Filtering induces correlation in fMRI resting state data

Catherine Davey[1,2], David B. Grayden[2], Gary F. Egan[1,3], Leigh A. Johnston[1,2]

[1]Department of Electrical and Electronic Engineering, University of Melbourne & NICTA Victorian Research Laboratory
[2]Howard Florey Institute, Florey Neuroscience Institutes. [3]Centre for Neuroscience, University of Melbourne, Australia

Introduction: It is standard practice in the analysis of fMRI resting state data to filter the data prior to calculating measures of dependence between voxels. Temporal band-pass filtering in the range of 0.01-0.08Hz is suggested for removal of low frequency drift and high frequency physiological noise [1]. Testing for significant dependence between voxel timeseries typically proceeds by one of three methods: 1) **Seed-voxel correlation tested via Fisher's z-transform**. e.g. examination of low-frequency correlations <0.08Hz [2], 2) **Seed-voxel correlation tested via t-test**. e.g. identification of significant correlation in band filtered data (0.01-0.08Hz) [3], or 3) **Instantaneous Granger Causality**. e.g. bivariate auto-regressive model of bandpass filtered timeseries [4]. We demonstrate that temporal filtering can induce artificial dependence for each of these measures. Furthermore, we derive corrections to each statistical test that account for the induced effect.

Theory: Let x and y denote two time series with a bivariate normal distribution, of length T, ρ_{xy} denote sample correlation and r_{xy} true correlation. Assume that a filter with normalised amplitude response a_k, for frequency k, is applied to each time series. Let $\rho^{(f)}_{xy}$ denote the sample correlation of the filtered data. Table 1 displays the dependence measures and tests, including our corrections for filtered data.

	Unfiltered	Filtered & corrected
Variance of sample correlation	$\mathrm{var}(\rho_{xy}) = \frac{(1-r_{xy}^2)^2}{T}$	$\mathrm{var}(\rho_{xy}^{(f)}) = (1-\rho_{xy}^{(f)2})^2 \left(\sum_{k=0}^{T} a_k\right)^{-1}$
Fisher's z-transform	$\mathrm{var}(z_{xy}) = T^{-1}$	$\mathrm{var}(z^{(f)}) = \left(\sum_{k=0}^{T} a_k\right)^{-1}$
T-test under null hypothesis	$t_{xy} = \rho_{xy}\sqrt{\frac{T-2}{1-\rho_{xy}^2}} \sim \mathrm{T}_{T-2}$	$t_{xy}^{(f)} = \rho_{xy}\sqrt{\frac{\sum_{k=0}^{T} a_k}{1-\hat{\rho}_{xy}^2}}$
Instantaneous Granger Causality, F_{xy}	$(T-3)F_{xy} \sim \chi^2(1)$	$\sqrt{\sum_{k=0}^{T} a_k}\, F_{xy} \sim \chi^2(1)$

Table 1: Dependence measures and their application to unfiltered and filtered data.

Simulated data: Sets containing 2000 pairs of time series from a bivariate normal distribution, characterised by time series length, T, variance, σ, and correlation, ρ, sampling frequency = 1s. Filter parameters: low-pass cut-off, f_l in [0,0.5) Hz , high-pass cut-off, f_h in (0,0.5] Hz. For IGC, an AR(1) model was used.

Experimental data: Resting state BOLD EPI scans of two healthy controls were acquired (3T Siemens Trio): TR=1.6s, FA=90, voxel size = 3.125x3.125x5.5 mm^3, 219 volumes. Images were motion corrected and smoothed (6x6x6mm^3). Left primary motor cortex (LMC) was identified for each subject via a GLM activation analysis on accompanying motor-task data, and timeseries averaged in surrounding ROI. Pre-filtering was implemented with a low-pass FIR filter, $f_h = 0.1Hz$.

Results: The simulation results demonstrate that the sample variance of correlation between filtered signals is predicted by the sum of the filter amplitude response and the true correlation (Fig. 1).

Fig. 2 shows the distribution of linear dependence measures applied to unfiltered, and filtered, noise. Corrected linear dependence measures have the same distribution as measures resulting from unfiltered data, in conformance with the null case from which the measures were drawn. Filtered but uncorrected measures show larger variance, increasing the false positive rate and inducing correlation.

Fig. 3 shows voxel-wise dependence measures relative to the LMC, for filtered and unfiltered data, for two different subjects. The results generated from corrected measures better reflect the functional connections expected in the motor cortex, and exhibit increased consistency across the dependence measures. The proportion of significant voxels resulting from filtered data with uncorrected measures suggest that artificial dependence is induced, while the lack of consistency across unfiltered maps suggest that noise is corrupting the results.

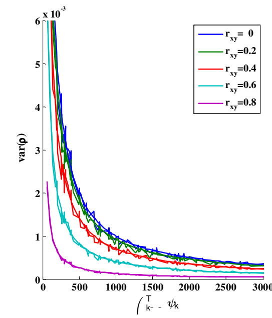

Figure 1: Variance of sample correlation via simulation (full), and predicted by filter amplitude response (dashed).

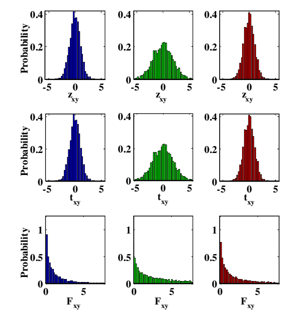

Figure 2: Dependence measures applied to unfiltered and filtered simulation data, with zero true correlation.

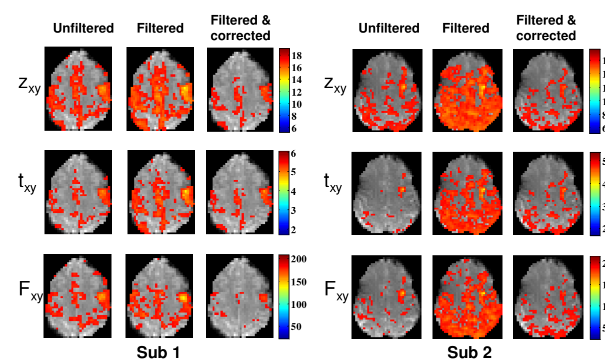

Figure 3: Dependence measures applied to unfiltered and filtered data.

Conclusions: Filtering timeseries can induce artificial dependence. We show that the sample variance of linear measures depends on the amplitude response of the filter; this dependence can be exploited to correct the dependence measures to counteract the induced dependence. Our results extend to multiple filters, and all filter types.

[1] Biswal *et al.* (1995) MRM 34:537-541 [2] Hampson *et al.* (2002) HBM 15:247-262. [3] Zhong *et al.* (2009) Brain Top. 22:134-144. [4] Abler *et al.* (2005) MRI 24:181-185.