and high heterogeneity of variance over voxels all estimates (rows). All estimators improve with time series length T and most degrade with increased heterogeneity. Both the sample mean ($\hat{\mu}_0^{\mathrm{DVARS}}$) and median ($\tilde{\mu}_0^{\mathrm{DVARS}}$) of DVARS$_t^2$ perform best, as does voxel-wise median of difference data variance ($\hat{\mu}_0^{\mathrm{DVARS}}$) for sufficient T, though $\hat{\mu}_0^{\mathrm{DVARS}}$ lacks robustness. No one variance estimator dominates.

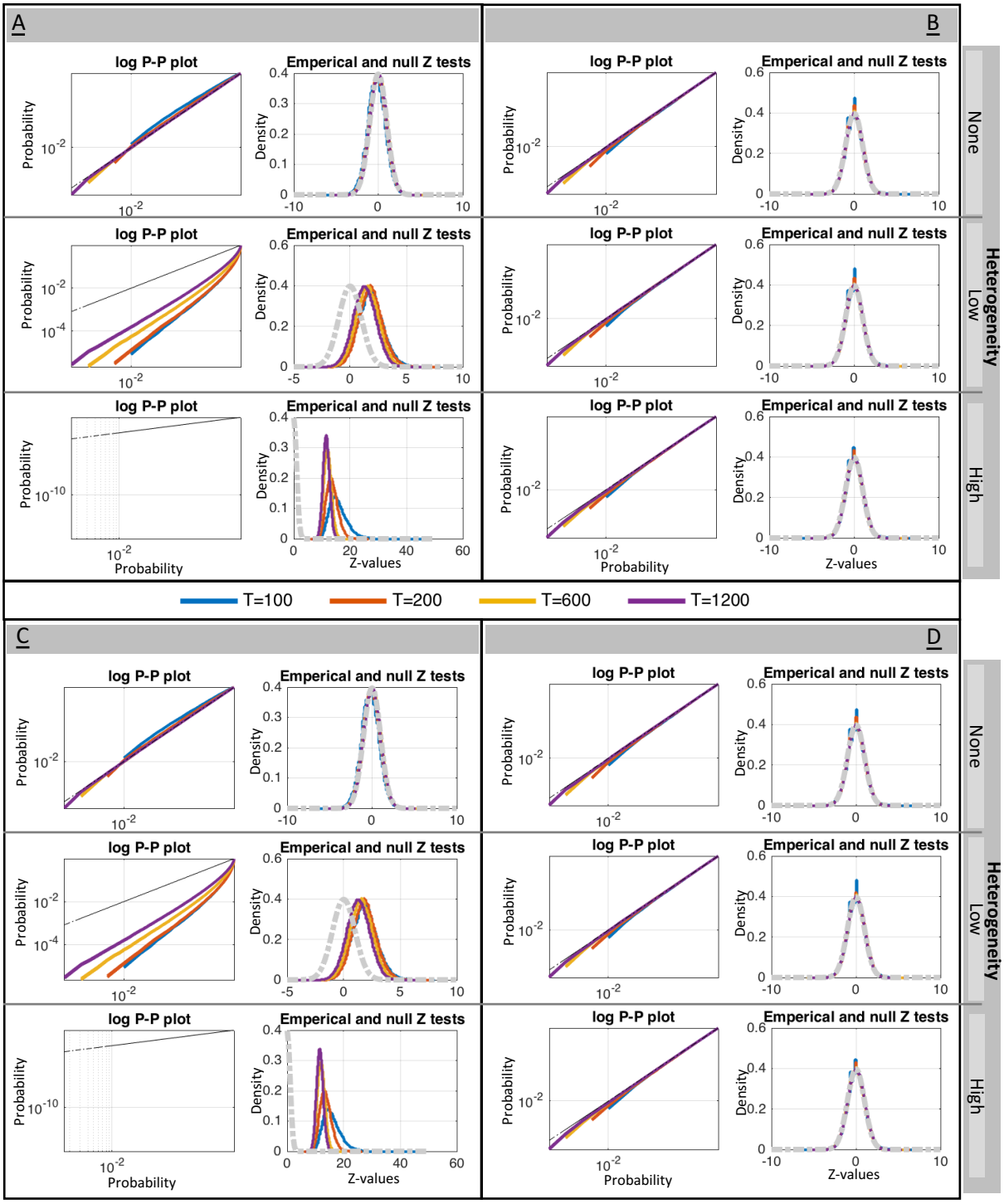


Figure 3: Simulation results for validity of DVARS p-values for different estimators of μ_0 and σ_0^2. The left two panels (A & C) use $\tilde{\mu}_0^D$, the two right panels (B & D) use $\tilde{\mu}_0^{\mathrm{DVARS}}$; the upper two panels (A & B) use variance based on hIQR with $d = 1$, the lower two panels (C & D) use hIQR with $d = 1/3$. P-P plots and histograms of approximate Z scores show that only use of $\tilde{\mu}_0^{\mathrm{DVARS}}$ gives reliable inferences, and power transformation parameter d seems to have little effect.

On the basis of these results, we elected to use $\tilde{\mu}_0^{\mathrm{DVARS}}$ as the only reliable option for the mean, and hIQR, $d = 1/3$ as a variance estimate.

4.2. Real Data

We first focus on selected results of two HCP subjects, then later summarize results for all HCP and ABIDE subjects.

4.2.1. Temporal Diagnostics: Comparison of DVARS measures

Figure 4 shows different DVARS-type measures for subject 118730 of the HCP cohort. The first six plots corresponds to the variants listed in Table 4; the bottom two plots show "DSE plots," plots of A_t, D_t, S_t and E_t variance components, upper plot with minimal pre-processing, lower with full pre-processing. The grey stripes indicate 19 data points identified as having significant DVARS after Bonferroni correction. Supplementary Table S3 shows values for all significant scans. We also show the results of this analysis for three more subjects (HCP subject 115320, ABIDE-NYU subject 51050 and 51055) in Supplementary Figures S1, S4 and S7.

In Figure 4, the largest D_t occurs at index 7 (i.e. 7th and 8th data points) where $\sqrt{D_t}$ is 4.074, corresponding to 70.155% of average variance and far in excess of the nominal IID value of 50%. The equivalent Z score is 36.330 and only indicates extreme statistical significance, while a more meaningful $\Delta\%D$-var of 41.198% indicates practical significance in terms of excess variance at this point. The least significant D_t occurs at index 726, with a Z score of 4.36 and $\sqrt{D_t}$ is 5.66; here a $\Delta\%D$-var of 4.95% indicates this is a relatively modest disturbance. In contrast, values of original DVARS or relative DVARS do not offer a meaningful interpretation.

The bottom panel of Figure 4 shows the DSE plot for fully pre-processed data. This data now exhibits the idealized behavior of IID data, with D-var and S variance components converging at 50% of average variance (see right-hand y-axis). Note how $\sqrt{D_t}$ is around 2.6 before clean up, and 2.5 after clean up, while $\sqrt{S_t}$ fell dramatically with cleaning, indicating that nuisance variance removed was largely of a "slow" variety. Also observe that clean up results in drops in total A_t variance where artifacts were observed, indicating variance removed by the regression procedure.

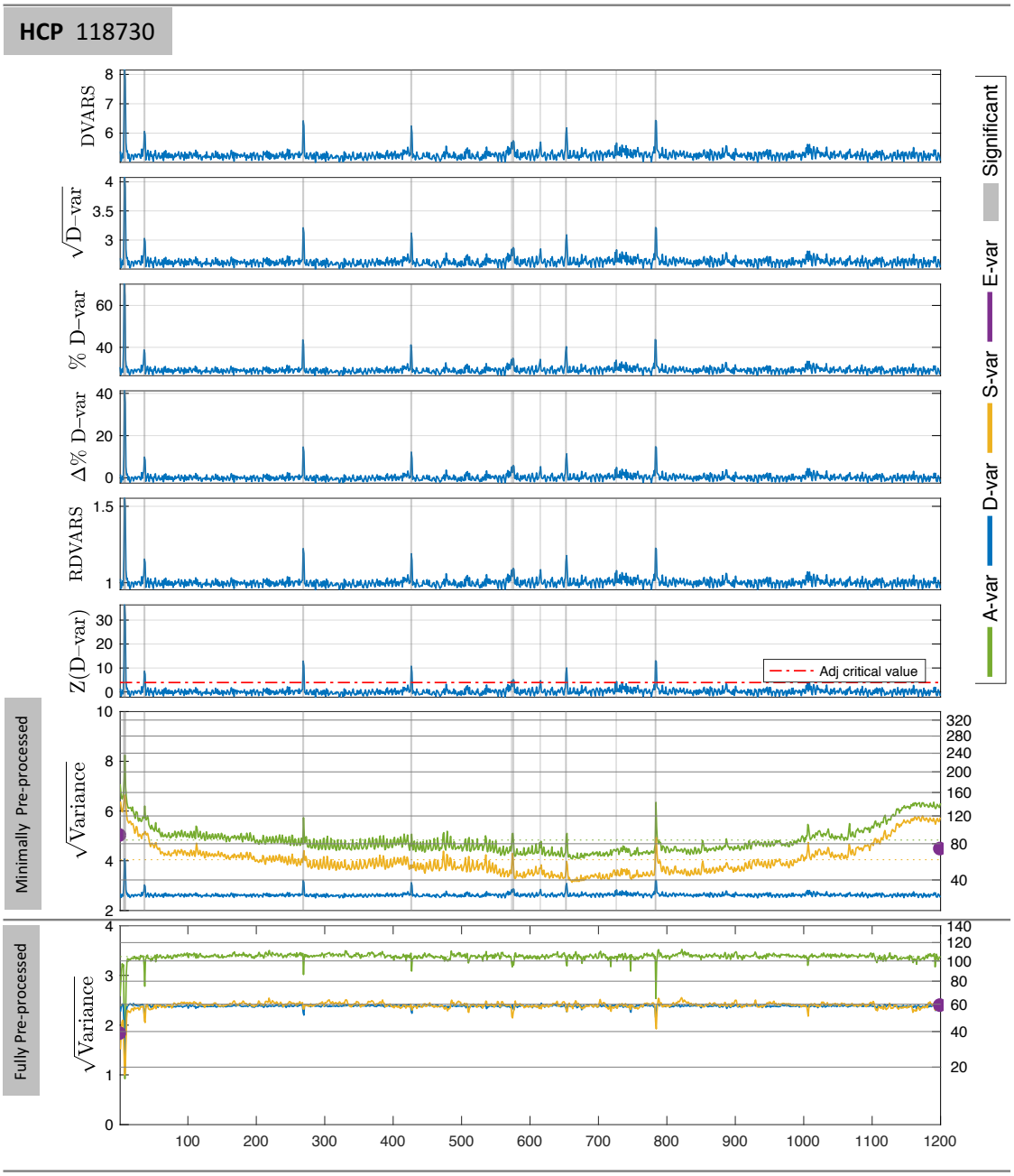


Figure 4: Comparison of different variants of DVARS-related measures on HCP 115320. The first six plots are variants of DVARS listed in Table 4; the Z(DVARS) plot shows the one-sided 5% Bonferroni significance threshold for 1200 scans, with vertical grey lines marking these significant scans. The bottom two plots show all 4 variance components, total A_t (green), fast D_t (blue), S_t slow (yellow), and edge E_t (purple), for minimally preprocessed (upper) and fully preprocessed (lower) data. For minimally preprocessed data D-var is about 25% of A-var (see right axis), far below S-var. For fully preprocessed data D-var and S-var converge to 50%A-var.

4.2.2. Temporal Diagnostics: Before and after clean-up

Figures 5 and 6 shows the minimally and fully pre-processed variance decompositions, respectively, of HCP subject 115320.

Figure 5, upper panel, shows that if the strict FD threshold were used 47% of scans would be flagged, while the lenient threshold appears to miss several important events. For example, around scans 775 and 875 there are two surges in $\sqrt{D_t}$, rising to about 60% and 40% average variance (excesses of 30% and 10%, respectively, from a baseline of about 30%) while FD remains low. The lower panel's pie chart shows that slow S-var explains just under 75% of total variance, and almost all of global variance; edge variance is also 1.5 above expected.

In Figure 6, the fully preprocessed data-set shows roughly equal of fast and slow variance, as reflected in the overlapping D_t and S_t time series (blue and yellow, respectively) and the pie and bar charts for total variance. Edge E-var has also dropped to fall in line with IID expectations. This convergence, however is not homogeneous over scans, and excursions of S-var are still found after scan 650. However, these are much reduced excursions of S_t (no more than 75% of average variance, compared to over 150% in Fig. 5).

Note that while significant DVARS are found, they are small in magnitude: Table 5 lists the 10 significant tests, none with $\Delta\%D$-var greater than 6%. If we used a $\Delta\%D$-var of 5% we would still mark 4 of these 10 significant; while we might hope for better performance from the FIX method, note the severe problems detected towards the end of the scan (Fig. 5).

The smallest significant $\Delta\%D$-var is 2.66%, which is smaller than the least significant scan detected in the minimally preprocessed data, 3.78%. This indicates the increased sensitivity in our procedure as the background noise in the data is reduced.

It is also worth noting that majority of the spikes detected in Figure 5 has been removed by motion correction algorithms (i.e. in this specific case, ICA-FIX; shown in Figure 6), however these algorithm may leave negative spikes which have to be detected via two-sided test and be removed.

Temporal diagnostics of before and after clean-up for three other subjects (HCP subject 118730, NYU-ABIDE subjects 51050 and 51050) also reported in Supplementary Materials. See Figure S2 and S3 for HCP subject 118730, Figure s5 and s6 for NYU-ABIDE subject 51050 and Figure s8 and s9 for NYU-ABIDE 51055.

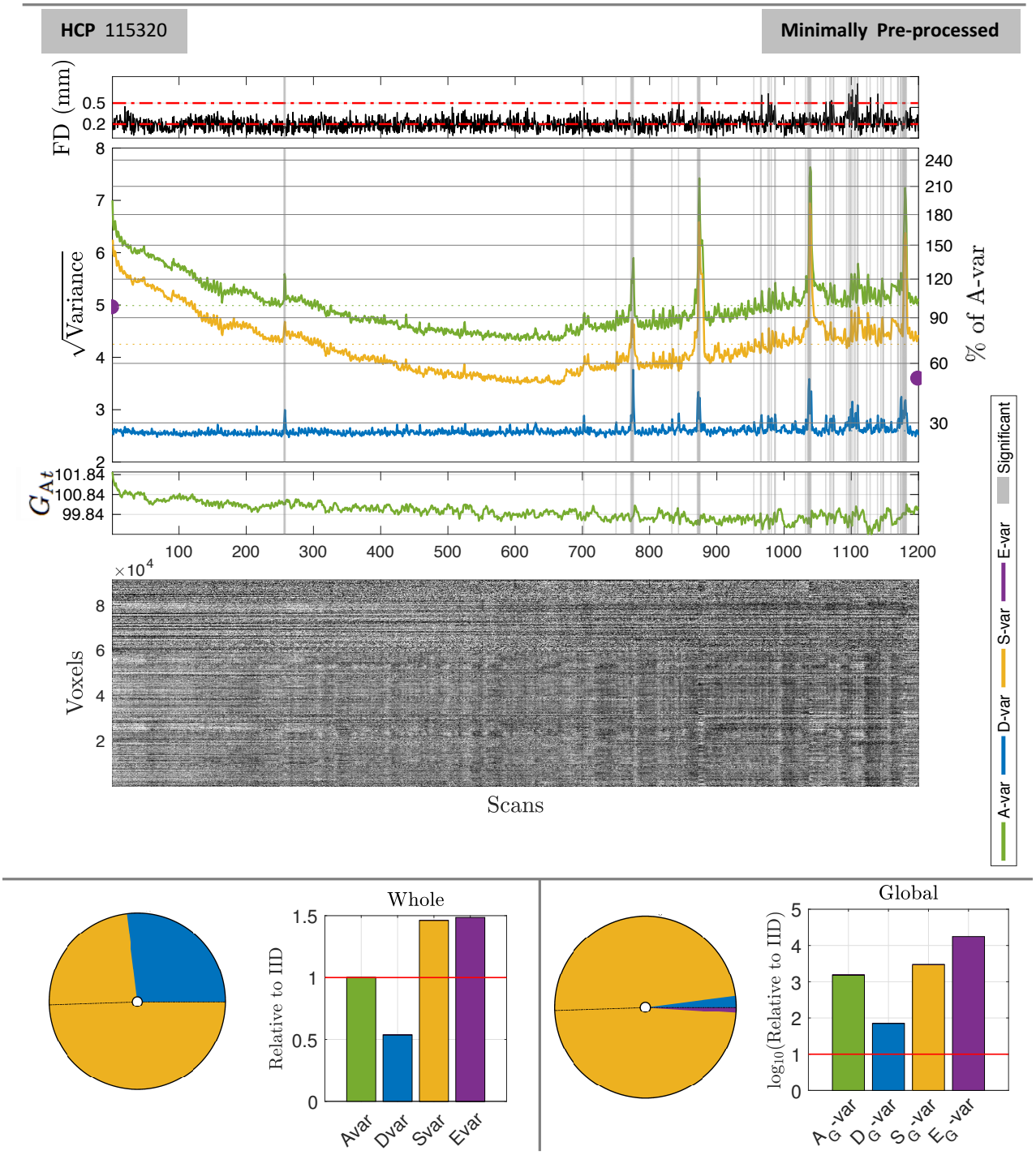


Figure 5: Illustration of DSE and DVARS inference for HCP 115320 minimally pre-processed data. The upper panel shows four plots, framewise displacement (FD), the DSE plot, the global variance signal $G_{\mathrm{A}t}$, and an image of all brainordinate elements. FD plots show the conventional 0.2mm and 0.5mm, strict and lenient thresholds, respectively. The bottom panel summaries the DSE ANOVA table, showing pie chart of the 4 variance components and a bar chart relative to IID data, for whole (left) and global (right) components. Many scans are marked as significant, reflecting disturbances in the latter half of the acquisition.

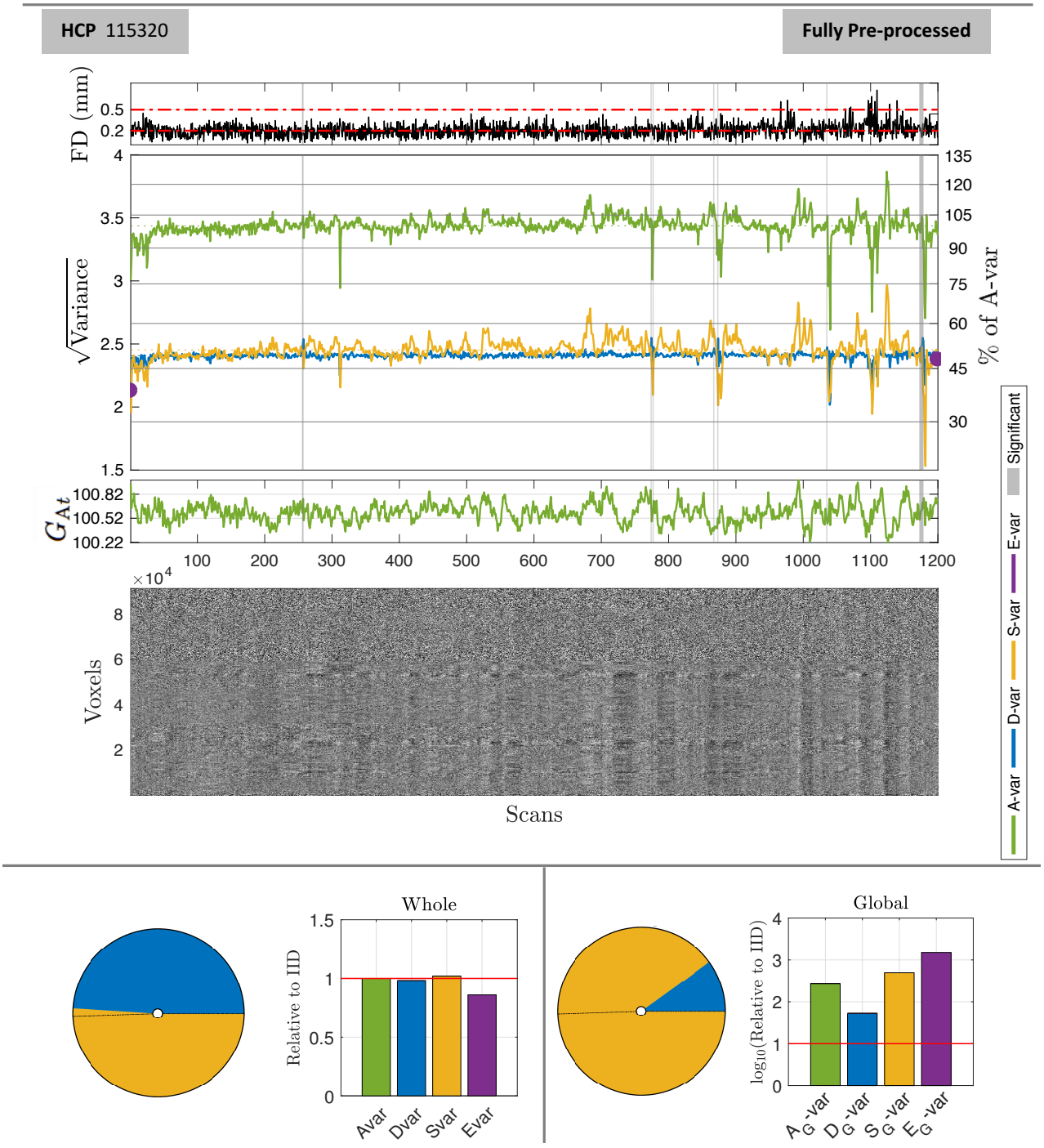


Figure 6: Illustration of DSE and DVARS inference for HCP 115320 fully pre-processed. Layout as in Fig. 5. Cleaning has brought S_t slow variance into line with D_t fast variance, each explaining about 50% of total variance. While some scans are still flagged as significant, $\%D$-var (right y axis) never rises above about 55%, indicating $\Delta\%D$-vars of 5% or less and possible lack of practical significance.

Table 5: List of all statistically significant D_t fast variance components in the fully pre-processed HCP 115320. Spikes which represent the highest (index 1177) and lowest (index 1035) was bold.

Scan	Index	DVARS	$\sqrt{\text{D-var}}$	%D-var	Δ%D-var	RDVARS	Z(D-var)	FD
256 & 257	256	4.982	2.4910	52.519	3.362	1.038	5.093	0.136
257 & 258	257	5.077	2.538	54.553	5.397	1.058	8.175	0.172
774 & 775	774	5.095	2.547	54.935	5.779	1.062	8.753	0.290
777 & 778	777	4.955	2.477	51.950	2.794	1.033	4.232	0.247
873 & 874	873	5.089	2.544	54.805	5.649	1.061	8.556	0.255
1035 & 1036	**1035**	**4.948**	**2.474**	**51.815**	**2.659**	**1.031**	**4.027**	**0.280**
1175 & 1176	1175	4.960	2.480	52.062	2.905	1.034	4.401	0.109
1176 & 1177	1176	4.953	2.476	51.926	2.769	1.032	4.195	0.104
1177 & 1178	**1177**	**5.096**	**2.548**	**54.964**	**5.807**	**1.062**	**8.796**	**0.301**
1178 & 1179	1178	5.049	2.524	53.952	4.795	1.052	7.263	0.132

The DSE ANOVA tables for minimally and fully preprocessed (Table 6) gives concise summaries of the data quality. The RMS values provide concrete values that can be used to build intuition for data from a given scanner or protocol. The total noise standard deviation falls from 5.015 to 3.437 with clean-up, but it is notable that the fast variance D-var falls only slightly from 2.598 to 2.406 (in RMS units), while slow variance falls dramatically from about 4.287 to 2.454. This indicates that much of the variance reduction in "cleaning" comes from removal of low frequency drifts and other slowly-varying effects. The magnitude of temporally structured noise is reflected by S-var explaining 73% of total variance, and after clean-up S-var and D-var fall into line around 50%. A measure of the spatially structured noise is the global A_{G}-var that, while small as a percentage, is seen to be about 1,500 that expected with IID before preprocessing, and falling to about 275 relative to IID after preprocessing. That the majority of A_{G}-var is due to S_{G}-var indicates that the global signal is generally low frequency in nature.

We also show the DSE ANOVA tables for three other subjects; HCP subject 118730 in Table S4, NYU-ABIDE subject 51050 and 51055 in Tables S5 and S6, respectively.

Table 6: DSE ANOVA Tables for HCP 115320. Minimally preprocessed data (top), fully preprocessed (bottom) are readily compared: Overall standard deviation drops from 5.015 to 3.437, while fast noise only reduces modestly from 2.598 to 2.406, indicating preprocessing mostly affects the slow variance. The IID-relative values for D, S and E for the fully preprocessed data are close to 1.0, suggesting successful clean-up in the temporal domain; the global signal, however, still explains about 275× more than would be expected under IID settings, indicating (likely inevitable) spatial structure in the cleaned data.

Minimally Preprocessed Data

Source	RMS	% of A-var	Relative to IID
A - All	5.015	100.000	1.000
D - Fast	2.598	26.837	0.537
S - Slow	4.287	73.039	1.462
E - Edge	0.176	0.124	1.486
A_{G} - All Global	0.415	0.684	1539.383
D_{G} - Fast Global	0.063	0.016	71.126
S_{G} - Slow Global	0.408	0.662	2,980.787
E_{G} - Edge Global	0.040	0.006	17,636.960

Fully Preprocessed Data

	RMS	% of A-var	Relative to IID
A - All	3.437	100.000	1.000
D - Fast	2.406	48.980	0.980
S - Slow	2.454	50.948	1.020
E - Edge	0.092	0.072	0.860
A_{G} - All Global	0.120	0.122	274.058
D_{G} - Fast Global	0.037	0.012	52.830
S_{G} - Slow Global	0.114	0.109	493.227
E_{G} - Edge Global	0.008	<0.001	1,508.473

Figure 7 illustrates the use of the DSE decomposition to summarize a group of subjects. The top shows the 20 HCP subjects, the bottom the 25 ABIDE subjects, with left showing total variance decomposition, right the decomposition for the global signal. For the HCP raw data, the D-var component ranges from just over 5% to 40%, successively converging to 50% with preprocessing. For the ABIDE data, failing to remove initial T1-saturated scans is immediately evident with edge variance E taking a large portion of variance from S, and the global signal unusually explaining well over 5% of variance. With the initial 3 scans removed, the usual pattern of D and S explaining most of the variance is seen.

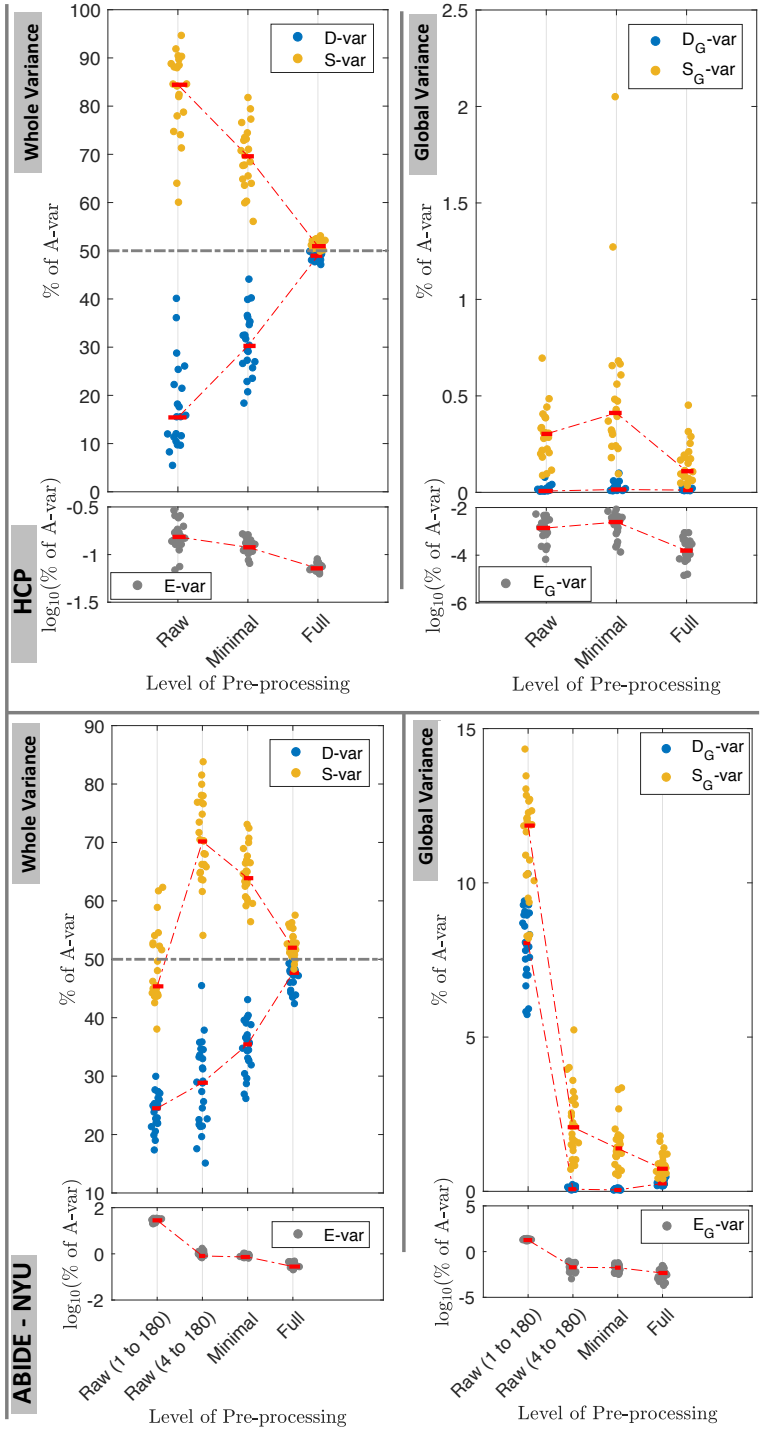


Figure 7: Illustration of group level contribution of each variance component for 20 HCP subjects (top) and 25 ABIDE subjects (bottom). The left panels show the total DSE %A-var values for whole variance, the right panels show the same for global variance. Successive preprocessing moves fast and slow variance components to equal proportions. For ABIDE, when the first 3 T1-saturated scans are not removed ("Raw (1 to 180)"), a large excess is seen on E-var (at the expense of S-var) and the global, both slow and fast components, explaining an unusually large portion of variance.

Figure 8 shows the image-wise D-var and S-var components, visualized as RMS. Like when observed in time-series form, the D-var images reflects a generally homogeneous 'noise floor', with CSF and major vasculature visible above this floor. The S-var image shows more structure, around the edges, and throughout the brain.

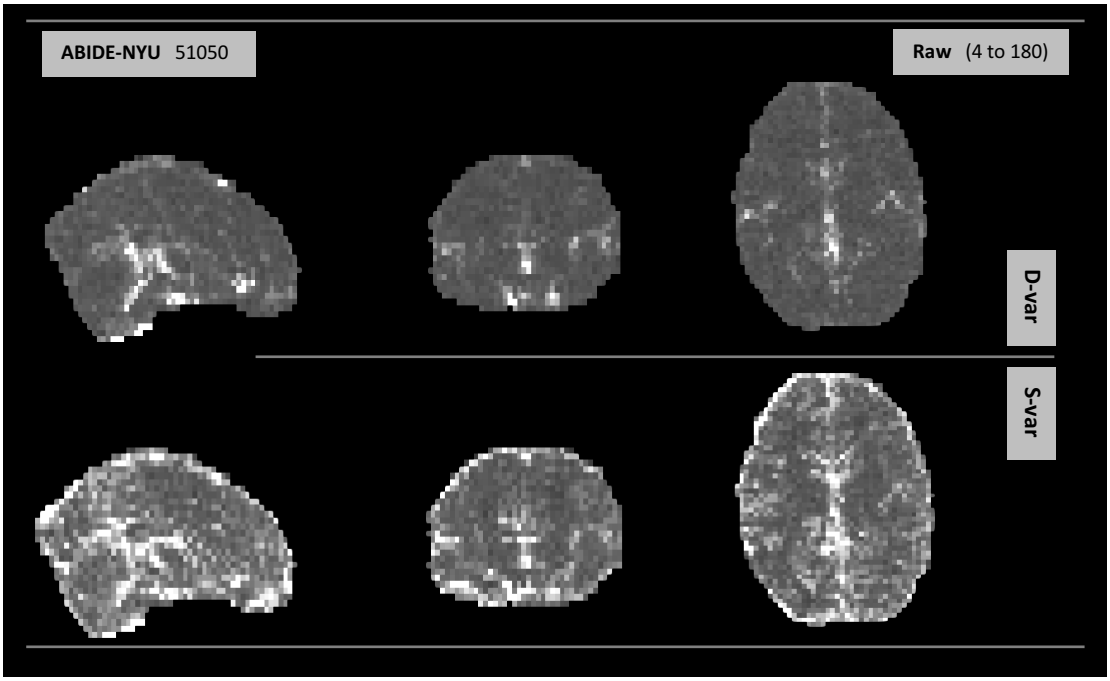


Figure 8: Fast D-var and slow S-var variance components computed voxelwise. The D-var image (top) shows a homogeneous appearance with vascular noise sources apparent, while S-var has appreciable edge components and is much more heterogeneous.

Finally, Table 7 explores the use of the estimated χ^2 degrees of freedom ν as an index of spatial effective degrees of freedom. Raw data, exhibiting substantial spatial structure, has $\nu = 287$, which increases to $\nu = 11,086$ for fully preprocessed data, still only about 5% of the actual number of voxels.

Table 7: Spatial effective degrees of freedom (EDF) for HCP subject 115320. As more spatial structure is removed with preprocessing, spatial EDF rises, but never to more than 5% of the actual number of voxels.

	Voxels	Spatial EDF	Spatial EDF / Voxels
Raw	162,768	287	0.176%
Minimally processed	224,998	1,660	0.738%
Full processed	224,998	11,086	4.928%

5. Discussion

We have provided a formal context for the diagnostic measure DVARS, showing DVARS_t^2 to be part of a decomposition of variance at each successive scan pair and over the whole 4D data. We have proposed DSE plots and DSE ANOVA tables that concisely summarize the interplay of the fast, slow, total and global variance, finding null expected values for each table entry. Our analysis shows that D-var (and DVARS) scales with overall noise variance, and is deflated by temporal autocorrelation. The DSE plots allow D-var to be judged relative S-var, checking for convergence to 50% of A-var as data approaches independence. We have found the null distribution and a practical null hypothesis testing procedure for DVARS. We complement the statistical significance of DVARS p-values with the practical significance of $\Delta\%D$-var. We illustrated these tools on exemplar HCP subjects, and used the ABIDE cohort to show how D and S converge with successive clean-up, and how E var can usefully detect T1 saturation effect when initial scans are not discarded.

5.1. Limitations

Our DVARS p-values depend critically on accurate estimates of μ_0 and σ_0^2. Despite finding exact expressions for the null mean and variance, we found the most practical and reliable estimates to be based on the sample DVARS_t^2 time series itself, using median for μ_0 and hIQR to find σ_0 (essentially identical results were found with $d = 1$). Of course this indicates that our inference procedure can only infer relative to the background noise level of the data, picking out extreme values that are inconsistent with our approximating χ^2 approximation.

We observe that as data becomes cleaner, and the background noise falls, we have greater power to identify extreme DVARS_t^2 values. This is a limitation that simply highlights the need for measures of practical significance, which provide with $\Delta\%D$-var.

The effective spatial degrees of freedom may prove to be a useful index of spatial structure in the data, but we stress this particular χ^2 degrees-of-freedom ν is specific to this setting and is unlikely to be useful in other contexts (e.g. as a Bonferroni correction over space).

Finally we do not suggest that our results here solve the fMRI diagnostic problem, nor are we enthusiastic advocates of scrubbing, removing and interpolating problem scans. Rather we have sought and we believe found greater insight into the behavior of this widely used fMRI diagnostic measure.

Software and Reproducibility

In this work majority of the analysis have been done on MATLAB 2015b and MATLAB 2016b, supported by FSL 5.0.9 for neuroimaging analysis.

Inference on DVARS as well as DSE variance decomposition techniques proposed in this paper is available via MATLAB scripts, found at `http://www.github.com/asoroosh/DVARS`. Elements of these methods have also been implemented in Python and is accessible via Nipype toolbox (Gorgolewski et al., 2011).

Results and figure scripts presented in this work is publicly available on `http://www.github.com/asoroosh/DVARS_Paper17`.

Acknowledgements

We thank Jonathan Power, Matt Glasser, Deanna Barch and Steve Peterson for feedback on the original standardized DVARS work, Javier Gonzalez-Castillo for input on interpreting our DVARS inferences, Abraham Snyder for input on history of DVARS and Michael Harms for his useful comments on the final draft of the manuscript.

Data were provided in part by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

Appendix A. DVARS History

As far as we are aware, DVARS was first used to compute frame censoring by Smyser et al. (2011). Power et al. 2012 reported the first systematic analysis of DVARS in relation to FD in resting state fMRI. However, at least as early as 2006, a web page at the Cambridge Cognitive Brain Unit maintained by Matthew Brett's titled "Data Diagnostics" offered tsdiffana.m, a Matlab script that produces the same measure (see `http://imaging.mrc-cbu.cam.ac.uk/imaging/DataDiagnostics`; when viewed on 28 October, 2012, the page listed the "last edited" data as 31 July 2006) and there are likely earlier uses in fMRI.

The idea of working with differences dates to at least 1941 in the statistics literature in work John van Neumann and colleagues (von Neumann et al., 1941). That work focused on estimation of "standard deviation from differences" when the mean slowly varied from observation to observation. They point out that the idea can traced back further, as early as 1870. In signal processing this estimator can be known as the Allan variance, developed as a robust variance estimator in the presence of 1/f noise (Allan, 1966). And in cardiology the "root mean square successive difference" is a standard measure of heart period variability (Berntson et al., 2005). For yet more background see Kotz et al. (1988).

Despite successive work on finding the exact distribution of this variance estimate (Harper, 1967), or using it in a test for the presence of autocorrelation (Cochrane and Orcutt, 1949), we are unaware of any study of the distribution of the individual differences averaged over a multivariate observation, as is the case in this fMRI application.

Appendix B. Plotting the global variance decomposition

The global variance components, at each time point, are just a single scalar value squared. Thus they may be more intuitively plotted in a signed RMS form. For example, instead of plotting variance A_{Gt}, D_{Gt} and S_{Gt}, the signed quantities

$$\begin{aligned} G_{\mathrm{A}t} =& \bar{Y}_t \\ G_{\mathrm{D}t} =& (\bar{Y}_t - \bar{Y}_{t+1})/2 \\ G_{\mathrm{S}t} =& (\bar{Y}_t + \bar{Y}_{t+1})/2 \end{aligned} \tag{B.1}$$

can be plotted. These set of three time series may seem arbitrary, but have the feature of the sum of squares of $G_{\mathrm{D}t}$ and $G_{\mathrm{S}t}$ sum to the mean-square $G_{\mathrm{A}t}$ and $G_{\mathrm{A}t,t+1}$.

Appendix C. Derivation of DSE variance decomposition

The decomposition of the average variance at time t and $t+1$, Eqn. (5), is based on a simple algebraic identity; for variables a and b,

$$a^2 + b^2 = \frac{1}{2}(a-b)^2 + \frac{1}{2}(a+b)^2. \tag{C.1}$$

This justifies a decomposition of the average variance at each voxel i, for each time $t = 1, \ldots, T-1$,

$$\frac{Y_{it}^2 + Y_{i,t+1}^2}{2} = \left(\frac{Y_{i,t+1} - Y_{it}}{2}\right)^2 + \left(\frac{Y_{it} + Y_{i,t+1}}{2}\right)^2. \tag{C.2}$$

Averaging this expression over voxels $i = 1, \ldots, I$ gives the decomposition for scan pair variance $A_{t,t+1}$ in Eqn. (5). Summing image variance $A_{t,t+1}$ over t, however,

$$\begin{aligned} \sum_{t=1}^{T-1} A_{t,t+1} =& \sum_{t=1}^{T-1} (A_t + A_{t+1})/2 \\ =& \frac{1}{2}A_1 + \sum_{t=2}^{T-1} A_t + \frac{1}{2}A_T \end{aligned} \tag{C.3}$$

misses 1/2 of edge terms, which are added to produce the fundamental DSE decomposition in Eqn. (9).

Appendix D. Derivation of DSE ANOVA Mean Squares

Here we set out the least restrictive model possible to justify our expected values for the DSE ANOVA table (Table 2). While the DSE ANOVA table and decompositions $A = D + S + E$ and $A_G = D_G + S_G + E_G$ are in mean-square (MS) units, below we develop the results in terms of sum-of-squares (SS) that, in each case, can be divided by $I \times T$ to obtain the MS.

All of the results follow from application of rules for expectations and variances of quadratic forms of mean zero vectors. For reference, if w is a mean zero random vector with covariance Σ, and B is a square matrix, then $\mathbb{E}(w^\top B w) = \mathrm{tr}(B\Sigma)$ and $\mathbb{V}(w^\top B w) = 2\,\mathrm{tr}(B\Sigma B\Sigma)$.

Appendix D.1. Model

In defining the the joint distribution of all $I \times T$ elements of the 4D data $\{Y_{it}\}$, we will always assume is that Y_{it} is mean zero and has constant variance over time, $\mathbb{V}(Y_{it}) = \mathbb{V}(Y_{it'})$ for $t \neq t'$, but allow variance to vary over space. For data organized as time series, length-T vectors Y_i, let

$$\begin{aligned} \mathbb{V}(Y_i) =& (\Sigma^S)_{ii}\Sigma^T_{ii}, \\ \mathbb{C}(Y_i, Y_{i'}) =& (\Sigma^S)_{ii'}\Sigma^T_{ii'}, \end{aligned} \tag{D.1}$$

where Σ^S is the $I \times I$ spatial covariance matrix, common to all time points, and $(\Sigma^S)_{ii}$ is the variance at the ith voxel, Σ^T_{ii} is the $T \times T$ temporal autocorrelation matrix for voxel i, $\mathbb{C}(\cdot)$ denotes covariance, and $\Sigma^T_{ii'}$ is the $T \times T$ temporal cross correlation matrix for voxels i and i'. This implies that, for data organized as images, length-I vectors Y_t,

$$\mathbb{V}(Y_t) = \Sigma^S. \tag{D.2}$$

When a time-space separable covariance structure is assumed then $\Sigma^T_{ii'} = \Sigma^T$ for all i, i'.

Appendix D.2. A-var Expected SS

Total SS $\sum_{it} Y_{it}^2$ has expected value

$$\begin{aligned} \mathbb{E}\left(\sum_{i=1}^{I} Y_i^\top Y_i\right) =& \sum_i (\Sigma^S)_{ii}\,\mathrm{tr}(\Sigma^T_{ii}) \\ =& \,\mathrm{tr}(\Sigma^S)T. \end{aligned} \tag{D.3}$$

Appendix D.3. D-var and E-var Expected SS.

The total D-var SS is $\sum_{i=1}^{I}\sum_{t=1}^{T-1}(Y_{i,t+1} - Y_{it})^2/4 = \sum_{i=1}^{I}(DY_i)^\top DY_i/4$ where

$$D = \begin{bmatrix} -1 & 1 & & & \\ & -1 & 1 & & \\ & & \ddots & \ddots & \\ & & & -1 & 1 \end{bmatrix} \tag{D.4}$$

is the $(T-1) \times T$ finite difference matrix. We have

$$\begin{aligned} \mathbb{E}(Y_i^\top D^\top DY_i) =& \,\mathrm{tr}(D^\top D(\Sigma^S)_{ii}\Sigma^T_{ii}) \\ =& 2(T-1) - (\Sigma^T_{ii})_{t,t+1} - 2\sum_{t=2}^{T-1}(\Sigma^T_{ii})_{t,t+1} - (\Sigma^T_{ii})_{T,T-1}, \end{aligned} \tag{D.5}$$

where notably the last expression only depends on the first off-diagonal of the temporal autocorrelation. To obtain more interpretable results we further assume that there is a constant lag-1 autocorrelation at each

voxel, $\rho_i = \left(\Sigma^T_{ii}\right)_{t,t+1}$, for $t = 1, \ldots, T-1$, which reduces (D.5) to $2(T-1)(1-\rho_i)$. This gives the expected total D-var SS as

$$\mathbb{E}\left(\sum_i Y_i^\top D^\top D Y_i/4\right) = \sum_i (\Sigma^S)_{ii}(T-1)(1-\rho_i)/2. \tag{D.6}$$

If we yet further assume constant temporal autocorrelation ρ, corresponding to our separable model, this SS simplifies to $\mathrm{tr}(\Sigma^S)(T-1)(1-\rho)/2$.

The expected SS for S-var is follows the same arguments with differencing matrix replaced with a running sum matrix $\mathrm{abs}(D)$, negating the three negative terms in Eqn. D.5, and reducing to $\mathrm{tr}(\Sigma^S)(T-1)(1+\rho)/2$ under spatially and temporally homogeneous lag-1 temporal autocorrelation.

Appendix D.4. E-var Expected SS.

The total SS E-var is $\sum_{i=1}^I \sum_{t=1,T} Y_{it}^2/2 = \sum_{t=1,T} Y_t' Y_t/2$, with expected value

$$\mathbb{E}\left(\sum_{t=1,T} Y_t' Y_t/2\right) = \mathrm{tr}(\Sigma^S). \tag{D.7}$$

Appendix D.5. A_G-var Expected SS.

The global time series is $\bar{Y}_t$ and total SS due to global is

$$\begin{aligned}\sum_{i=1}^I \sum_{t=1}^T \bar{Y}_t^2 =& I \sum_t (\mathbf{1}^\top Y_t/I)^2 \\ =& \sum_t (\mathbf{1}^\top Y_t)^2/I,\end{aligned} \tag{D.8}$$

where $\mathbf{1}$ is a vector of ones. The expectation of the squared term is $\mathbb{V}(\mathbf{1}^\top Y_t) = \mathbf{1}^\top \Sigma^S \mathbf{1}$, and thus the expected SS is

$$\frac{T}{I}\, \mathbf{1}^\top \Sigma^S \mathbf{1}\,. \tag{D.9}$$

Appendix D.6. D_G-var and S_G-var Expected SS.

Write the global differenced time series as $\bar{Y}_t^D = \mathbf{1}\, Y_t^D/I$ where $Y_t^D = (Y_{t+1} - Y_t)$ for $t = 1, \ldots, T-1$. The total SS due to half differenced global $D_{\mathrm{G}t}$ is then

$$\sum_{i=1}^I \sum_{t=1}^{T-1} (\bar{Y}_t^D)^2/4 = \sum_{t=1}^{T-1} (\mathbf{1}^\top Y_t^D)^2/(4I). \tag{D.10}$$

To find the expectation of the squared term, note that

$$\mathbb{V}(Y_t^D) = 2(\Sigma^S - \Sigma^S \circ \Sigma^{ST}_{t,t+1}), \tag{D.11}$$

where $\circ$ is the Hadamard product and $\Sigma^{ST}_{t,t+1}$ is the spatiotemporal covariance matrix, elements extracted from the temporal cross correlation matrix as per $(\Sigma^{ST}_{tt'})_{ii'} = (\Sigma^{T}_{ii'})_{t,t'}$, and that

$$\mathbb{V}(\mathbf{1}^\top Y_t^D) = 2\left(\mathbf{1}'\,\Sigma^S\,\mathbf{1} + \sum_{ii'} \Sigma^S_{ii'}(\Sigma^T_{ii'})_{t,t+1}\right). \tag{D.12}$$

The final expression for the expected SS is then, with successive assumptions

$$\begin{aligned}\sum_{t=1}^{T-1} \mathbb{V}(\mathbf{1}^\top Y_t^D)/(4I) &= \sum_{t=1}^{T-1} \mathbf{1}'\,\Sigma^S\,\mathbf{1}(1-\Sigma^T_{t,t+1})/(2I) \\ &= (T-1)\,\mathbf{1}'\,\Sigma^S\,\mathbf{1}(1-\rho)/(2I),\end{aligned} \tag{D.13}$$

where first equality comes from assuming a separable covariance structure and the second from a common lag-1 autocorrelation.

The result for S_G-var follows similarly.

Appendix D.7. E_G-var Expected SS.

The total SS E_G-var is $\sum_{i=1}^{I}\sum_{t=1,T} \bar{Y}_t^2/2$, and following same arguments as for A_G-var has expected value

$$\frac{1}{I}\,\mathbf{1}^\top\,\Sigma^S\,\mathbf{1}\,. \tag{D.14}$$

Results for the non-global terms in the decomposition $A_N = D_N + S_N + E_N$ follow as difference of respective total and global terms.

Appendix E. Derivation of DVARS Null Distribution

As results are more naturally defined for squared quantities, we seek a null distribution for

$$\mathrm{DVARS}_t^2 = {Y_t^D}^\top Y_t^D/I, \tag{E.1}$$

where $Y_t^D = Y_{t+1} - Y_t$ as above. While an expression of the mean of DVARS can be obtained from Eqn. (D.11), note also

$$\mathbb{E}(\mathrm{DVARS}_t^2) = \mathrm{tr}(\mathbb{V}(Y_t^D))/I. \tag{E.2}$$

That is, the expected value of DVARS_t^2 is simply the variance of each voxel in the differenced data, averaged over voxels. The natural estimator of this is the sample mean (or robust equivalent) of the sample variance image (or robust equivalent) of the differenced 4D data.

The variance is more involved

$$\mathbb{V}\left(\mathrm{DVARS}_t^2\right) = 2\,\mathrm{tr}\left(\mathbb{V}(Y_t^D)\,\mathbb{V}(Y_t^D)\right)/I^2, \tag{E.3}$$

in particular depending on the entirety of the $I \times I$ difference image variance matrix. For the most restrictive assumptions considered above $\mathbb{V}(Y_t^D) = 2(1-\rho)\Sigma^S$ and thus

$$\mathbb{V}\left(\mathrm{DVARS}_t^2\right) = 8(1-\rho)^2 \frac{\mathrm{tr}(\Sigma^S \Sigma^S)}{I^2}. \tag{E.4}$$

This dependence on the full spatial covariance demands the empirical approaches to variance estimations taken in the body of the paper.

Only at this point do we invoke a normality assumption, and make use of the classic chi-square approximation for sums-of-squared normal variates (Satterthwaite, 1946). In this approach we equate the mean and variance of $c \times \mathrm{DVARS}_t^2$ ($c\mu_0$ & $c^2\sigma_0^2$) and χ_ν^2 (ν & 2ν) and solve for c and ν, giving the multiplier $c = 2\mu_0/\sigma_0^2$ and degrees-of-freedom $\nu = 2\mu_0^2/\sigma_0^2$ as found in Section 2.4.

Appendix F. Power Transformations to Improve DVARS Variance Estimation

The robust IQR-based variance estimate reflects a normality assumption, equating the sample IQR with that of a standard normal. DVARS_t^2, as a sum-of-squares and as reflected by its χ^2 approximation, may exhibit positive skew. Hence we consider power transformations of DVARS_t^2 that may improve symmetry and the accuracy of the IQR variance estimate. While the asymptotically optimal power transformation to normality for χ^2 is known to be the $d = 1/3$ cube-root transformation (Hernandez and Johnson, 1980), our test statistic is only approximately χ^2 and, in particular, variance heterogeneity can worsen the approximation.

To obtain a quantity that should be more symmetric consider the power transformation

$$W_t = \left(\mathrm{DVARS}_t^2\right)^d. \tag{F.1}$$

IQR-based estimates of the variance of W, σ_W^2, will hopefully be more accurate than such estimates on DVARS^2. However, ultimately we seek estimates of the variance of DVARS^2, and so for a given d we compute

$$\begin{aligned} \mathbb{V}(\mathrm{DVARS}_t^2) &= \mathbb{V}(W_t^{1/d}) \\ &= \frac{1}{d}\mu_W^{2(1/d-1)}\sigma_W^2, \end{aligned} \tag{F.2}$$

where the last expression is the delta method variance of $W_t^{1/d}$, and μ_W is the mean of W_t (which we robustly estimate with the median of W_t).

References

Allan, D.W., 1966. Statistics of Atomic Frequency Standards. Proceedings of the IEEE 54, 221–230. doi:10.1109/PROC.1966.4634.

Berntson, G.G., Lozano, D.L., Chen, Y.J., 2005. Filter properties of root mean square successive difference (RMSSD) for heart rate. Psychophysiology 42, 246–252. doi:10.1111/j.1469-8986.2005.00277.x.

Burgess, G.C., Kandala, S., Nolan, D., Laumann, T.O., Power, J.D., Adeyemo, B., Harms, M.P., Petersen, S.E., Barch, D.M., 2016. Evaluation of denoising strategies to address motion-correlated artifacts in resting-state functional magnetic resonance imaging data from the human connectome project. Brain Connectivity 6, 669–680.

Cochrane, D., Orcutt, G.H., 1949. Application of Least Squares Regression to Relationships Containing Auto- Correlated Error Terms. Journal of the American Statistical Association 44, 32–61. URL: http://www.jstor.org/stable/2280349, doi:10.2307/2280349.

Cole, D.M., Smith, S.M., Beckmann, C.F., 2010. Advances and pitfalls in the analysis and interpretation of resting-state FMRI data. Frontiers in systems neuroscience 4, 8. URL: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2854531{&}tool=pmcentrez{&}rendertype=abstracthttp://www.ncbi.nlm.nih.gov/pubmed/20407579, doi:10.3389/fnsys.2010.00008.

Craddock, R., Benhajali, Y., Chu, C., Chouinard, F., Evans, A., Jakab, A., Khundrakpam, B., Lewis, J., Li, Q., Milham, M., et al., 2013. The neuro bureau preprocessing initiative: open sharing of preprocessed neuroimaging data and derivatives. Frontiers in Neuroinformatics (Neuroinformatics 2013) .

Di Martino, A., Yan, C.G., Li, Q., Denio, E., Castellanos, F.X., Alaerts, K., Anderson, J.S., Assaf, M., Bookheimer, S.Y., Dapretto, M., et al., 2014. The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. Molecular psychiatry 19, 659–667.

Glasser, M.F., Sotiropoulos, S.N., Wilson, J.A., Coalson, T.S., Fischl, B., Andersson, J.L., Xu, J., Jbabdi, S., Webster, M., Polimeni, J.R., Van Essen, D.C., Jenkinson, M., 2013. The minimal preprocessing pipelines for the Human Connectome Project. NeuroImage 80, 105–124. URL: http://dx.doi.org/10.1016/j.neuroimage.2013.04.127, doi:10.1016/j.neuroimage.2013.04.127, arXiv:NIHMS150003.

Gorgolewski, K., Burns, C.D., Madison, C., Clark, D., Halchenko, Y.O., Waskom, M.L., Ghosh, S.S., 2011. Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in python. Frontiers in neuroinformatics 5, 13.

Harper, W.M., 1967. The Distribution of the Mean Half-Square Successive Difference. Biometrika 54, 419–433. URL: http://www.jstor.org/stable/2335034.

Hernandez, F., Johnson, R., 1980. The large-sample behavior of transformations to normality. Journal of

American Statistical Association 75, 855–861. URL: http://www.tandfonline.com/doi/abs/10.1080/01621459.1980.10477563, doi:10.1080/01621459.1980.10477563.

Kotz, S., Johnson, N., Read, C., 1988. Successive Differences in. Number v. 2 in Encyclopedia of Statistical Sciences, Wiley. URL: http://onlinelibrary.wiley.com/book/10.1002/0471667196.

von Neumann, J., Kent, R.H., Bellinson, H.R., Hart, B.I., 1941. The mean square successive difference. The Annals of Mathematical Statistics 12, 153–162. URL: http://www.jstor.org/stable/2235765.

Nichols, T., 2013. Notes on Creating a Standardized Version of DVARS , 1–5arXiv:1704.01469.

Power, J.D., Barnes, K.A., Snyder, A.Z., Schlaggar, B.L., Petersen, S.E., 2012. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. NeuroImage 59, 2142–2154. URL: http://dx.doi.org/10.1016/j.neuroimage.2011.10.018, doi:10.1016/j.neuroimage.2011.10.018.

Salimi-Khorshidi, G., Douaud, G., Beckmann, C.F., Glasser, M.F., Griffanti, L., Smith, S.M., 2014. Automatic denoising of functional MRI data: combining independent component analysis and hierarchical fusion of classifiers. NeuroImage 90, 449–68. URL: http://www.ncbi.nlm.nih.gov/pubmed/24389422, doi:10.1016/j.neuroimage.2013.11.046.

Satterthwaite, F.E., 1946. An Approximate Distribution of Estimates of Variance Components. Biometrics Bulletin 2, 110–114. URL: http://www.jstor.org/stable/3002019.

Smith, S.M., Vidaurre, D., Beckmann, C.F., Glasser, M.F., Jenkinson, M., Miller, K.L., Nichols, T.E., Robinson, E.C., Salimi-Khorshidi, G., Woolrich, M.W., Barch, D.M., Uurbil, K., Van Essen, D.C., 2013. Functional connectomics from resting-state fMRI. Trends in cognitive sciences 17, 666–82. URL: http://www.ncbi.nlm.nih.gov/pubmed/24238796, doi:10.1016/j.tics.2013.09.016.

Smyser, C.D., Snyder, A.Z., Neil, J.J., 2011. Functional connectivity MRI in infants: Exploration of the functional organization of the developing brain. NeuroImage 56, 1437–1452. URL: http://dx.doi.org/10.1016/j.neuroimage.2011.02.073, doi:10.1016/j.neuroimage.2011.02.073, arXiv:NIHMS150003.

Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E., Yacoub, E., Ugurbil, K., Consortium, W.M.H., et al., 2013. The wu-minn human connectome project: an overview. Neuroimage 80, 62–79.

Table 1: Time series visualisation of the DSE variance decomposition. A-var is to the total variance at time point t, D-var, S-var and E-var correspond to the fast, slow and edge variance terms. Global and non-global variance components sum to the total components. All of these terms, given as mean squared quantities, are best reported and plotted in root mean squared (RMS) units (see Appendix B for more on plotting global variance).

Name	Notation	Value	Range	x-axis loc.
A-var	A_t	$\frac{1}{I}\sum_{i=1}^{I} Y_{it}^2$	$t = 1, \ldots, T$	t
D-var	D_t	$\frac{1}{I}\sum_{i=1}^{I}(Y_{it} - Y_{i,t+1})^2/4$	$t = 1, \ldots, T-1$	$t + \frac{1}{2}$
S-var	S_t	$\frac{1}{I}\sum_{i=1}^{I}(Y_{it} + Y_{i,t+1})^2/4$	$t = 1, \ldots, T-1$	$t + \frac{1}{2}$
E-var	E_t	$\frac{1}{I}\sum_{i=1}^{I} Y_{it}^2/2$	$t = 1, T$	t
Global A-var	A_{Gt}	$\bar{Y}_t^2$	$t = 1, \ldots, T$	t
Global D-var	D_{Gt}	$(\bar{Y}_t - \bar{Y}_{t+1})^2/4$	$t = 1, \ldots, T-1$	$t + \frac{1}{2}$
Global S-var	S_{Gt}	$(\bar{Y}_t + \bar{Y}_{t+1})^2/4$	$t = 1, \ldots, T-1$	$t + \frac{1}{2}$
Global E-var	E_{Gt}	$\bar{Y}_t^2/2$	$t = 1, T$	t
Non-Global A-var	A_{Nt}	$\frac{1}{I}\sum_i(Y_{it} - \bar{Y}_t)^2$	$t = 1, \ldots, T$	t
Non-Global D-var	D_{Nt}	$\frac{1}{I}\sum_i(Y_{it} + Y_{i,t+1} - (\bar{Y}_t - \bar{Y}_{t+1}))^2/4$	$t = 1, \ldots, T-1$	$t + \frac{1}{2}$
Non-Global S-var	S_{Nt}	$\frac{1}{I}\sum_i(Y_{it} + Y_{i,t+1} - (\bar{Y}_t + \bar{Y}_{t+1}))^2/4$	$t = 1, \ldots, T-1$	$t + \frac{1}{2}$
Non-Global E-var	E_{Nt}	$\frac{1}{I}\sum_i(Y_{it} - \bar{Y}_t)^2/2$	$t = 1, T$	t

Table 2: DSE ANOVA table giving a mean squared (MS) variance decompositions of resting-state fMRI data. The first row shows how the total MS can be split into 3 terms, in the second through 4th columns, $A = D + S + E$. The first column likewise shows how total MS can be decomposed in to that explained by a spatially global time series (second row) and a non-global or residual-global component (third row), $A = A_G + A_N$. Likewise, each row and column sums accordingly: $A_G = D_G + S_G + E_G$, $D = D_G + D_N$, etc. Terms are shown here as MS for brevity, but are best reported in root mean squared (RMS) units. Unless noted otherwise, summations are over the full range of possible values. See Table 1 for definitions of the time series variables.

	A-var	D-var	S-var	E-var
Whole	$A = \frac{1}{T}\sum_t A_t$	$D = \frac{1}{T}\sum_t D_t$	$S = \frac{1}{T}\sum_t S_t$	$E = \frac{1}{T}\sum_{t=1,T} E_t$
Global	$A_G = \frac{1}{T}\sum_t A_{Gt}$	$D_G = \frac{1}{T}\sum_t D_{Gt}$	$S_G = \frac{1}{T}\sum_t S_{Gt}$	$E_G = \frac{1}{T}\sum_{t=1,T} E_{Gt}$
Non-Global	$A_N = \frac{1}{T}\sum_t A_{Nt}$	$D_N = \frac{1}{T}\sum_t D_{Nt}$	$S_N = \frac{1}{T}\sum_t S_{Nt}$	$E_N = \frac{1}{T}\sum_{t=1,T} E_{Nt}$

Table 3: Expected values of the DSE ANOVA table under nominal models. First two rows show expected mean squared (MS) values under the simplified separable noise model, for whole and global variance; Σ^S is the $I \times I$ spatial covariance matrix and $\mathrm{tr}(\Sigma^S)$ is its trace, the sum of voxel-wise variances. Third and fourth rows show expected MS normalized relative to the total variance A-var. Final two rows show the expected normalized MS under a naive, default model of independent and identically distributed (IID) data in time and space. This shows that D-var and S-var are equal under independence and, when normalized, approximately 1/2. Due to the $1/I$ term, the global will generally only ever explain a small portion of variance, but, spatially structured noise reflected by the average summed spatial covariance, $\mathbf{1}^\top \Sigma^S \mathbf{1} / I$, can enlarge global variance.

	A-var	D-var	S-var	E-var
Whole Sep.	$\frac{1}{I}\,\mathrm{tr}(\Sigma^S)$	$\frac{1}{2}\frac{1}{I}\frac{T-1}{T}(1-\rho)\,\mathrm{tr}(\Sigma^S)$	$\frac{1}{2}\frac{1}{I}\frac{T-1}{T}(1+\rho)\,\mathrm{tr}(\Sigma^S)$	$\frac{1}{I}\frac{1}{T}\,\mathrm{tr}(\Sigma^S)$
Global Sep.	$\frac{1}{I^2}\,\mathbf{1}^\top \Sigma^S\,\mathbf{1}$	$\frac{1}{2}\frac{1}{I^2}\frac{T-1}{T}(1-\rho)\,\mathbf{1}^\top \Sigma^S\,\mathbf{1}$	$\frac{1}{2}\frac{1}{I^2}\frac{T-1}{T}(1+\rho)\,\mathbf{1}^\top \Sigma^S\,\mathbf{1}$	$\frac{1}{I^2}\frac{1}{T}\,\mathbf{1}^\top \Sigma^S\,\mathbf{1}$
Whole Sep. Rel.	1	$\frac{1}{2}\frac{T-1}{T}(1-\rho)$	$\frac{1}{2}\frac{T-1}{T}(1+\rho)$	$\frac{1}{T}$
Global Sep. Rel.	$\frac{1}{I}\,\mathbf{1}^\top \Sigma^S\,\mathbf{1} / \mathrm{tr}(\Sigma^S)$	$\frac{1}{2}\frac{1}{I}\frac{T-1}{T}(1-\rho)\,\mathbf{1}^\top \Sigma^S\,\mathbf{1} / \mathrm{tr}(\Sigma^S)$	$\frac{1}{2}\frac{1}{I}\frac{T-1}{T}(1+\rho)\,\mathbf{1}^\top \Sigma^S\,\mathbf{1} / \mathrm{tr}(\Sigma^S)$	$\frac{1}{I}\frac{1}{T}\,\mathbf{1}^\top \Sigma^S\,\mathbf{1} / \mathrm{tr}(\Sigma^S)$
Whole IID Rel.	1	$\frac{1}{2}\frac{T-1}{T}$	$\frac{1}{2}\frac{T-1}{T}$	$\frac{1}{T}$
Global IID Rel.	$\frac{1}{I}$	$\frac{1}{2}\frac{1}{I}\frac{T-1}{T}$	$\frac{1}{2}\frac{1}{I}\frac{T-1}{T}$	$\frac{1}{I}\frac{1}{T}$

Table 4: Form and interpretation of various DVARS variants, expressed as functions of original DVARS_t. Here $\{Y_{it}\}$ are the 4D data, A is the overall mean square variance, μ_0 is the expected DVARS_t^2 under a null model, and $P(\mathrm{DVARS}_t)$ is the p-value for DVARS^2, and Φ^{-1} is the inverse cumulative distribution function of a normal.

Name	Expression	Interpretation
DVARS	$\mathrm{DVARS}_t = \sqrt{\sum_i (Y_{it} - Y_{i,t+1})^2 / I}$	Standard deviation of difference image
$\sqrt{}$ D-var	$\mathrm{DVARS}_t / 2$	Fast component of noise, as standard deviation
$\%D$-var	$\mathrm{DVARS}_t^2 / (4A) \times 100$	Fast noise, as % of average noise variance
$\Delta\%D$-var	$(\mathrm{DVARS}_t^2 - \mu_0)/(4A) \times 100$	Excess fast noise, as % of average noise variance
Rel. DVARS	$\mathrm{DVARS}_t / \sqrt{\mu_0}$	DVARS as a multiple of null mean
Z(D-var)	$\Phi^{-1}(1 - P(\mathrm{DVARS}_t))$	DVARS p-value as Z-score