A Multiscale Analysis of the Temporal Characteristics of Resting-State fMRI Data

Cheolwoo Park, Nicole A. Lazar, Jeongyoun Ahn
Department of Statistics
University of Georgia
Athens, GA 30602

and Andrew Sornborger
Faculty of Engineering and Department of Mathematics
University of Georgia
Athens, GA 30602

August 20, 2010

Abstract

In this paper, we conduct an investigation of the null hypothesis distribution for functional magnetic resonance imaging (fMRI) time series using multiscale analysis tools, SiZer (Significance of Zero crossings of the derivative) and wavelets. Most current approaches to the analysis of fMRI data assume simple models for temporal (short term or long term) dependence structure. Such simplifications are to some extent necessary due to the complex, high-dimensional nature of the data, but to date there have been few systematic studies of the dependence structures under a range of possible null hypotheses, using data sets gathered specifically for that purpose. We aim to address some of these issues by analyzing the detrended data with a long enough time horizon to study possible long-range temporal dependence. Our multiscale approach shows that even for resting-state data, *i.e.* "null" or ambient thought, some voxel time series cannot be modeled by white noise and need long-range dependent type error structure. This finding suggests the use of different time series models in different parts of the brain in fMRI studies.

Key words: Hurst parameter; Long–range dependence; Principal component analysis; SiZer; Wavelet spectrum.

1 Introduction

In the analysis of functional magnetic resonance imaging (fMRI) data it is critical to precisely define and understand the nature of the null hypothesis. To clarify the issue, consider a number of realistic experimental scenarios. In the first, a scientist is interested in discovering which areas of the brain are activated in response to a visual stimulus. The way this is often done is to alternate periods of stimulus presentation with control periods of fixation or rest. Simple statistical models, such as a t-test, are then carried out to identify the voxels in the brain at which there are statistically significant differences between the levels of activation during the periods of task and the periods of rest. In a second scenario, a researcher compares activation of a schizophrenic brain to that of a "normal" brain, again via a simple statistical model. A third scenario might involve determining, via a series of brain scans, when a stroke patient has recovered full function. In all of these cases, certain assumptions are made about the null state, which in the first situation can be classified as "ambient thought", but in the other two situations is not as straightforward: What constitutes "normal" against which to compare a pathological brain? In the absence of baseline measurements on a stroke patient (*i.e.* measurements taken prior to the trauma), how can recovery to full function be determined? Lacking precise knowledge about the distribution of the brain activity at rest, or more generally under a given null hypothesis, statements about power and other properties of statistical tests can only be speculative and will tend to be misleading.

Typical assumptions regarding the null distribution of temporal structure are simple, such as autoregressive of low order, or multiscale potentially long-range dependence, to reflect general linear models, temporal models and wavelet-based models respectively. Luo and Nichols (2003) develop a tool for visualizing and diagnosing fMRI data to check the assumptions of fMRI time series. They introduce methods to find and characterize violations of the assumptions of linear models using either a working correlation model of independence or an arbitrary covariance structure. While their tools are useful, they only consider weakly correlated error such as AR(1). Lund et al. (2006) improve their approach using a nuisance variable regression model and find that the AR(1) model is inadequate near larger arteries due to oscillatory noise. An AR(p) model has been considered by many authors for explaining temporal correlation; see Bullmore et al. (1996); Locascio et al. (1997); Woolrich et al. (2001); Penny et al. (2003); Gautama and Van Hulle (2005).

Several research groups have already pioneered the possibility of long-range dependence in fMRI data and the use of wavelets; see Bullmore et al. (2003) for a general review. Wavelet analysis is of interest in the analysis of fMRI data because of its multiresolutional nature, its adaptivity to non-stationary or local features in the data, its ability to decorrelate autocorrelated data, its usefulness

in denoising and compressing large imaging data sets (not just fMRI), and its fast computational speed. In a recent paper Maxim et al. (2005) study the differences between patients with early Alzheimer's disease and age-matched comparison subjects by Hurst parameter estimation. However, since most data that are collected experimentally involve only a relatively small number of time points, on the order of one to three hundred at best (representing a total scanning time of well under half an hour in most instances), it has not been possible to fully investigate the dependence structure of the null data, regardless of how that has been defined. For instance, with only a hundred or so time points, wavelet spectra or long-range temporal dependence cannot be adequately evaluated, and so the nature of the dependence remains at best a supposition. Even when such an evaluation is attempted, as in work by Bullmore et al. (2003), any long-term effects present during times of "background thought" – not to mention the actively working brain – will be missed for lack of information at the relevant time scale.

A related issue to understanding the temporal noise structure in the data is whether the same structure can be - or should be - used to describe the entire brain. Purdon et al. (2001) note that noise structure varies over the brain, and thus it makes sense to have different models for different parts at the brain although the authors did not go farther than suggesting the possibility. Also, Long et al. (2005) study accurate characterization of the noise processes using two time-varying procedures for examining non-stationary noise structure.

In this paper we aim to address some of the outstanding issues made apparent by these previous works. Specifically, our focus is on the understanding of the null hypothesis in fMRI time series, for which we consider three distinct elements. First, in order to characterize the temporal structure itself in a thorough way, we collect data only for this purpose. Our working data set comprises 1500 time points collected on a subject who is not performing any task while in the MR scanner, a "null" data set of close to an hour in length. We focus on resting-state data to avoid any possible signal-related sources which could contaminate our assessment procedure on the temporal structure itself. Second, we consider a wide range of dependence structures from white noise to long-range dependence. Our data set is long enough to allow for an in-depth study of the dependence structure in resting-state data, the first step in a thorough evaluation of the distributional characteristics of the null hypothesis. Third, we propose to use the multiscale tools SiZer (Significance of Zero crossings of the derivative) and wavelet spectrum to study the temporal structure of data with a sufficient number of time points. Our multiscale approach combines exploratory analyses with statistical inference and allows us to study a wider variety of noise structures.

It is also becoming popular to use resting-state data to explore the "default mode" of brain function, or "resting-state connectivity" (see, for instance, Raichle and Snyder (2007); Fox and

Raichle (2007) for some recent reviews). Here, data are collected while the patient is at rest to study the background network of the brain when no specific cognitive task is being carried out. Understanding the temporal correlation structure of the resting brain is relevant to the study of resting-state connectivity, therefore, aside from being of interest in and of itself.

The paper is organized as follows. Section 2 introduces our main statistical tools, SiZer and the wavelet spectrum. Data description, pre-processing, and detrending of the data are presented in Section 3. A multiscale analysis based on our statistical tools is provided in Section 4. We discuss our conclusions and future work in Section 5.

2 Multiscale Analysis Methods

In this section, we describe two multiscale statistical methods, SiZer that is based on a scale space approach and the wavelet spectrum. Both have proved to be very useful in the analysis of complex data, for example Internet traffic data (Park et al., 2010). In the fMRI context, Worsley et al. (1996) and Worsley and Evans (1997) apply a scale space approach to determine the amount of smoothing before the analysis is performed. In many studies the images are smoothed to a full width at half maximum (FWHM) of a certain scale (e.g. 20mm) to reduce the impact of misregistration of the data, to allow for anatomical variability among subjects, and to improve the signal to noise ratio. Worsley et al. (1996) and Worsley and Evans (1997) propose a computationally efficient search over a range of filter widths to find local maxima in location and scale space. We point out that our scale space approach carries different meanings and goals from theirs because we use SiZer and the wavelet spectrum as goodness-of-fit test tools to find a proper time series model for each voxel.

2.1 SiZer

SiZer analysis (Chaudhuri and Marron, 1999) is a visualization method which enables statistical inference for discovery of meaningful structure within a data set, while doing exploratory analysis via smoothing. SiZer addresses the question of which features observed in a smooth (*e.g.* a local linear fit) represent important underlying structure, and are not simply artifacts of the sampling noise. The method is based on scale-space ideas from computer vision; see Lindeberg (1994). In SiZer, scale-space is a family of kernel smooths indexed by the scale, which is the smoothing parameter or bandwidth h. SiZer considers a wide range of bandwidths, thereby avoiding the classical problem of bandwidth selection. This approach uses all the information that is available in the data at each scale.

SiZer visually displays the significance of features over location x and scale h, using a color map. Each pixel shows the result of a hypothesis test for the slope of the smoothed curve at the point indexed by the horizontal location x, and by the bandwidth corresponding to the row h. Each point (x, h) is colored on a gray scale, according to the behavior of the curve at that point.

As an example, consider regression with a fixed design. Given data $(x_i, Y(i))$ for $i = 1, \ldots, N$, the model is

$$Y(i) = g(x_i) + \epsilon_i, \tag{2.1}$$

where $g(\cdot)$ is a regression function and the ϵ_i's are identically and independently distributed with $E(\epsilon_i) = 0$ and $\mathrm{Var}(\epsilon_i) = \sigma^2$.

SiZer applies local linear fitting (Fan and Gijbels, 1996) to obtain a family of kernel estimates and derivatives of the regression function. The regression function g is locally approximated by a linear equation and it is fitted by minimizing

$$\sum_{i=1}^{N} \{Y(i) - (\beta_0 + \beta_1(x_i - x_0))\}^2 K_h(x_i - x_0) \tag{2.2}$$

over $\beta = (\beta_0, \beta_1)'$, where $K_h(\cdot) = K(\cdot/h)/h$. The kernel function K is the standard normal density function. By using a Taylor expansion, it is easy to show that the solution of the above equation provides estimates of the regression function and its first derivatives at x_0 for different bandwidths, that is $\hat{\beta}_0 \approx g_h(x_0) = \int g(x) K_h(x - x_0) dx$, and $\hat{\beta}_1 \approx g_h'(x_0) = \int g'(x) K_h(x - x_0) dx$.

From this solution, we can construct the family of smooths parameterized by h and the confidence intervals that underlie the SiZer analysis. These are of the form

$$\hat{g}_h'(x_0) \pm q \times \hat{sd}(\hat{g}_h'(x_0)),$$

where q is an appropriate quantile that accounts for multiple testing corrections. We use the quantile proposed by Hannig and Marron (2006), which is based on extremal value theory. Chaudhuri and Marron (1999) discuss the estimation of the standard deviation. The significant features are determined based on whether these confidence intervals are above (below) zero.

Conventional SiZer is a useful tool to find meaningful structure in a data set, but its usefulness is limited in the case of dependent data because it assumes independent errors and compares against a white noise null hypothesis. In cases of dependent data, significant features which are due to the presence of dependence appear in the plot. For time series, $x_i = i$ and $Cov(\epsilon_i, \epsilon_j) = \gamma(|i - j|)$. Let X be the design matrix of the local linear fit at $x_0 = i_0$, and let $W = diag\{K_h(i - i_0)\}$. Then, the variance of the local polynomial estimator is given by

$$V(\hat{\beta}|X) = (X^T W X)^{-1} (X^T \Sigma X)(X^T W X)^{-1},$$

where, for the assumed correlation structure, Σ is the kernel weighted covariance matrix of the errors with generic element

$$\sigma_{ij} = \gamma(|i-j|)K_h(i-i_0)K_h(j-i_0).$$

For SiZer to fulfill its potential to flag significant trends in time series, its underlying confidence intervals must be adjusted to properly account for the correlation structure of the data. Since there might exist stationary stochastic processes whose realizations can look quite similar to significant trends in a SiZer map, the distinction between 'significant features' and 'dependence structure' becomes a nontrivial problem. Dependent SiZer (Park et al., 2004) uses a true autocovariance function γ of an assumed model and conducts a goodness of fit test. This tool enables us to see how different the behavior of the data is from that of the assumed model. The only difference between original SiZer and dependent SiZer is that the latter compares the data with the specified model rather than with white noise.

Implementation of dependent SiZer requires estimation of the parameters involved in the autocovariance function. For example, let us consider an AR(p) process:

$$e_t = \phi_1 e_{t-1} + \phi_2 e_{t-2} + \cdots + \phi_p e_{t-p} + \epsilon_t, \tag{2.3}$$

where $\phi_1, \ldots, \phi_p$ are the autoregression coefficients and ϵ_t are independent and identically distributed random variables with mean 0 and variance σ^2. For a simple case, the autocovariance function γ of an AR(1) process with coefficient ϕ is given by

$$\gamma(l) = \frac{\sigma^2}{1-\phi^2}\phi^l, \quad l = 0, 1, 2, \ldots.$$

A common model which can be used to describe long-range dependence in the errors is *fractional Gaussian noise*, the increment of fractional Brownian motion. Fractional Gaussian noise is a stationary Gaussian process with mean zero and autocovariance function

$$\gamma(l) = \sigma^2(|l+1|^{2H} + |l-1|^{2H} - 2|l|^{2H})/2, \quad l = 0, 1, 2, \ldots,$$

where H, $H \in (0, 1)$, is the Hurst long–range dependence parameter. Thus, one needs to estimate ϕ and σ^2 or H and σ^2 to conduct a goodness of fit test using dependent SiZer. In this paper, σ^2 and ϕ are conventionally estimated by the sample variance and $\hat{\gamma}(1)/\hat{\sigma}^2$, respectively, and the estimation of H is explained in Section 2.2. The details of the multiple comparison test procedure in dependent SiZer can be found in Park et al. (2004) and Rondonotti et al. (2007).

The conventional SiZer Matlab code is available on the web at `http://www.unc.edu/~marron/marron_software.html` and the dependent SiZer Matlab code is available from the first author.

2.2 Wavelet Spectrum

Abry and Veitch (1998) propose a wavelet based estimator of the Hurst parameter H. Let ψ be a wavelet function with moment order L, that is (Daubechies, 1992)

$$\int u^l \psi(u) du = 0, \quad l = 0, 1, \ldots, L-1 \tag{2.4}$$

and $\psi_{j,k}(u) = 2^{-j/2}\psi(2^{-j}u - k),\ k \in Z$. The wavelet coefficients of $Y_t \in L^2(R)$ at time t, $d_{Y_t}(j,k)$ are defined as $d_{Y_t}(j,k) = \langle Y_t, \psi_{j,k} \rangle$ where $\langle \cdot, \cdot \rangle$ represents the inner product of $L^2(R)$. If we perform a time average of the $|d_{Y_t}(j,k)|^2$ at a given scale, that is,

$$\mu_j = \frac{1}{n_j} \sum_{k=1}^{n_j} d^2_{Y_t}(j,k),$$

where n_j is the number of coefficients at scale j, then $E \log_2(\mu_j) \sim (2H-1)j + C$, where C depends only on H. Define the variable S_j for some appropriately chosen range $j = j_1, \ldots, j_2$ as $S_j \equiv \log_2(\mu_j) - a_j$, where a_j is a bias correction term. Then the slope of a weighted linear regression applied to (j, S_j) for $j = j_1, \ldots, j_2$, provides an estimate of the Hurst parameter.

An advantage of applying the wavelet method is that the vanishing moments allow us to ignore polynomial trends in the data up to order $L-1$. Thus Hurst parameter estimation is very robust against simple trends. Another advantage of the wavelet-based method is that for a stationary process, the wavelet coefficients tend to be approximately uncorrelated (see Abry and Veitch (1998) and, for example, Stoev et al. (2005)). Since fMRI data are correlated in time, this decorrelation property is particulary useful. For more detail about the wavelet spectrum, see Abry and Veitch (1998); Stoev et al. (2005); Veitch and Abry (1999).

The Matlab code for the wavelet spectrum is available on the web at `http://www.cubinlab.ee.unimelb.edu.au/~darryl/secondorder_code.html`.

3 Resting-State Data

The fMRI data that we investigate were obtained on a 1.5 Tesla magnet. The aim was to obtain a long record that had enough dynamical range to assess the temporal correlation of fMRI data in a resting subject. Three 64×64 voxel slices with repetition time $T_R = 2$ seconds were imaged. The three axial slices (4mm thickness) were separated by 10mm and the most inferior slice displays the frontal sinus (slightly above the eye). Data were collected for approximately 50 minutes, resulting in 1500 imaged volumes. In order to decrease any noticeable degree of magnetization non-equilibrium at the beginning of the scan, 30 extra images were obtained at the beginning of

the session and discarded. The overall stability of the brain was good (not much head motion) even before registration. A total slow translation of the brain by approximately 1mm and a maximal rotation of approximately 0.2 degrees were corrected for by registration of the data. The subject's gaze was focused on a fixation point and no stimulus was presented for the duration of the experiment. Although this fixation could not remove all non-stationarity in the data entirely, it was meant to limit non-stationarity to the extent possible. The 1500 images of this dataset is an order of magnitude longer than is standard in fMRI data collection, allowing for better estimation of the temporal dependence structure.

Our main interest is measurement errors in time, but fMRI data are made up of signal, which for our purposes means changes in the Blood Oxygen Level-Dependent (BOLD) signal due to coherent neural activity, and several other types of signal not of interest that can include: machine-generated patterned noise, physiological noise from subjects (this includes heartbeat, respiration and vasomotor oscillations that are also part of the BOLD signal). Due to the correlation of the components of the fMRI signal, it can be challenging to distinguish the measurement error specifically from the rest. However, as our goal is to study the behavior of that error, we expend some effort on making the distinction, in order to obtain, to the extent possible, the temporal correlation in the measured data. Our approach consists of pre-processing, followed by linear and PCA detrending. In Section 4, we analyze the pre-processed and detrended data using our multiscale approach. All data analysis is done in Matlab.

A simple pre-processing step is to identify, and if necessary remove, obvious anomalies or outliers in the data. We explore the mean response over all 12288 voxels at each time point and find an outlier in the middle of the series at time point 587. Further investigation shows an anomaly that has the appearance of being machine-generated, as evidenced by a "rippling" effect outside of the brain. The outlier needs to be remedied since the artifact will affect our assessment of the measurement error in time. Therefore, we replace this image with the average of the images at time points 586 and 588. This is a common approach to handling"spike" outliers of the type we observe in the data. These images are not presented here to save space; refer to Park et al. (2008) for more details about our pre-precessing step.

Each voxel time series can be viewed in a nonparametric regression setting as in (2.1). Our goal is to study ϵ and characterize its temporal (in)dependence structure. For this purpose, we need to remove the trend $g(\cdot)$ in (2.1). In fMRI data, linear trends are often observed and we remove them by fitting linear regressions. Let V be the 1500-by-4096 data matrix, where each row is one frame of the brain image and column V_j is the time series of the jth voxel. Note that we have a data matrix for each slice. The linear detrending can be done, for each V_j, by fitting $\widehat{V}_j(i) = a_j + b_j i$ at

time point i. Let $\widehat{V} = [\widehat{V}_1, \widehat{V}_2, \cdots, \widehat{V}_{4096}]$ and let $V_L = V - \widehat{V}$ be the linearly detrended data.

Fig. 1 about here.

Figure 1 (a) shows the three slices at time point $t = 900$, which is arbitrarily chosen for illustration purposes. The brain is clearly visible. Figure 1 (b) shows the linearly detrended images at the same time point. This simple linear detrending procedure removes most of the signal, but there still remain spatially coherent features, which outline the shape of the brain. The changes in the BOLD signal and the physiological noise might be correlated spatially, which can affect the assessment of the temporal dependence.

Our next step is to remove this spatially covarying structure using PCA. The idea is that PCA should capture nonlinear trend patterns commonly shared by the voxels after the linear detrending. A similar idea is employed by Huang et al. (2002) who use PCA to remove baseline artifacts in fMRI data. The PCA procedure projects the data onto the orthogonal subspace spanned by a few important principal component vectors. In other words, we decompose $V_L - \bar{V}_L = \sum_{k=1}^{K} s_k \mathbf{u}_k \mathbf{v}_k' + V_R$, where $\bar{V}_L$ is the mean matrix whose rows are 4096 replicates of the mean voxel value of each image, $\mathbf{u}_k$ and $\mathbf{v}_k$ are the kth left and right eigenvectors, and s_k is the kth singular value. Here K is the number of important eigenvalues (*i.e.* the eigenvalues departing from the expected distribution of a PCA of spatially independent data), which can be determined from the scree plot. The residuals, V_R, which are the PCA detrended data, are now the voxel values without linear trends and co-varying nonlinear trends. Thus, we use V_R to model the temporal correlation structure. A further study of our PCA detrending procedure including a simulation study can be found in Park et al. (2008).

Figure 1 (c) shows the PCA-based detrended image of the three slices at the same time point, $t = 900$. Note that the shape of the brain is no longer present. This suggests that the remaining spatially covarying signal components are removed with PCA detrending. To summarize, with our detrending procedure we remove all linear trends and as many spatially covarying components as possible from the data, resulting in a residual dataset that can be used for obtaining more accurate assessment of the temporal structure.

4 Multiscale data analysis

In this section, we model the detrended data V_R, *i.e.* the estimated temporal measurement error, by applying SiZer and wavelet spectrum analysis. Creating a SiZer plot or a wavelet spectrum in

Matlab takes less than 30 seconds on a PC Pentium 4.

Fig. 2 about here.

We choose two different voxels inside the brain for illustration purposes. The first and second voxels come from the middle and left parts of the first slice, respectively. Figure 2 shows conventional SiZer plots of the time series for the two selected voxels. The curves in the top panels display a family of local linear estimates obtained from the observations, some of which are displayed as jittered dots. These curves are indexed by different levels of smoothing, *i.e.* they are the $\hat{\beta}_0$ estimates of the regression function obtained with different values of h (see equation (2.2) and related discussion). Figure 2 (a) shows no significant features while (b) has several large and small oscillations. The lower panels of Figure 2 display the SiZer maps of the two voxels. The horizontal locations in the SiZer map are time; the vertical locations correspond to the same logarithmically spaced bandwidths that are used for the smooths in the top panel. Black (white) colors in the SiZer map indicate that the curves are significantly increasing (decreasing) at the point indexed by the horizontal location and the bandwidth corresponding to that row, as compared to white noise. Intermediate gray colors mean that the trend shows no increase or decrease that is significantly different from those typical of the white noise at that particular location. The dotted white curves in the SiZer maps show effective window widths for each bandwidth, as intervals representing $\pm 2h$, *i.e.*, ± 2 standard deviations of the Gaussian kernel. The SiZer map in Figure 2 (a) shows only gray, which means that the voxel is consistent with white noise. By contrast, the map in Figure 2 (b) has many features, indicating that this voxel has trends that differ significantly from white noise.

Fig. 3 about here.

To better understand the source of the significant features discovered in Figure 2 (b), Figure 3 shows the autocorrelation functions (ACF) of the same two voxel time series. The horizontal lines in these plots are 95% confidence bands under the null hypothesis of no autocorrelation. Figure 3 (a) shows no significant deviation, consistent with the SiZer plot in Figure 2 (a). However, Figure 3 (b) shows a strong long autocorrelation, which might be responsible for the significant features in Figure 2 (b).

This leads us to the wavelet spectrum approach for analyzing long-range dependent time series as introduced in Section 2.2. The Hurst parameter, H, obtained from the wavelet spectrum measures the degree of dependency. Values of H close to 0.5 indicate weak correlation (white noise), while

H close to 1 suggests strong long-range dependence.

Fig. 4 about here.

Figure 4 shows the wavelet spectra of the two sample voxels. We use Daubechies wavelets with two vanishing moments, *i.e.* $L = 2$ in (2.4), for all wavelet spectra. The choice of the number of vanishing moments L can be an issue in general since higher moments can reduce more non-stationarities in the data. This means that if the detrending procedure in Section 3 leaves some unexpected residual trends in the images, the wavelet spectrum with large L can make this up since it is robust to nonlinear trends. We tried $L = 3$ and our conclusions remained the same. The vertical segments on this plot are the estimated 95% confidence intervals of the log-mean-energy statistics of the wavelet spectrum. The estimates of the Hurst parameter are also reported on the plot. The wavelet spectrum in Figure 4 (a) looks flat and the estimated H is very close to 0.5, indicating white noise. On the other hand, the spectrum in Figure 4 (b) shows an upward linear trend and the estimated H is 0.69, an indication of long-range dependence. This finding is consistent with the SiZer and autocorrelation plots, indicating that the time series is different from white noise.

It is worth emphasizing again that estimation of the Hurst parameter would be unreliable with only a couple of hundred time points. This is because fewer time points result in fewer scales in the wavelet spectrum, and a weighted least squares fit with a small number of observations cannot be trusted. In our analysis, we have 1500 time points resulting in eight scales in the wavelet spectra with two vanishing moments, as seen in Figure 4. If we had, for example, 200 time points (typical for many fMRI studies even of noise), the spectra would have at best five scales. Thus, gathering longer time series is crucial in assessing the characteristics of long-range dependence.

Fig. 5 about here.

Finally, we estimate the Hurst parameter for every voxel in the data set. There are only a few voxels with $\hat{H} > 1$, which indicates non-stationary behavior of the time series. The non-stationary behavior is worth investigation, but is beyond the scope of this paper. Thus, we replace those estimates with 1. Also, there are some voxels with $\hat{H} < 0.5$, but either their 95% confidence bands cover 0.5 or there is no indication of deviation from independent white noise by SiZer analysis. Thus, we replace these with 0.5. We truncate $\hat{H}$ values at 0.5 and 1 because we are mainly interested in long range dependence, which is captured between those values. Furthermore, too large and/or small values change the scale of the images, which might cause misinterpretation of the maps. The

Hurst parameter estimates are arranged in three slices and the maps are shown in Figure 5.

In the top panels, voxels with $\hat{H} = 0.5$ are colored black, with $\hat{H} = 1$ are colored white, and with $0.5 < \hat{H} < 1$ are colored intermediate gray (the larger, the lighter). Interestingly, the shape of the brain is recovered in the maps. In the bottom panels, similar maps are displayed overlaid on the subject's brain. Voxels with $0.5 \leq \hat{H} \leq 0.65$ are colored black, with $0.65 < \hat{H} \leq 0.8$ are colored intermediate gray, and with $0.8 < \hat{H} \leq 1$ are colored white. The lighter region ($0.65 < \hat{H} \leq 1$) corresponds largely to the gray matter of the brain, which is where activity occurs, indicating that the working brain exhibits long-range dependence. As might be expected, the regions outside the brain are almost completely white noise. A few light voxels are found, for example in the upper middle part of the third slice, and they might be an indication of possible non-stationarities.

It is important to remember here that we have removed spatial covariance from the dataset. Therefore, the spatial map of the non-coherent temporal correlation structure indicates that the residual must be modeled with spatially non-uniform noise. Since the noise maps out obvious brain structures, the implication is that this noise is most likely due to non-coherent aspects of the BOLD signal that vary from place to place in the brain.

Fig. 6 about here.

Figure 6 (a) gives the distribution of the Hurst parameter estimates shown in Figure 5. The estimates from the three slices are combined together. The plot shows that about 75% of the voxels have an estimated Hurst index of less than 0.6. Since the regions outside the brain are almost all white noise, we draw the distribution of the Hurst parameter estimates inside the brain only in Figure 6 (b). We can see that more than half of the voxels are weakly correlated, but about 17% of the voxels have H estimates greater than 0.7. Thus, our multiscale approach clearly shows that even for resting data some voxel time series in the brain cannot be modeled by white noise or simple short-term dependence such as AR(1) and need long-range dependent type error structure, for example fractional Gaussian noise (fGn).

One possible way of finding specific models for these voxel error time series is to construct dependent SiZer. Since dependent SiZer conducts a goodness of fit test with an assumed time series model, it can help guide us to appropriate models for the data.

Fig. 7 about here.

Figure 7 shows two different types of time series modeling: (a) AR(1) and (b) fGn, for the voxel shown in Figure 2 (b). Since the family of smooths will be exactly the same as in Figure 2

(b), we only report the SiZer maps. Figure 7 (a) shows the SiZer map of AR(1) with its estimated coefficient $\hat{\phi} = 0.19$. The SiZer map reports the tests under the null hypothesis of AR(1) with the specified parameter instead of white noise. Compared to Figure 2 (b) some features in the small and middle scales which were flagged as being significant due to the dependent structure, have disappeared. However, since most big features are still significant, AR(1) is not a good model for the error. Figure 7 (b) shows the SiZer map of fGn with the Hurst parameter $\hat{H} = 0.7$, which is estimated from the wavelet spectrum in Figure 4 (b). Now, all the significant features in Figure 2 (b) and Figure 7 (a) have disappeared. Thus, fGn with $H = 0.7$ is a valid model for this voxel error. Figures 7 (a) and (b) show a fourth color, darker gray, in the lower corners of the SiZer maps. This color represents areas where the data are too sparse for SiZer to be effective. Thus, no statistical test is performed in these regions.

To save space, we do not report all the results, but we found similar lessons from the rest of the data. As shown in Figure 5, some regions of the brain can be modeled by white noise (Figure 2 (a)) or weakly correlated models as many researchers have assumed, but some parts of brain should be modeled by long-range dependent process (Figure 7 (b)), not by white noise (Figure 2 (b)) nor AR(1) (Figure 7 (a)).

To check the consistency of the results of this analysis, we collected data under a similar imaging paradigm, using the 3T scanner at the Biomedical Imaging Research Center at UGA. Again, three slices with 1500 time points were collected from a subject, and the Hurst parameters were estimated for all $64 \times 64 \times 3 = 12288$ voxels after the linear and PCA detrending. The Hurst parameter map from the new data looked similar to Figure 5 and the lesson learned remained the same. We do not report the result of the second data set to save space.

5 Discussion

The goal of this paper was to assess the temporal measurement error structure of resting-state data using multiscale tools. We collected longer time series (1500 time points) than usual fMRI data (usually a couple of hundred time points) since measuring long-range dependence requires long time series for accurate assessment.

Our multiscale approach clearly shows that even for resting data, *i.e.* "null" or ambient thought, different parts of the brain need different models, for example some voxel time series cannot be modeled by white noise or an AR(1) type model and need long-range dependent type error structure. In practice most fMRI time series consist of only a few hundred time points, and thus it would be hard to demonstrate existence of long-range dependence, but our findings provide possible models

for fMRI voxel time series. In our analysis we modeled these voxels as fGN with different Hurst parameters. The fact that one cannot use a single noise/null model for all the brain should be considered for other fMRI data analyses. One could use a long-range dependent model for the noise dependence structure and estimate the Hurst parameter from the data. An alternative approach, which would be more appropriate when there are not many data points, would be using the wavelet transform. The wavelet transform will decorrelate long-range dependence and then the noise will be independently or weakly correlated in the wavelet domain. In this domain, one could use, for example, a generalized linear model to identify the activated area (Fadili and Bullmore, 2002). One can consider alternative models to account for temporal correlations, for example, AR with different orders. In such a case SiZer tools can be applied to determine proper orders of voxels. We suggest the comparison of different time series models as future work.

Our SiZer analysis has focused on assessing the structure of temporal correlation in voxel time series. However, there also exist spatial correlations between voxels. Thus, the spatial counterpart of the autocorrelation function, the variogram, should be considered. This can be used to construct a spatial SiZer based on a local linear fit of the variogram. The estimated variogram will be taken into account for constructing the covariance matrix of the test statistics in the new SiZer.

The results of our analysis can be further strengthened by analyzing multiple subjects. We plan to extend the ideas developed in this paper to multiple subjects and also to comparing groups of subjects. The application of multiscale methods to individual subject is straightforward, however, discovery of their commonalities and comparison of groups will be a challenging component of the future work.

Acknowledgement

We would like to thank Nathan Yanasak for his assistance with data acquisition, Cynthia Krafft and Qingyang Li for their help with image creation and interpretation of results, and Jennifer McDowell for her help with the physiological interpretation of aspects of the fMRI signal.

We also thank the referees for their helpful comments. Lazar's work was partially supported by the M.G. Michael Award, Franklin College of Arts and Science, University of Georgia. Ahn's work was partly supported by NSF Grant DMS-0805758.

References

Abry, P. and Veitch, D. (1998). Wavelet analysis of long–range dependent traffic. *IEEE Transactions on Information Theory*, 44:2–15.

Bullmore, E., Brammer, M., Williams, S., Rabe-Hesketh, S., Janot, N., David, A., Mellers, J., Howard, R., and Sham, P. (1996). Statistical methods of estimation and inference for functional MR image analysis. *Magnetic Resonance in Medicine*, 35:261–277.

Bullmore, E., Fadili, J., Breakspear, M., Salvador, R., Suckling, J., and Brammer, M. (2003). Wavelets and statistical analysis of functional magnetic resonance images of the human brain. *Statistical Methods in Medical Research*, 12:375–399.

Chaudhuri, P. and Marron, J. S. (1999). SiZer for exploration of structures in curves. *Journal of the American Statistical Association*, 94:807–823.

Daubechies, I. (1992). *Ten Lectures on Wavelets*. SIAM, Philadelphia.

Fadili, M. and Bullmore, E. (2002). Wavelet-generalized least squares: A new BLU estimator of linear regression models with $1/f$ errors. *NeuroImage*, 15:217–232.

Fan, J. and Gijbels, I. (1996). *Local Polynomial Modelling and Its Applications*. Chapman & Hall, London.

Fox, M. and Raichle, M. (2007). Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Neuroscience*, 8:700–711.

Gautama, T. and Van Hulle, M. M. (2005). Estimating the global order of the fMRI noise model. *NeuroImage*, 26:1211–1217.

Hannig, J. and Marron, J. S. (2006). Advanced distribution theory for SiZer. *Journal of the American Statistical Association*, 101:484–499.

Huang, C.-C., Liou, M., and Cheng, P. E. (2002). Baseline correction of functional MR time courses with PCA. *Proceedings of the 9th International Conference on Neural Information Processing*, 1:433–435.

Lindeberg, T. (1994). *Scale Space Theory in Computer Vision*. Kluwer, Boston.

Locascio, J., Jennings, P., Moore, C., and Corkin, S. (1997). Time series analysis in the time domain and resampling methods for studies of functional magnetic resonance brain imaging. *Human Brain Mapping*, 5:168–193.

Long, C., Brown, E., Triantafyllou, C., Aharon, I., Wald, L., and Solo, V. (2005). Nonstationary noise estimation in functional MRI. *NeuroImage*, 28:890–903.

Lund, T., Madsen, K., Sidaros, K., Luo, W., and Nichols, T. (2006). Non-white noise in fMRI: Does modelling have an impact? *NeuroImage*, 29:54–66.

Luo, W. and Nichols, T. E. (2003). Diagnosis and exploration of massively univariate neuroimaging models. *NeuroImage*, 19:1014–1032.

Maxim, V., Sendur, L., Fadili, J., Suckling, J., Gould, R., Howard, R., and Bullmore, E. (2005). Fractional Gaussian noise, functional MRI and Alzheimer‘s disease. *NeuroImage*, 25:141–158.

Park, C., Hernandez Campos, F., Le, L., Marron, J. S., Park, J., Pipiras, V., Smith, F. D., Smith, R. L., Trovero, M., and Zhu, Z. (2010). Long-range dependence analysis of Internet traffic. *To appear in Journal of Applied Statistics*.

Park, C., Lazar, N. A., Ahn, J., and Sornborger, A. (2008). Do different parts of the brain have the same dependence structure? A multiscale analysis of the temporal and spatial characteristics of resting fMRI data. Technical Report, Department of Statistics, University of Georgia.

Park, C., Marron, J. S., and Rondonotti, V. (2004). Dependent SiZer: goodness of fit tests for time series models. *Journal of Applied Statistics*, 31:999–1017.

Penny, W., Kiebel, S., and Friston, K. (2003). Variational Bayes inference for fMRI time series. *NeuroImage*, 19:727–741.

Purdon, P. L., Solo, V., Weisskoff, R. M., and Brown, E. N. (2001). Locally regularized spatiotemporal modeling and model comparison for functional MRI. *NeuroImage*, 14:912–923.

Raichle, M. and Snyder, A. (2007). A default mode of brain function: A brief history of an evolving idea. *NeuroImage*, 37:1083–1090.

Rondonotti, V., Marron, J. S., and Park, C. (2007). SiZer for time series: a new approach to the analysis of trends. *Electronic Journal of Statistics*, 1:268–289.

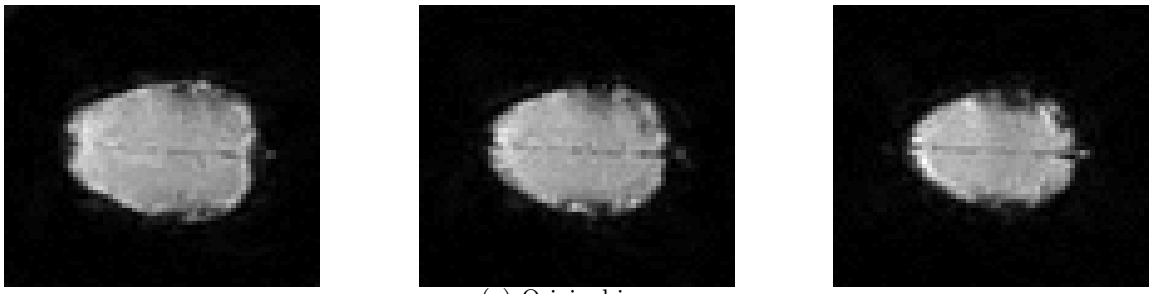
Stoev, S., Taqqu, M., Park, C., and Marron, J. S. (2005). On the wavelet spectrum diagnostic for hurst parameter estimation in the analysis of Internet traffic. *Computer Networks*, 48:423–445.

Veitch, D. and Abry, P. (1999). A wavelet based joint estimator for the parameters of LRD. *IEEE Transactions on Information Theory*, 45:878–897.

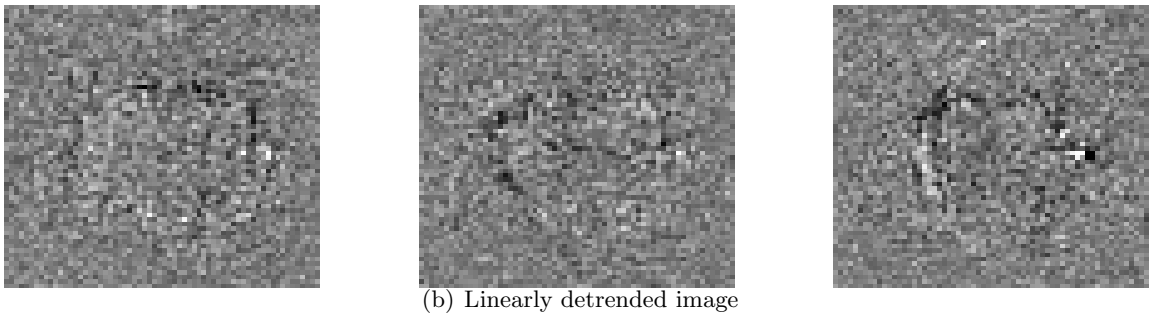
Woolrich, M. W., Ripley, B. D., Brady, J. M., and Smith, S. M. (2001). Temporal autocorrelation in univariate linear modeling of fMRI data. *NeuroImage*, 14:1370–1386.

Worsley, K. J., Marrett, S., Neelin, P., and Evans, A. C. (1996). Searching scale space for activation in PET images. *Human Brain Mapping*, 4:74–90.

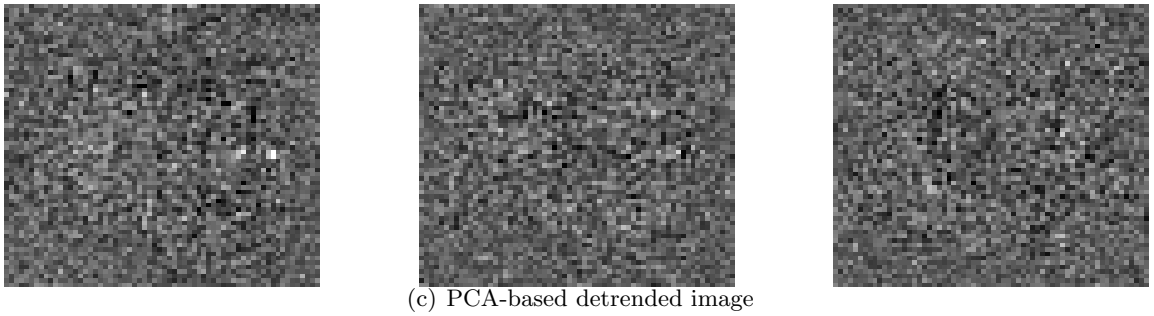
Worsley, K. J., W. M. and Evans, A. C. (1997). ICA of functional MRI data: An overviewscale space searches for a periodic signal in fMRI data with spatially varying hemodynamic response. In *Proceedings of BrainMap '95 Conference*.



(a) Original image

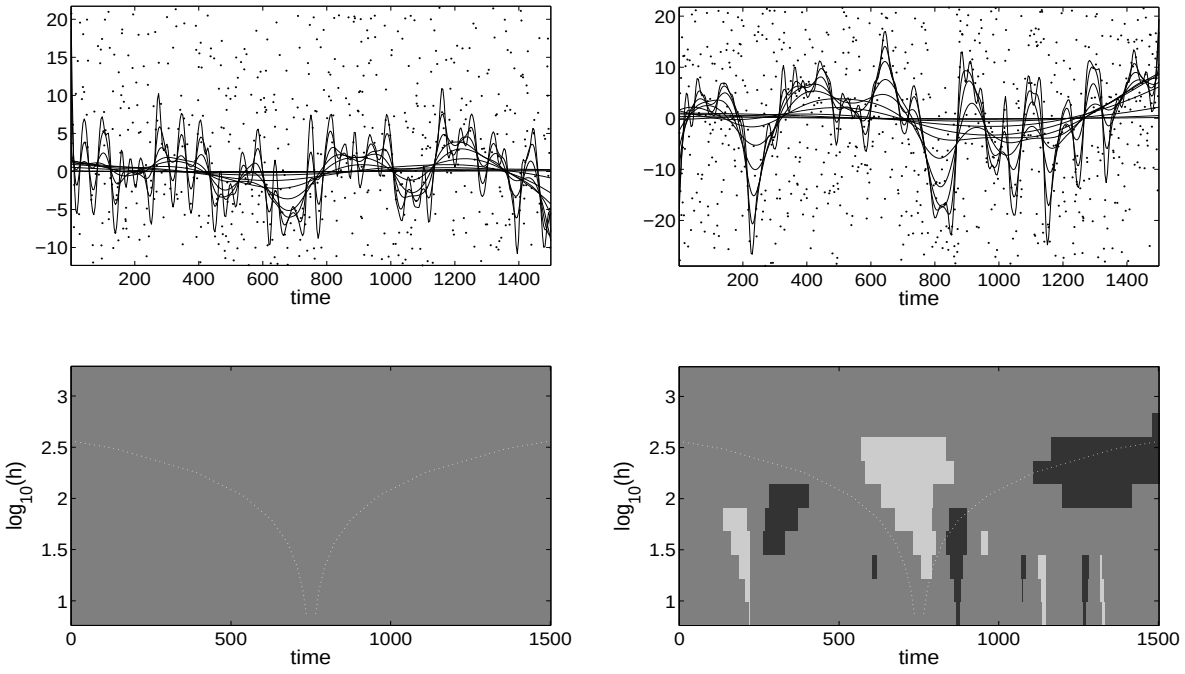


(b) Linearly detrended image



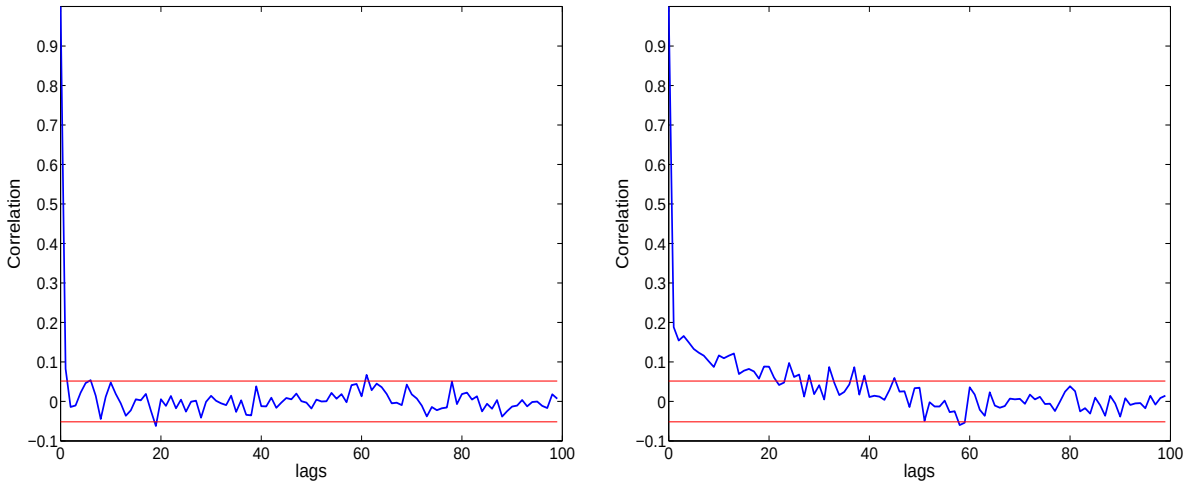
(c) PCA-based detrended image

Figure 1: Brain images at time point 900: (a) original image, (b) linearly detrended using a linear fit, and (c) PCA detrended using first eight PCs. Each row is drawn at the same scale. The shape of the brain is not seen in (c).



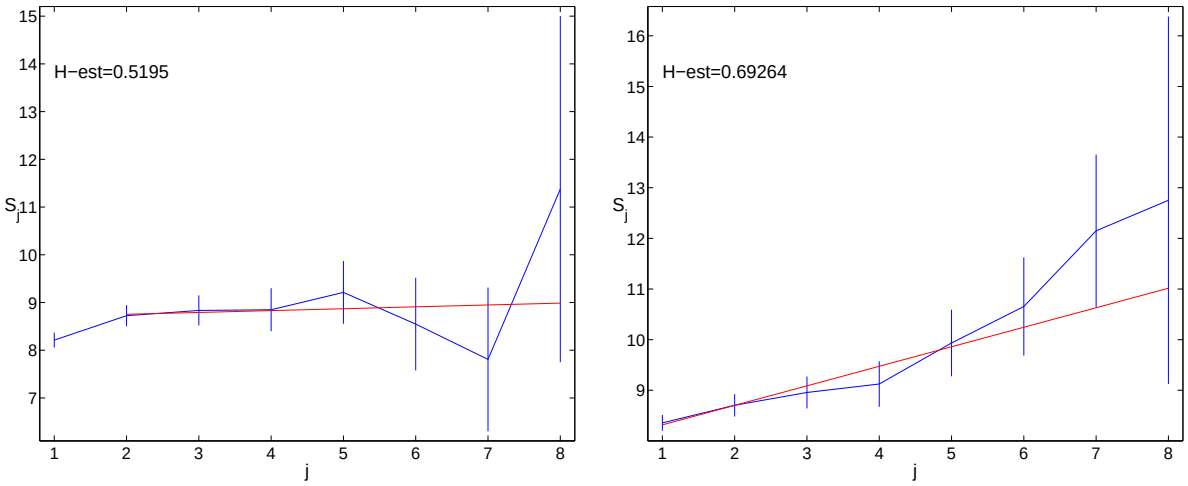
(a) (b)

Figure 2: SiZer plots of two voxels. (a) shows no significant features meaning that the voxel can be modeled by white noise. (b) shows several significant features meaning that the voxel cannot be modeled by white noise.



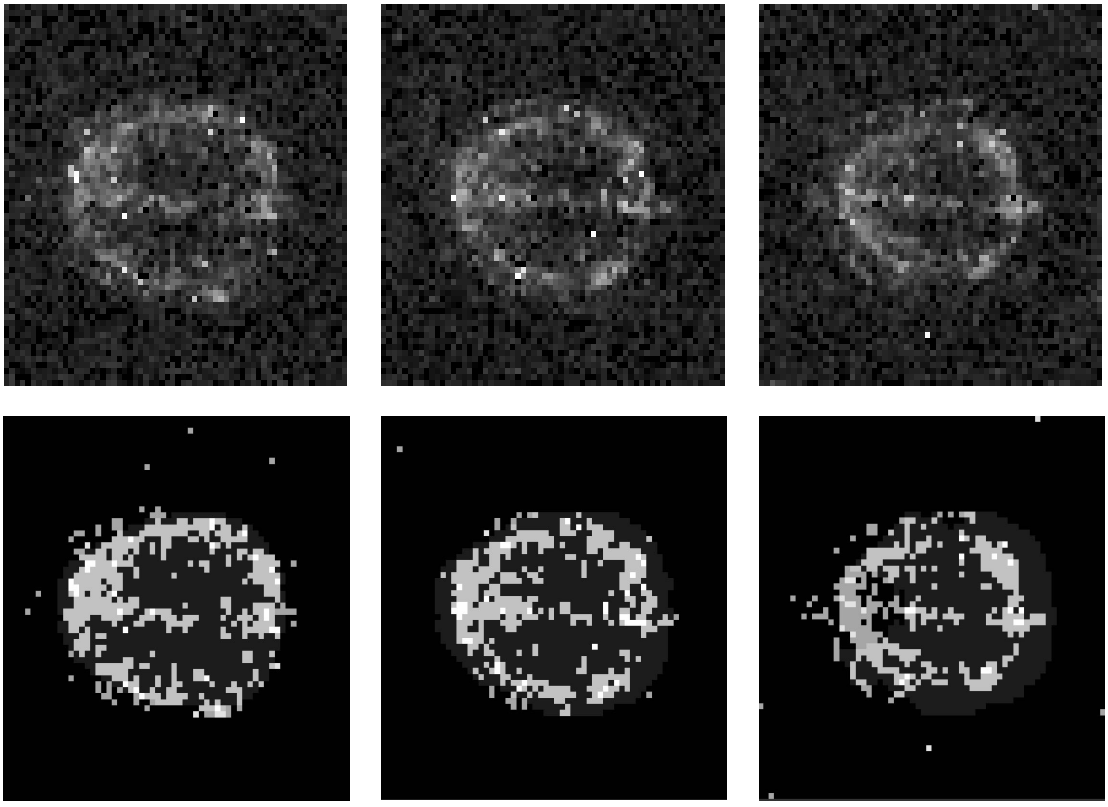
(a) (b)

Figure 3: Autocorrelation functions of the two voxels in Figure 2. (a) ACF decreases quickly and remains within the 95% confidence band of no correlation. (b) ACF decreases slowly, an indication of long-range dependence



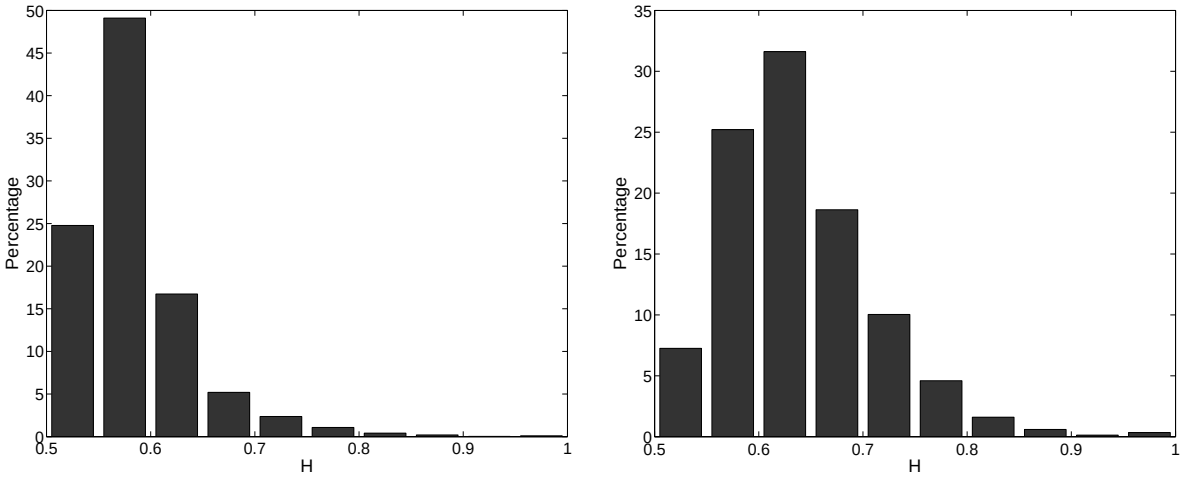
(a) (b)

Figure 4: Wavelet spectra of the two voxels in Figure 2. (a) The spectrum shows no trend and the estimated H is close to 0.5. (b) The spectrum shows a linear trend and the estimated H is close to 0.7, which indicates the existence of long-range dependence.



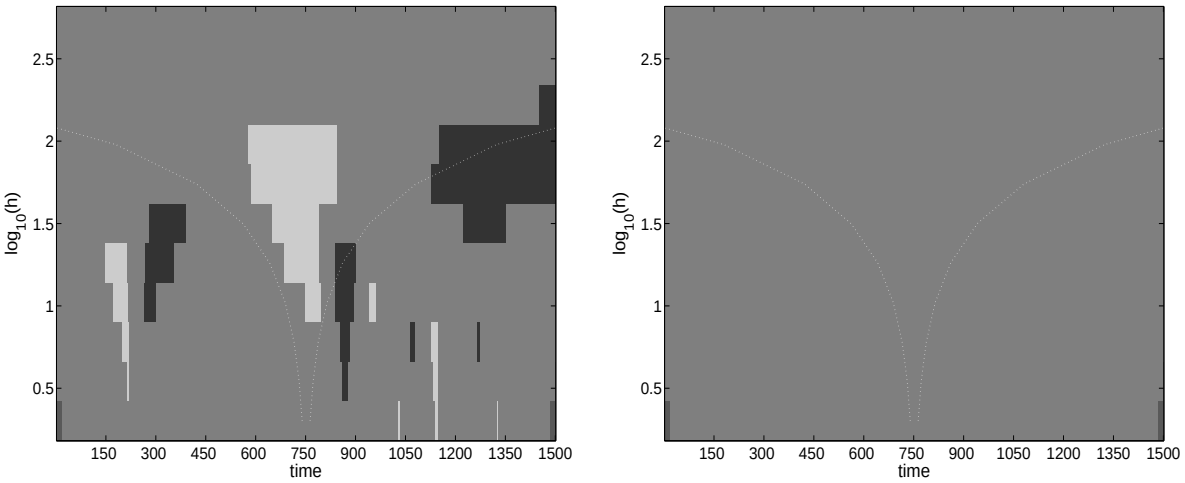
(a) Slice 1 (b) Slice 2 (c) Slice 3

Figure 5: Top panels: the maps of the Hurst parameter estimates (between 0.5 and 1) for the three slices. Bottom panels: the maps of the Hurst parameter estimates overlaid on the subject's brain for the three slices. Regions with brighter colors have higher $\hat{H}$, a suggestion of long-range dependence.



(a) All voxels (b) Voxels inside the brain

Figure 6: Histogram of the Hurst parameter estimates (all three slices combined).



(a) AR(1) (b)fGn $H = 0.7$

Figure 7: Dependent SiZer maps of the voxel shown in Figure 2 (b) under the null hypothesis of (a) AR(1) and (b) fGn with $H = 0.7$. Under AR(1), the SiZer map shows many significant features, as when the null hypotheses is white noise, Figure 2 (b). Under fGn with $H = 0.7$, no features are found, meaning that the voxel can be modeled by fGn with $H = 0.7$.