A Non-Linear Model for Phase-Only fMRI Data

Christopher P. Meller[1] and Daniel B. Rowe[1,2,*]
Division of Biostatistics[1] and Department of Biophysics[2]

Division of Biostatistics
Medical College of Wisconsin

Technical Report 50

December 2004

Division of Biostatistics
Medical College of Wisconsin
8701 Watertown Plank Road
Milwaukee, WI 53226
Phone:(414) 456-8280

A Non-Linear Model for Phase-Only fMRI Data

Christopher P. Meller[1] and Daniel B. Rowe[2*]
Division of Biostatistics[1] and Department of Biophysics[2]
Medical College of Wisconsin
Milwaukee, WI USA

Abstract

Magnitude-only data models have dominated fMRI analysis since the inception of the technology. However, when such analyzes are done half the data is disregarded which may or may not contain valuable information. This overlooked data is called the phase and is restricted to values between $-\pi$ and π. When phase time series are analyzed, ordinary least squares (OLS) regression has been the technique of choice. However, OLS models perform poorly when "wrap-around" or low SNR is present. We have explored alternatives to the OLS model which will account for the angular response of the phase while also allowing us the flexibility to develop similar hypothesis tests. We adopt a model by Fisher and Lee (1992) to our analysis of the generally discarded fMRI phase time series and show its improvement over the OLS model for both parameter estimation and data prediction for various conditions. We fit the model to both simulated data and acquired data from an actual fMRI experiment and found an improvement in parameter estimation along with modeling for the Fisher and Lee method in the simulated data while detailing potential benefits when used with experimental acquired data. Finally, we look at a map of statistics relating association of the observed voxel phase time courses with a reference function in our acquired data and show the possible detection of biological information in the generally discarded phase.

1 Introduction

Functional magnetic resonance imaging (fMRI) is an invaluable tool used by clinicians and physicians alike to investigate and diagnose biological phenomena in both animals and humans. However, the effectiveness of fMRI is directly related to the appropriateness of the model used to analyze the data. We will look at modeling the phase along with making inference dealing with possible experiment related phases changes termed, phase "activations."

*Corresponding Author: Daniel B. Rowe, Department of Biophysics, Medical College of Wisconsin, 8701 Watertown Plank Road, Milwaukee, WI 53226, dbrowe@mcw.edu.

Those areas of the brain that exhibit cognitive function to different types of stimuli is of interest to investigators. When neurons in a given location synapse for brain function they use up bound oxygen in the hemoglobin of oxygenated blood. The physiological response in the brain is to overcompensate with a surplus of oxygenated blood in that location. It is the change in the magnetic field in this location due to an increase of oxygenated blood that is measured and not neuronal firing. A statistically significant increase in oxygenated blood in a location or voxel associated with the presentation of a stimulus is called activation [1].

Currently, the most common method being used to detect activation in fMRI is a general linear model with magnitude-only data. However as Rowe and Logan (2004) pointed out, this model disregards all the phase data and any information it may contain. We will look at modeling the generally discarded phase data by using a model which is more flexible than the ordinary least squares (OLS) regression model, the current standard for this type of analysis. Specifically, we will implement a model developed by Fisher and Lee (1992) on both simulated and actual experimental fMRI time series data and lay the framework for making inferences about the possible phase activations. In doing this we will have accounted for the angular nature of the response and provided an improvement in dealing with the "wrap-around" issue. In the past the only way to account for "wrap-around" in the data was to artifically "unwrap" it and proceed with an OLS analysis.

2 Background

It is important for any investigator to understand the underlying distributional assumptions in fMRI data. The fMRI voxel time courses are complex-valued with normally distributed noise on each component. If we denote (y_R, y_I) as the real and imaginary observed signals, (ρ_R, ρ_I) as the real and imaginary true signals, and (η_R, η_I) as the respective noise components then,

$$\begin{aligned} y_R &= \rho_R + \eta_R \\ y_I &= \rho_I + \eta_I. \end{aligned} \tag{2.1}$$

The noise is specified be normally distributed with mean zero and covariance matrix Σ denoted, $(\eta_R, \eta_I)' \sim N(0, \Sigma)$ where $\Sigma = \sigma^2 I_2$. Introducing a phase imperfection, θ that couples the means of the real and imaginary components, we can rewrite (ρ_R, ρ_I) as $(\rho \sin\theta, \rho \cos\theta)$. Until further mention it should be understood that we are dealing with a single point in a time series and are omitting the subscript, t.

2.1 Magnitude-Only Data

The above mentioned complex-valued data is commonly transformed into polar coordinates, $r = \sqrt{y_R^2 + y_I^2}$ and $\phi = \tan^{-1}(y_I/y_R)$, referred to as the magnitude and phase, respectively. It is important to understand that the distributional specifications are on the real and imaginary parts of the signal and not on the

magnitude (or phase). Rowe and Logan (2004) derived the joint distribution of r and ϕ and found its probability density function (PDF) had the following form [11],

$$f(r,\phi) \quad = \quad \frac{\rho}{2\pi\sigma^2} e^{-\frac{1}{2\sigma^2}\left[r^2+\rho^2-2\rho r\cos(\phi-\theta)\right]}, \quad r,\rho,\sigma^2>0;\ -\pi<\phi,\theta\leq\pi. \tag{2.2}$$

As a result we know that when the above transformation is made and the magnitude is marginalized, i.e. the phase data is discarded, the resulting distribution is a Ricean,

$$f(r) \quad = \quad \frac{r}{\sigma^2} e^{-\frac{r^2+\rho^2}{2\sigma^2}} \int_0^{2\pi} e^{-\frac{r\rho}{\sigma^2}\cos(\alpha)} d\alpha, \quad r,\rho,\sigma^2>0. \tag{2.3}$$

The integral term is generally denoted by $I_0(r\rho/\sigma^2)$, a zeroth order modified Bessel function of the first kind. If we define signal-to-noise ratio (SNR) to be ρ/σ we see that the magnitude is Rayleigh distributed for SNR=0. If the SNR is "large," say greater than 10, then the distribution limits to a normal with mean ρ and variance σ^2. Additionally, Rowe and Logan use a Taylor expansion to show that the magnitude-only data can be correctly modeled by an OLS regression for large SNR [11].

2.2 Phase-Only Data

Similar to the magnitude we derive the marginal distribution for the phase and observe the following density function,

$$f(\phi) = \frac{e^{-\frac{\rho^2}{2\sigma^2}}}{2\pi}\left[1+\frac{\rho\ \cos(\phi-\theta)}{2\pi\sigma} e^{\frac{\rho^2\ \cos^2(\phi-\theta)}{2\sigma^2}} \Phi\left(\frac{\rho\ \cos(\phi-\theta)}{\sigma}\right)\right], \quad -\pi<\phi,\theta\leq\pi;\ \rho,\sigma>0. \tag{2.4}$$

We define $\Phi(\cdot)$ to be the cumulative density function for the standard normal distribution. Parallel with the magnitude, the phase has a limiting normal distribution with mean θ and variance $(\sigma/\rho)^2$ for large SNR. This limiting distribution allows OLS to be used to fit a model to this data. However, phase data lies within an interval of $(-\pi,\pi]$ which leaves it vulnerable to "wrap-around" and can cause the OLS model fit to be questionable. We define "wrap-around" to be the observation of successive phase measurement points with jumps greater than $|2\pi|$.

Consistent with the Talyor series argument of Rowe and Logan dealing with the magnitude measurements, a similar argument can be made to argue in support of a normal limiting distribution for the phase measurements in the time series. Introducing the subscript, t, for each time point combined with the limiting distribution we are able to model,

$$\phi_t = \tan^{-1}\left[\frac{\rho_t\sin(\theta_t)+\eta_{I_t}}{\rho_t\cos(\theta_t)+\eta_{R_t}}\right], \quad t=1,\ldots,n, \tag{2.5}$$

with a normal distribution and an OLS model.

3 Models

As previously stated, phase data is rarely analyzed in fMRI, and the times it is, OLS is the modeling methodology of choice [2]. Before an OLS model can be fit, the time series needs to be artifically "unwrapped." The process of artifically "unwrapping" the data and then fitting an OLS model to it will be

described shortly. Menon (2002) looked at phase time series after unwrapping and noticed an approximate linear association with the magnitude data which he believed was attributed to large blood vessels [9]. The OLS model will be briefly summarized then an alternative model will be introduced along with its advantages and disadvantages.

3.1 Ordinary Least Squares

The OLS phase-only data model can be written in the following form,

$$\phi_t = z_t'\gamma + \delta_t, \tag{3.1}$$

where ϕ_t are observed phase angle measurements at each time point, t, for an arbitrary voxel, z_t' is a specific row of the design matrix Z in Equation 3.2, γ are the fixed but unknown phase regression coefficients, and δ_t is the measurement error. This measurement error is assumed to originate from the following normal distribution, $N(0, \tau_t^2)$ where $\tau_t^2 = \sigma^2/\rho_t^2$, and the design matrix for our example Z is constructed in the following way for three covariates,

$$Z = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ \vdots & \vdots & \vdots \\ 1 & 17 & -1 \\ \vdots & \vdots & \vdots \\ 1 & 256 & -1 \end{pmatrix}. \tag{3.2}$$

The above design matrix includes a column of 1's to model the intercept, a column of counting numbers to account for a possible linear trend over time [13], and a third column with alternating sets of 16 1's and -1's to account for the a task related function. Sets of 16 are used in our matrix to coincide with both our simulated and acquired data which have stimulus lengths of 16 seconds. We define SNR by $\rho_t/\sigma = (\beta_0 + \beta_1 t + \beta_2 z_{2t})/\sigma$ but $\beta_1 t + \beta_2 z_{2t}$ is generally very small when compared to β_0 within the brain and zero outside the brain so we utilize the following approximation, $\rho_t/\sigma \approx \beta_0/\sigma$ and thus the variance becomes constant over time. When we describe the phase-only data with the OLS model we can make inferences using standard linear contrasts tests, H_0: $D\gamma = 0$ vs H_1: $D\gamma \neq 0$. Our contrast matrix, D, is defined to be a $w \times (q+1)$ matrix. This testing procedure leads to the unrestricted, $(\hat{\gamma}, \hat{\tau}^2)$, and restricted, $(\tilde{\gamma}, \tilde{\tau}^2)$, maximum likelihood estimates for γ and τ^2 having the familiar form,

$$\hat{\gamma} = (Z'Z)^{-1}Z'\phi \tag{3.3}$$

$$\hat{\tau}^2 = \frac{(\phi - Z\hat{\gamma})'(\phi - Z\hat{\gamma})}{n} \tag{3.4}$$

$$\tilde{\gamma} = \Psi\hat{\gamma} \tag{3.5}$$

$$\tilde{\tau}^2 = \frac{(\phi - Z\tilde{\gamma})'(\phi - Z\tilde{\gamma})}{n} \tag{3.6}$$

and $\Psi = I - (Z'Z)^{-1}D'[D(Z'Z)^{-1}D']^{-1}D$. Using the asymptotic χ^2_w distribution of -2 log of a likelihood ratio test we note that the test statistic is

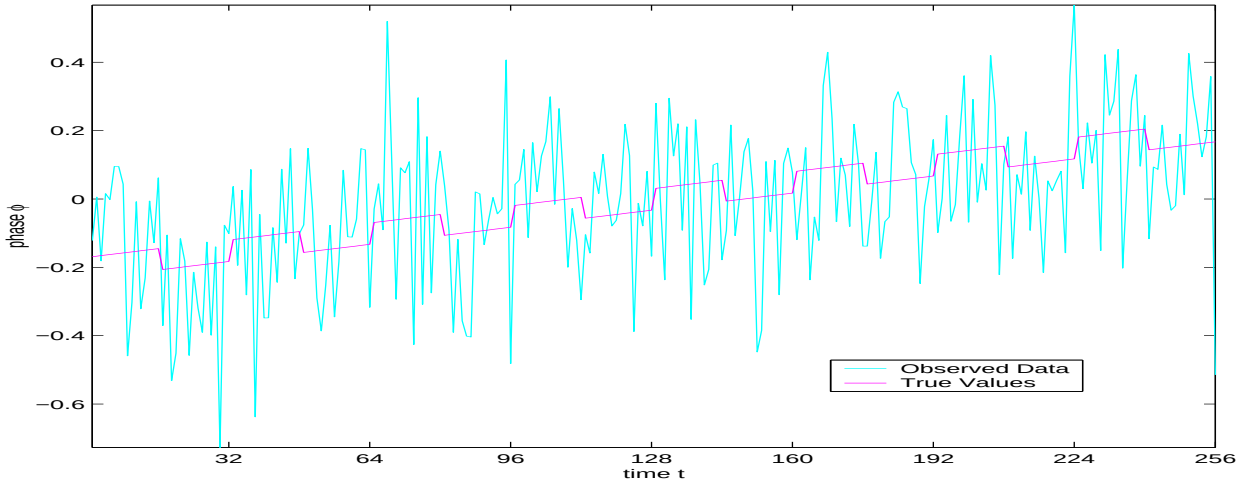
$$-2\log\lambda = n\log\left(\frac{\tilde{\tau}^2}{\hat{\tau}^2}\right). \tag{3.7}$$

The degrees of freedom, w, is equal to the full row rank of D [10]. We also may test individual hypotheses by using the large sample normal statistics due to the number of time points present.

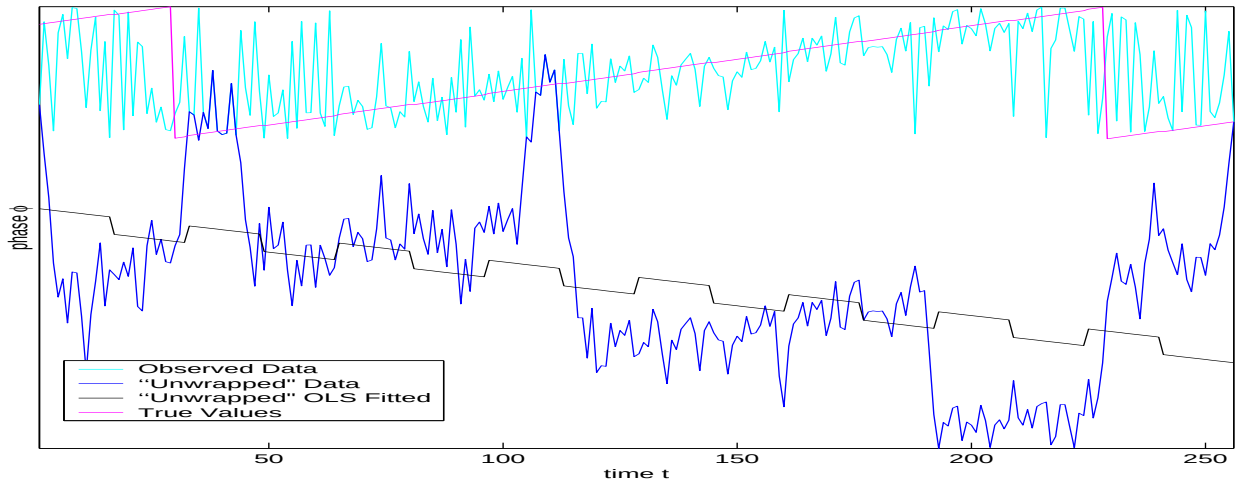
The angular response of the phase-only OLS model can cause modeling problems at the wrap around junction if not accounted for properly thus resulting in poor parameter estimation and incorrect inferences. When a combination of the three following conditions are valid we can describe the phase data with little concern using our standard OLS regression model. These conditions are 1) a large SNR, 2) our baseline angle, γ_0, is not near either the upper boundary of π or the lower boundary of $-\pi$, and 3) the linear trend is small enough that the data does not rise or lower beyond π and $-\pi$. The large SNR assumption makes the probability that successive measurements have a large difference between them to be very small. A simulated time series following the aforementioned conditions is plotted in Figure 1(a) with both the observed data (cyan) and true vales (magenta).

However, in real acquired fMRI data the mean phase in voxels varies greatly and these stringent conditions are in general not met. Often within an fMRI data set phase angles are observed close to $\pm\pi$. An example of phase data which "wraps" around the phase boundary is seen in Figure 1(b) where the cyan time series is the original data and the blue time series is the observed data "unwrapped." As previously mentioned, the method generally used to deal with this issue is "unwrapping." Unwrapping is the process of beginning with the first observation, proceeding through the time series and "flagging" an observation in which the next point in the time series have an absolute difference greater than or equal to a predefined value, generally π, then shifting the rest of the time series by $\pm 2\pi$. The process is repeated to the end of the time series. However, when high SNR becomes suspicious the model fails and the investigator needs another model to work with.

Often times simply "unwrapping" the data and fitting an OLS regression line to the data is sufficient as in Figure 1(a). This situation can be defined as "control" data in which no "wrap-around" is present and the SNR is large. However, as we see in Figure 1(b) this isn't always true. We see the fitted OLS regression line (black) for the "unwrapped" data (blue) doesn't model the data correctly and often this will lead to incorrect inferences. The fitted regression line has a negative slope while the true value of the slope coefficient is positive. This implies the need for another model that can account for the angular response. Simply "unwrapping" the data does not guarantee OLS to be a correct method to deal with the "wrap-around" issue, especially when the SNR is low. Instead, we explore various angular models and look at fitting the data one. Implementing an angular model procedure to deal with the angular nature of the response will allow us to relax the large SNR requirement.



(a) "Control" fMRI Phase Data



(b) Original, "Unwrapped" Phase Data, and OLS Fitted Line

Figure 1: Part a: Shows data where no "unwrapping" is needed. Part b: Shows an Example of an improper OLS fitted line (Black) to "unwrapped" phase time series (Blue) along with the original data (Cyan) and the true values (magenta).

3.2 Circular Model

As shown previously, fitting an OLS regression model to artifically "unwrapped" data may not be an ideal method and is especially poor for low and moderate SNR data. We explore models currently in the statistical literature that deal with angular responses and linear predictors given by Gould [5], Johnson and Wehrly [14] along with Fisher and Lee [4]. The type of model with linear explanatory variables and an angular response is referred to as linear-circular regression [6]. We will describe these three models and show that the one by Fisher and Lee gives us the most flexibility because it is easily adapted to our fMRI phase data. This will be done by discussing the benefits and limitations of each model.

The Von Mises distribution, also known as the circular normal distribution, is the underlying distribution in each of the angular models we will discuss [6]. This distribution specifically deals with random variables on the interval of $(-\pi, \pi]$ and is the conditional distribution for the phase given the magnitude. These facts alone should give the investigator some confidence using models based on a Von Mises distribution when exploring fMRI phase time series because of this close relationship. The Von Mises distribution, $f(\phi)$ for ϕ, as derived in the Appendix has the following PDF with mean θ and concentration κ,

$$f(\phi) = \frac{e^{\kappa \cos(\phi - \theta)}}{2\pi I_0(\kappa)}, \quad -\pi \leq \phi, \theta < \pi, \quad 0 \leq \kappa < \infty, \tag{3.8}$$

where $I_0(\cdot)$ is a zeroth order modified Bessel function of the first kind. The von Mises distribution has a limiting normal distribution with mean θ and variance $1/\kappa$ when κ is large.

As describe in Jammalamadaka and SenGupta (2001), Gould (1969) introduced a model which estimated the overall mean direction of each data point using the following form,

$$\theta_t = \gamma_0 + \sum_{j=1}^{m} \gamma_j z_{jt} (\text{mod } 2\pi)$$

with $j = 1, \ldots, m$, where m is the number of covariates, and $t = 1 \ldots n$, n being the number of time points in a single voxel. The modulus, denoted by mod, is defined to be the remainder of each $\gamma_j z_{jt}$ quantity when dividing by the argument, 2π. The log likelihood for this angular regression with a single regressor, γ, i.e. $m = 1$, case is

$$\log L = -n \log 2\pi \ + \ \kappa \sum_{t=1}^{n} \cos(\phi_t - \gamma_0 - z_{1t}\gamma_1) - n \log I_0(\kappa).$$

As discussed by Johnson and Wehrly, this model has some advantages [7]. These advantages include the ability to test $\gamma_1 = 0$ and the ability to estimate the unknown parameters using an iterative process which Gould laid out in his paper. Yet, this model has certain limitations that arise when considering it for our fMRI phase time series analysis. This model assumes the angle continuously "wraps" around resulting in what is referred to as the "barber's pole" form. However, in fMRI data the phase angle rarely wraps around more than once unless the SNR is extremely low as would by seen outside the brain. Johnson and Wehrly also discussed the fact that the MLEs are non-unique in this model because the likelihood function has infinitely many equally high peaks.

Johnson and Wehrly (1978) also introduced an alternative model to be used for angular data with linear predictors. Their model assumed that the predictor variable, z_{1t}, is random and has a known distribution. This allows them to derive the joint distribution of ϕ_t and z_{1t},

$$f(\phi_t, z_{1t}) \quad = \quad 2\pi g\left[2\pi(F_1(\phi_t) - F_2(z_{1t}))\right] f_1(\phi_t) f_2(z_{1t}),\ 0 \leq \phi_t < 2\pi,\ -\infty < z_{1t} < \infty$$

where $F(\phi_t), F(z_{1t}), f(\phi_t), f(z_{1t})$ are marginal cumulative density functions (CDFs) and marginal PDFs for ϕ_t and z_{1t} respectively. The function $g(\cdot)$ is an a priori defined density on a unit circle. This density $g(\cdot)$ is often chosen to be the Von Mises distribution because it is the most well known. The authors accordingly chose $g(\cdot)$ to be the Von Mises distribution for their example, making the conditional distribution of ϕ_t given z_{1t} to be the following,

$$f(\phi|Z) \quad = \quad [2\pi I_0(\kappa)]^{-n}\, e^{\kappa \Sigma_{t=1}^{n} \cos(\phi_t - \gamma_0 - 2\pi F(z_{1t}))}$$

where $\phi = (\phi_1, \ldots, \phi_n)'$. The authors also show that the maximum likelihood estimates for the mean angle, γ_0, and the concentration, κ, are well defined for this model. This model does differ from Gould's angular model because it restricts ϕ_t from "wrapping" around more than once. In other words, the model by Johnson and Wehrly doesn't allow more than one complete rotation of the phase on the "barber pole" as z_{1t} varies over time. Often times this is a reasonable assumption as in fMRI phase data.

However, this model does have drawbacks of its own including the need to know the distribution of z_{1t}. We often do not know the PDF or CDF of our predictor variables and tend to view them as fixed numbers that don't originate from a known distribution. Also, this model only allows for a single predictor z_{1t}. We may like to choose Z as a design matrix as describe in the previous section thus allowing us to include both a possible linear trend and reference function into our model.

Fisher and Lee (1992) generalize the Johnson and Wehrly model to allow for multiple predictor variables and relaxed the need for distributional assumptions on the design matrix [4]. Their model assumes the angular observations, $\phi_1, \ldots, \phi_n$, are independent and follow a Von Mises distribution with a constant concentration parameter, κ, that is the same across the series. In other words, $\phi_1, \ldots, \phi_n$ each originate from a Von Mises distribution with mean θ_t and concentration κ. Since they keep the concentration parameter fixed over time they model using only mean directions. They mention a model that fixes the mean and models the concentration but this is not appropriate for functional "phase" activation. The fixed concentration model is given by

$$\theta_t \quad = \quad \gamma_0 + g(u_t'\gamma) \tag{3.9}$$

where the design matrix U, with the t^{th} row $u_t' = (u_{1t}, \ldots, u_{(q+1)t})$, is comprised of all columns except the first in the design matrix Z, as defined in the previous section. Of course U can have any combination of columns the investigator chooses but we will focus on our choice. There is no need to include the baseline column in U because the intercept is already estimated within the model. The "link" function $g(\cdot)$ has the

purpose of mapping the result to the range of $-\pi$ to π. As stated by the authors this link function corrects the non-identifiability problem of MLEs that is present in the Gould model [4]. One possible link function that we will use which is given by Fisher and Lee is,

$$g(\cdot) \quad = \quad 2\tan^{-1}\left(\text{sgn}(\cdot)\,|\cdot|^{\lambda}\right) \tag{3.10}$$

where $\text{sgn}(\cdot)$ is the operator that returns the sign of its argument and the transformation parameter, λ, can be estimated from the data similar to the Box-Cox transformation [3, 4].

Fisher and Lee give us equations to obtain parameter estimates and make inferences using the mean model $\theta_t = \gamma_0 + g(u_t'\gamma)$ with the link function defined as $g(\cdot) = 2\tan^{-1}(\cdot)$. Of course the link function can be chosen differently but we will illustrate fitting the Fisher and Lee model with this particular choice. First, we define the natural log likelihood, denoted log, of the Von Mises distribution,

$$\log L = -n\log 2\pi \;\; - n\log I_0(\kappa) \;\; + \kappa\sum_{t=1}^{n}\cos(\phi_t - \gamma_0 - g(u_t'\gamma)). \tag{3.11}$$

Second, the authors define the following

$$\begin{aligned}
v_t &= \sin(\phi_t - \gamma_0 - g(u_t'\gamma)),\\
v &= (v_1,\ldots,v_n)',\\
U &= (u_1,\ldots,u_n)'\\
G &= \text{diag}(g'(u_1'\gamma),\ldots,g'(u_n'\gamma))\\
S &= \frac{1}{n}\sum_{t=1}^{n}\sin(\phi_t - g(u_t'\gamma))\\
C &= \frac{1}{n}\sum_{t=1}^{n}\cos(\phi_t - g(u_t'\gamma))\\
R &= (S^2 + C^2)^{1/2}.
\end{aligned}$$

In the above, variables have the following dimensions: v is an $n \times 1$ vector, G is a $n \times n$ matrix; while S, C, and R are scalars. The function, $g'(\cdot)$ is defined to be the derivative of the link function. Next, the MLEs are found by solving the following equations,

$$U'Gv \;\; = \;\; 0 \tag{3.12}$$

$$R\sin(\hat{\gamma}_0) \;\; = \;\; S, \tag{3.13}$$

$$R\cos(\hat{\gamma}_0) \;\; = \;\; C, \tag{3.14}$$

$$A(\hat{\kappa}) \;\; = \;\; R \tag{3.15}$$

where $A(\kappa) = I_1(\kappa)/I_0(\kappa)$. Fisher and Lee describe an iterative procedure for finding a solution to these equations. They also noted that centering the individual columns of U around their means will optimize the numeric calculations. We begin with an initial value for $\hat{\gamma}$ and calculate values for S, C, and R from the

above equations. An updated value of $\hat{\gamma}$, denoted by $\hat{\gamma}^*$ is then found by solving the following equation for $\hat{\gamma}^*$,

$$(U'G^2U)(\hat{\gamma}^* - \hat{\gamma}) \quad = \quad U'G^2y \tag{3.16}$$

where $y = (y_1, \ldots, y_n)'$ and $y_t = \frac{v_t}{[A(\hat{\kappa})g'(u_t'\hat{\gamma})]}$. We also find an updated estimate for γ_0 at each iteration by solving $\gamma_0 = \tan^{-1}(S/C)$ and κ from $A(\kappa)$. The updated estimate of $\hat{\gamma}^*$ is recursively placed back into formulas for a set number of iterations or until the iterative values differ by less than some pre-defined amount.

Fisher and Lee also give us a solution to find the large sample asymptotic variance of the estimated coefficient vector,

$$var(\hat{\gamma}) \quad = \quad \frac{1}{\hat{\kappa}A(\hat{\kappa})}\left\{\left(U'G^2U\right)^{-1} + \frac{\left(U'G^2U\right)^{-1}U'gg'U\left(U'G^2U\right)^{-1}}{\left(n - g'U\left(U'G^2U\right)^{-1}U'g\right)}\right\},$$

which will allow us to draw inferences on our γ's where g is a vector whose elements are the diagonal elements of G. They also describe the asymptotic variance for $\hat{\gamma}_0$ and $\hat{\kappa}$ to be equal to $1/(nA'(\hat{k}))$ and $[2(n-q)\hat{\kappa}A(\hat{\kappa})]^{\frac{1}{2}}$, respectively where $A'(\kappa)$ is the derivative of the ratio of Bessel functions with respect to $\hat{\kappa}$. The variance of the asymptotic normal limiting distribution of the Von Mises has variance $1/\kappa$ [6].

This now provides us with an estimate for the phase regression coefficients, γ, and their variances. We can then use a normal approximation to test the hypothesis, H_0: $\gamma_m = 0$ verses H_1: $\gamma_m \neq 0$ using the constructed test statistics, $\hat{\gamma}_m/\sqrt{var(\hat{\gamma}_m)}$. Alternatively one could set up linear contrast hypothesis tests and obtain the $-2\log$ of the ratio of the unrestricted likelihood over the restricted and use the asymptotic χ^2_w distribution to make inferences.

4 Simulation

We generated data to simulate activation in a voxel which is similar to that observed from a bilateral finger tapping fMRI block design experiment as described in the next Section [11]. The simulated time series consisted of $n = 256$ points where the true values for the data are known before random noise according to a pre-specified distribution is added.

Simulated fMRI data is constructed according to a general non-linear model as described by Rowe and Logan (2005) which for the magnitude consists of an intercept β_0; a time trend coefficient β_1; and a coefficient β_2 for a reference function, related to a block experimental design [12]. We also include regression coefficients on our phase change which consists of γ_0, γ_1, and γ_2. These specified γ's are also determinant of an intercept, trend, and a reference function. This allows the complex-valued data to have the following form,

$$y_t = [(\beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t})\cos(\theta_t) + \eta_{Rt}] + i\,[(\beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t})\sin(\theta_t) + \eta_{It}] \tag{4.1}$$

where

$$\theta_t = (\gamma_0 + \gamma_1 u_{1t} + \gamma_3 u_{2t}) \tag{4.2}$$

$t = 1, \ldots, n$, and $(\eta_{Rt}, \eta_{It}) \sim N(0, I_2\sigma^2)$. After creating the data we obtained the respective phase time series by taking the four quadrant arctangent of the imaginary component over the real component which we have previously shown to have the complicated distribution given in Equation 2.4 which is normal for large SNR.

For the current simulations we looked at two basic cases for the fMRI phase time series and fit both the standard OLS model along with the Fisher and Lee angular model to the data. We looked at the two fits graphically and examined the parameter estimates along with the variances of the parameter estimates for regression coefficients γ with each model. The two cases include a "control" case where there is no "wrap-around" present and a "test" case where the issue of "wrap-around" is present, both at equal SNR and magnitude parameters. We use the first time series to compare the two models when OLS could be used with little concern and verify the Fisher and Lee model is working properly. The latter phase time series will demonstrate to the reader that the OLS model poorly fits the data when the stringent conditions are not met. We will denote estimates of the phase regression coefficients under the unconstrained alternative hypothesis for the OLS model to be $\hat{\gamma}_{OLS}$ and those from the Fisher and Lee model with $\hat{\gamma}_{FL}$

The first time series has an SNR=10 with the true values being $(\beta_0, \beta_1, \beta_2) = (0.25, 0.008, 0.05)$, $(\gamma_0, \gamma_1, \gamma_2) = (0, 0.2, 8)$, and $\sigma^2 = 0.05$. We obtain estimates for the OLS model to be $\hat{\gamma}_{0_{OLS}} = -0.011$, $\hat{\gamma}_{1_{OLS}} = 0.175$, and $\hat{\gamma}_{2_{OLS}} = 10.805$. The estimates for the Fisher and Lee angular regression are $\hat{\gamma}_{0_{FL}} = -0.011$, $\hat{\gamma}_{1_{FL}} = 0.088$, and $\hat{\gamma}_{2_{FL}} = 5.396$. The respective variances are $var(\hat{\gamma}_{0_{OLS}}) = 0.0001$, $var(\hat{\gamma}_{1_{OLS}}) = 0.0004$, and $var(\hat{\gamma}_{2_{OLS}}) = 9.699$ for the OLS model while they are $var(\hat{\gamma}_{0_{FL}}) = 0.000072$, $var(\hat{\gamma}_{1_{FL}}) = 0.0001$, and $var(\hat{\gamma}_{2_{FL}}) = 2.411$ for the angular model. When no "wrap-around" is present, the Fisher and Lee model yields nonbaseline phase coefficients that are a factor of two smaller than those of the OLS model and a corresponding factor of four difference in coefficient variances although the model fits are identical. The OLS model estimate $\hat{\sigma}^2$ was equal to 0.037 while the Fisher and Lee model estimate of $1/\hat{\kappa}$ equaled 0.0363. In Figure 2(a) the true phase signal is shown with the black line, the observed time series (with noise) is plotted with cyan and the Fisher and Lee fitted model is plotted as red. Since that the OLS and Fisher and Lee models give us identical curves when no "wrap-around" is present we omitted the overlapping plot of the OLS curve. The decisions based on test the statistics for H_0: $\gamma_2 = 0$ vs H_1: $\gamma_2 \neq 0$ also agreed when using an asymptotic 5% two sided critical value of 1.96. The test statistic values were 3.469 for the OLS model and 3.476 for the Fisher and Lee model.

The second time series also has an SNR=10 with identical β and σ^2 values as the previous simulation but now with true values of $(\gamma_0, \gamma_1, \gamma_2) = (-1, 3.5, 40)$. We obtained the estimates for the OLS model to be $\hat{\gamma}_{0_{OLS}} = 0.351$, $\hat{\gamma}_{1_{OLS}} = 0.323$, and $\hat{\gamma}_{2_{OLS}} = 17.127$. The estimates for the Fisher and Lee angular regression are $\hat{\gamma}_{0_{FL}} = -1.012$, $\hat{\gamma}_{1_{FL}} = 2.951$, and $\hat{\gamma}_{2_{FL}} = 36.414$. The respective variances are $var(\hat{\gamma}_{0_{OLS}}) = 0.013$, $var(\hat{\gamma}_{1_{OLS}}) = 0.04$, and $var(\hat{\gamma}_{2_{OLS}}) = 874.597$ for the OLS model while they are $var(\hat{\gamma}_{0_{FL}}) = .00042$, $var(\hat{\gamma}_{1_{FL}}) = 0.011$, and $var(\hat{\gamma}_{2_{FL}}) = 53.471$ for the Fisher and Lee angular model. The OLS model estimate $\hat{\sigma}^2$ was equal to 0.2997 while the Fisher and Lee model estimate of $1/\hat{\kappa}$ equaled 0.1905. The decisions based

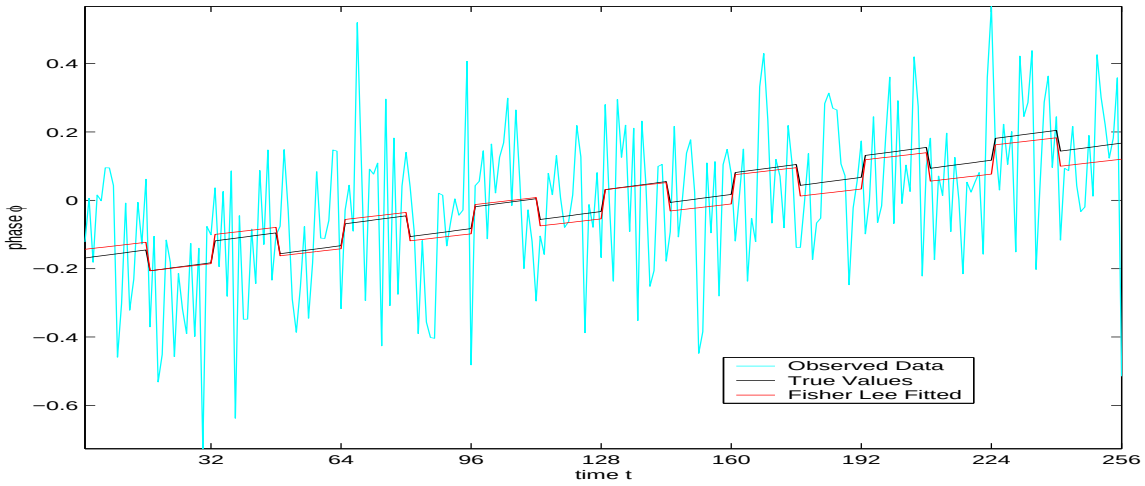
on the test statistics for H_0: $\gamma_2 = 0$ vs H_1: $\gamma_2 \neq 0$ do not agree when using an asymptotic two sided 5 % critical value of 1.96. The test statistic values were 0.579 for the OLS model and 4.9792 for the Fisher and Lee model.

By looking at Figure 2(b) we can see that reference function or third column of U has a strong correlation with the phase time series values and we would agree with the Fisher and Lee decision that it is statistically significant. In Figure 2(b) the true noiseless phase signal is shown with the black line, the simulated observed time series (with noise) is plotted with cyan, the OLS fitted model is signified with blue, and the Fisher and Lee fitted model is plotted as red. We see that the OLS and Fisher and Lee models no longer give us consistent fits. The OLS model doesn't fit the data well and is situated between most of the observations. The estimated values for the γ's were much closer to the true values for the Fisher and Lee model compared with the OLS model. Also, the variances for the Fisher and Lee estimates are much smaller especially for $\hat{\gamma}_{2_{FL}}$ which is the coefficient of primary interest. Further extending our exploration of the γ_2 estimates for both models we generated 10,000 random data sets with the same parameter values and compared the estimates for γ_2 using both models. The sample mean for all 10,000 estimates using the Fisher and Lee model along with the OLS model were 32.41 and 9.64, respectively. The sample variance for these 10,000 estimates of γ_2 were 13.28 and 209.22 for the Fisher and Lee along with the OLS model, respectively. This suggests that the Fisher and Lee angular regression model is an improved alternative to OLS when modeling fMRI phase time series because in has the potential for more accurate and precise estimates when dealing with a reference function coefficient.

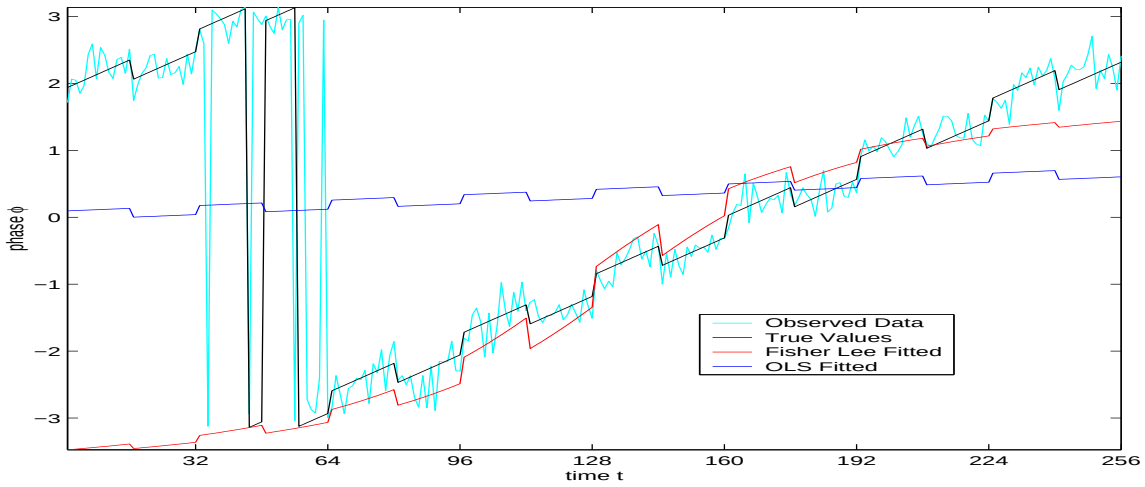
5 Real Data Example

We now will compare the two models using actual data acquired during a sequential bilateral finger tapping fMRI block design experiment. There are a total of 4096 voxel time series comprised of $n = 256$ observations per voxel. These 256 observations consists of alternating sets of 16 observations on and off. An axial slice through the motor cortex is examined here. Scanning used a 1.5T GE Signa scanner, where 5 axial slices of 64×64 where acquired. Voxels were 3.125×3.125mm in plane and 5mm thick with TR=1000ms, TE=47ms, and filtering was applied to remove low frequency and respiration noise.

Before examining the phase-only data, we looked at the magnitude-only data. An OLS model to the magnitude-only time series in each voxel with design matrix $X = Z$ as previously described and magnitude-only regression coefficients $\beta = (\beta_0, \beta_1, \beta_2)'$. We computed activation t-statistics in each voxel testing the hypothesis of the coefficient corresponding to the reference function in the last column of X being zero. The bilateral activation in the motor cortex regions for the magnitude-only OLS model can be seen in Figure 3 along with the activation along the midline in the supplemental motor area. In Figure 3(a) are unthresholded t-statistics while Figure 3(b) contains t-statistics that were Bonferroni thresholded Bonferroni corrected for multiple comparisons as described in Logan and Rowe (2004) with 5% family wise error (FWE) rate we

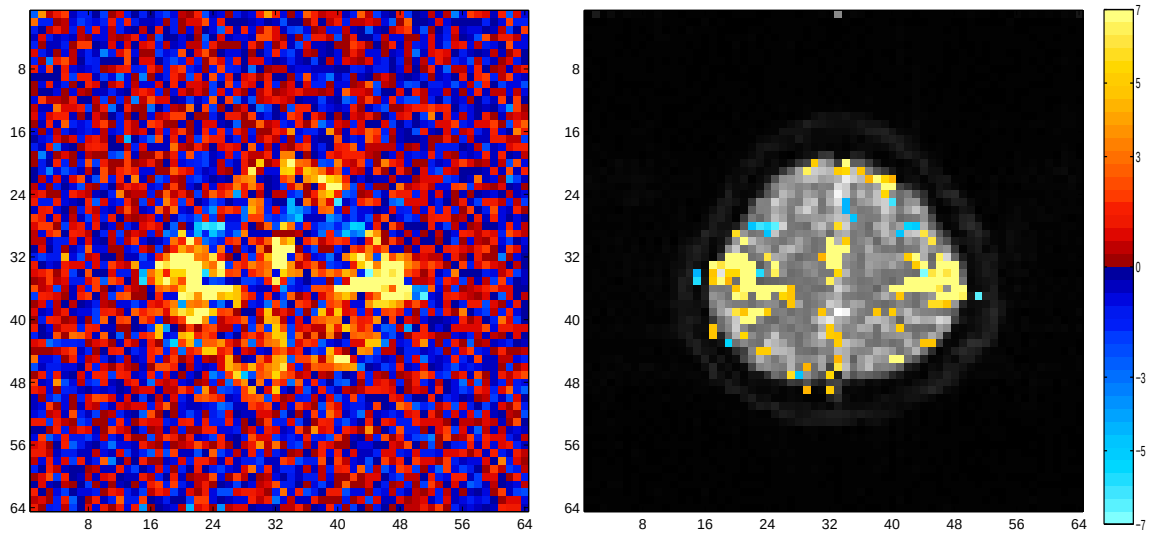


(a) OLS Fitted Line



(b) Both OLS and Fisher and Lee Fitted Line

Figure 2: Part a: True signal (Black), observed signal (Cyan), and Fisher and Lee fitted line (Red). Part b: True signal (Black), observed signal (Cyan), OLS fitted line (Blue), and Fisher and Lee fitted line (Red)



(a) (b)

Figure 3: Part a shows the unthresholded OLS test statistics for H_0: $\beta_2 = 0$ using the magnitude-only data. Part b shows the 5% Bonferroni FWE thresholded statistics

present the activation maps in Figures 5(c) and 5(d) [8].

If the phase-only data contains no information regarding possible biological phenomena in the brain we would anticipate seemingly random activations above the threshold. If the phase-only data is to contain information regarding possible biological phenomena in the brain we would anticipate seeing "phase" activations with some sort of pattern. One possible pattern is to be in similar locations as the "magnitude-only" activations. The association between magnitude-only and phase-only time series observed by Menon for areas with large blood vessels would suggest "phase" activations can be found in similar places as "magnitude" activations given such blood vessels are present [9]. Any similarities between the statistics for the magnitude-only and phase-only models would strengthen the idea that valuable information is discarded when magnitude-only data is analyzed solely.

We fit each time series with both an OLS regression and a Fisher and Lee angular regression then compared the results. Due to the erratic nature of phase data outside of the brain we observed an increased number of statistically significant "phase" activations in this area with the Fisher and Lee model. These are observed with the Fisher and Lee model and not with the OLS model because the Fisher and Lee model is able to fit a more accurate model because of the lower SNR restriction. These seemingly random "phase" activations outside the brain are obviously not due to any biological phenomena but may be solely the result of a more robust model fit.

An issue which arises when implementing the Fisher and Lee model with our given link function is convergence of the γ estimates for a small number of voxel time series. When a large concentration of the

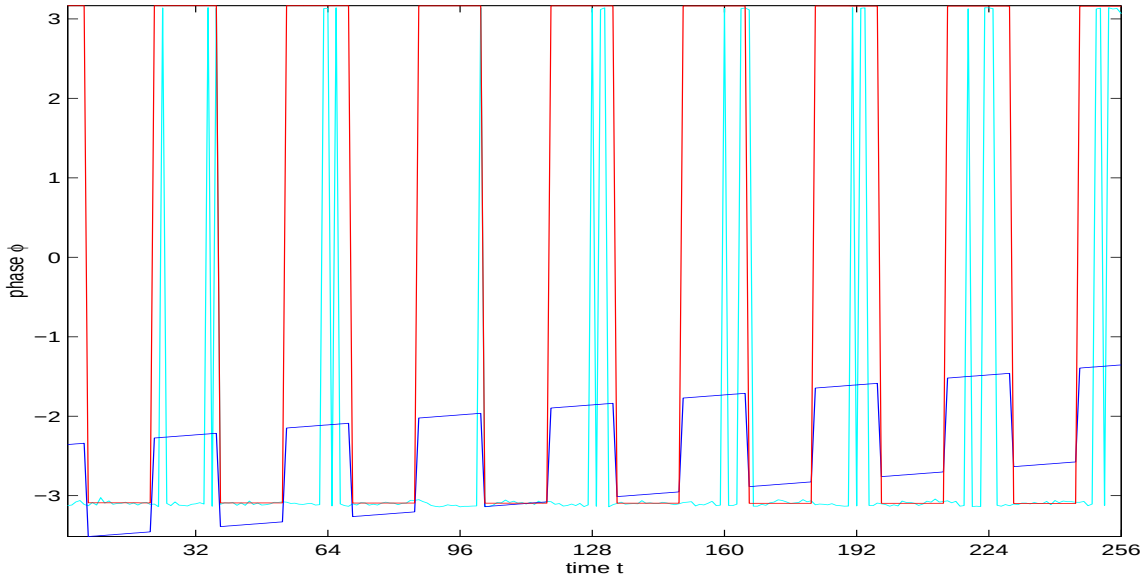


Figure 4: Actual acquired fMRI phase time series (Cyan), OLS fitted values (Blue), along with Fisher and Lee fitted values (Red).

points in a given series are near $\pm\pi/2$, the limits of our arctangent function, the Fisher and Lee model does not always to converge. However, this issue is easily corrected by subtracting the angular mean from the entire series for those particular voxels before fitting with the Fisher and Lee method. The investigator is then able to obtain estimates and make inferences using this improved model. An example which shows the robustness of the Fisher and Lee model can be seen in Figure 4 where the observed phase time series from an actual experimental voxel is shown in cyan while the fitted series for the OLS along with the Fisher and Lee models are shown with blue and red, respectively. It is easily seen that the Fisher and Lee angular regression accurately models the large "spikes" in the data during the finger tapping task and stays within the boundaries of $-\pi$ and π. Yet, the OLS fit doesn't accurately model the correct intensity of the "spikes" and actually has fitted values below the range of $-\pi$. It should be noted that this figure is truly a three dimensional cylinder instead of a two dimensional plot where the "spikes" are observed jittering around the connected $\pm\pi$ boundary. In other words, the plot takes the long way around the cylinder to connect points. The model by Fisher and Lee correctly estimates a very small amplitude coefficient for the reference function γ_2 while the OLS model estimages a very large one.

Since we are interested in possible associations between the phase-only data and a pre-defined reference function we constructed large sample asymptotic z-values to make inference about γ_2. We monitored these statistics for all 4096 voxels in both the OLS and the Fisher and Lee models. In Figures 3(a) and 3(b) we show the unthresholded statistics map for both OLS and the Fisher and Lee models. We see the statistics within the brain appear to agree almost identically because the time series don't exhibit large variation similar to the region outside the brain. It should be noted that some of the phase time series on the border

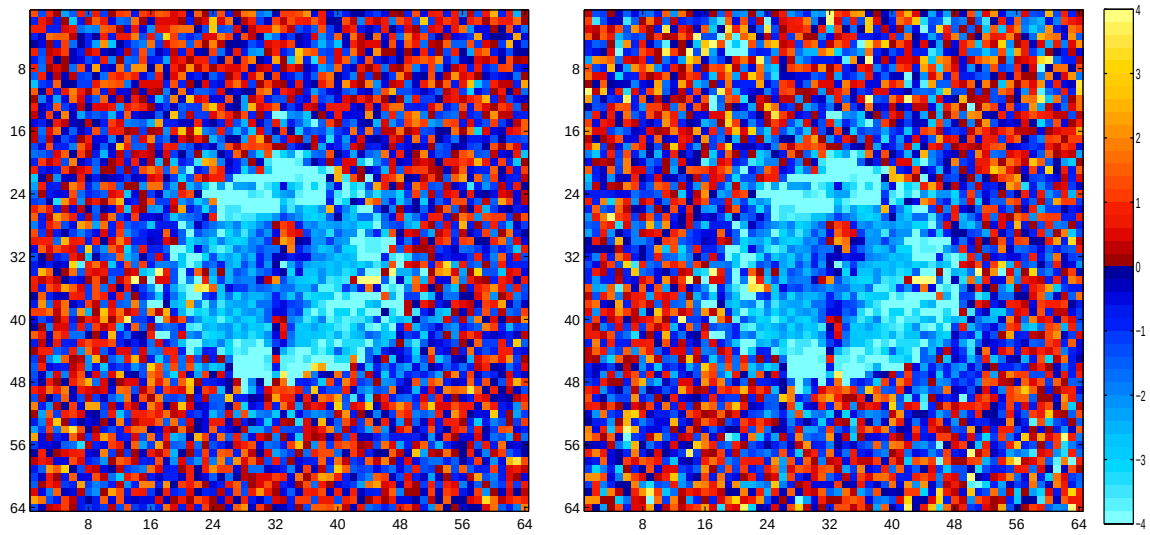
of the brain had angular means near the $\pm\pi/2$ value and needed to be centered to obtain convergence. After applying a threshold that was Bonferroni corrected for multiple comparisons as described in Logan and Rowe (2004) with 5% Bonferroni FWE rate we present the activation maps in Figures 5(c) and 5(d) [8]. Again we see similar conclusions to the OLS model with the exception of additional "phase" activations outside the brain. Again, we attribute this to the robustness of the Fisher and Lee model when applied to very noisy data. Since the Fisher and Lee model more accurately characterizes troublesome data we would expect more statistical significance to be found for the reference function coefficient, γ_2. Troublesome voxels can be seen in the brain as well as seen by Figure 4 which is the "yellow" voxel on the right side of the brain in Figure 5(c) that is declared active with the OLS model but not whith the Fisher and Lee model in Figure 5(d). The focal positive "magnitude" activations showed an association with the "phase" activations which would imply that there is information in the phase-only data.

6 Conclusion

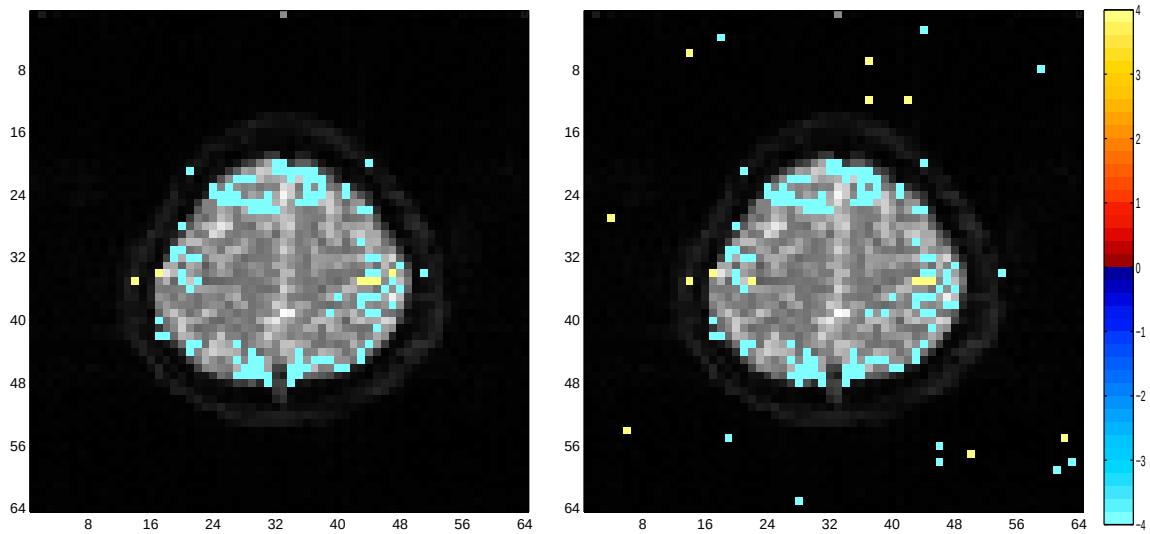
Modeling fMRI phase time series with OLS regression results in some troublesome phenomena which include poor fit, incorrect parameter estimation, and inaccurate test statistics. Most of these problem arise from the issue of "wrap-around" in the time series. We discussed three linear-circular regression techniques and concluded that the model proposed by Fisher and Lee (1992) was the best choice for the problem at hand. It allowed us to define a design matrix which could account for several regressors including a linear trend and a reference function from which to make and test hypothesis using a large sample asymptotic z statistic. We were able to implement a numeric algorithm given by Fisher and Lee to obtain our coefficient estimates along with the variances of the coefficient estimate and thus test hypotheses.

Simulations were performed for both a "control" case with no "wrap-around" and a "test" case with "wrap-around" present. In the case of no "wrap-around" we found that the fitted line for the Fisher and Lee model was nearly identical to that of the OLS model. The test statistics also were very similar and the conclusion for each agreed. This would imply the models are equivalent under these conditions. In the generated time series where "wrap-around" was present we noted that the estimates for the Fisher and Lee model were much closer to the true values compared with the OLS model. Also, the test statistics for the Fisher and Lee model differed from those for the OLS model and led to different conclusions with respect to γ_2 for the case with "wrap-around." We also showed during our second set of simulations that the γ_2 estimates were more accurate and precise when we compared the Fisher and Lee model to the OLS model for the 10,000 simulated phase-only data sets created with the same parameters. This should give the reader confidence in implementing this model when the data contains "wrap-around."

The actual phase time series we presented solidified our findings that the Fisher and Lee model is an excellent choice for fitting with fMRI phase-only data. We showed a specific voxel example where OLS was unable to match the modeling accuracy of the Fisher and Lee model and actually contained fitted values



(a) (b)



(c) (d)

Figure 5: Parts a and b Shows the unthresholded OLS and Fisher and Lee test statistics for H_0: $\gamma_2 = 0$ using the phase-only data. Parts c and d Shows the thresholded test statistics for both OLS and Fisher and Lee models using the same hypothesis.

not consistent with the fMRI phase angular property. Fisher and Lee's model is less suseptible to low SNR problems and if applied to other real data examples with more noise we would expect it to perform at as well as OLS. In real fMRI data that is of a higher resolution with smaller voxels, we expect there to be a larger difference between the two models. This newly implemented angular model does detect temporal correlations between the phase time course and a reference function in many of the same places the OLS model does. We suggest using the Fisher and Lee angular model as both a comparative tool along with a exploratory method for fMRI phase data. It may be more beneficial for the investigator to implement this new angular method on a voxel-wise basis where the fitted values can be compared to both the observed data and the OLS fit allowing the appropriate modeling decision to be made. Finally, as we have shown with simulated data the Fisher and Lee model has less assumptions that need to be met and often times describes the data better compared with OLS regression.

A Conditional Distribution

We will show and verify that the conditional distribution of the phase is a general Von Mises (Circular Normal) distribution. To derive the conditional distribution we need both the joint distribution, $f(\phi, r)$ and the marginal, $f(r)$ which were derived by Rowe and Logan (2004).

$$f(\phi, r) = (2\pi\sigma^2)^{-1} r e^{-\frac{r^2+\rho^2}{2\sigma^2}} e^{-\frac{r\rho\cos(\phi-\theta)}{\sigma^2}} \tag{A.3}$$

We set $\sigma^2 = \frac{1}{\kappa}$ to obtain a more standard notation and get,

$$f(\phi, r) = \left(\frac{2\pi}{\kappa}\right)^{-1} r e^{-\frac{(r^2+\rho^2)\kappa}{2}} e^{r\rho\kappa\cos(\phi-\theta)}. \tag{A.4}$$

By definition of conditional distributions we need to divide by the marginal distribution of r

$$\begin{aligned} f(\phi|r) &= \frac{\frac{1}{2\pi} e^{-\frac{\kappa}{2}(r^2-2\rho r\cos(\phi-\theta)+\rho^2)}}{r e^{-\frac{\kappa}{2}(r^2+\rho^2)} I_0(r\rho\kappa)} \\ &= \frac{e^{\kappa^*\cos(\phi-\theta)}}{2\pi I_0(\kappa^*)}. \end{aligned} \tag{A.5}$$

where $\kappa^* = r\rho\kappa$. This derived conditional distribution known distribution of a Von Mises (Circular Normal) with mean θ and concentration of $r\rho\kappa$ [6].

References

[1] P.A. Bandettini, A. Jesmanowicz, E.C. Wong, and J.S. Hyde. Processing strategies for time-course data sets in functional MRI of the human brain. *Magn. Reson. Med.*, 30(2):161–173, 1993.

[2] J. Borduka, A. Jesmanowicz, J.S. Hyde, H. Xu, Estkowski L., and S.-J. Li. Current-induced magnetic resonance phase imaging. *J. of Magnetic Resonance*, 137:265–271, 1999.

[3] G.E.P. Box and D.R. Cox. An analysis of transformations. *Journal of the Royal Statistical Society*, 26:296–311, 1964.

[4] N.I. Fisher and A.J. Lee. Regression models for an angular response. *Biometrics*, 48:665–677, 1992.

[5] A.L. Gould. A regression technique for angular variates. *Biometrics*, 25:683–700, 1969.

[6] S.R. Jammalamadaka and A. SenGupta. *Topics in Circular Statistics*. World Scientific Publishing Co., Singapore, 2001.

[7] R.A. Johnson and T.E. Wehrly. Some angular-linear distributions and related regression models. *Journal of the American Statistical Association*, 73(353):602–606, 1978.

[8] B.R. Logan and D.B. Rowe. An evaluation of thresholding techniques in fMRI analysis. *NeuroImage*, 22(1):95–108, 2004.

[9] R.S. Menon. Postacquisition suppression of large-vessel BOLD signals in high-resolution fMRI. *Magn. Reson. Med.*, 47(1):1–9, 2002.

[10] N. Ravishanker and D.K. Dey. *A First Course in Linear Model Theory*. Chapman & Hall/CRC Press, Boca Raton, FL, USA, 2002.

[11] D.B. Rowe and B.R. Logan. A complex way to compute fMRI activation. *NeuroImage*, 23(3):1078–1092, 2004.

[12] D.B. Rowe and B.R. Logan. Complex fMRI analysis with unrestricted phase is equivalent to a magnitude-only model. *NeuroImage*, 24(2), 2005.

[13] A.M. Smith, B.K. Lewis, U.E. Ruttimann, F.Q. Ye, T.M. Sinnwell, Y. Yang, J.H. Duyn, and J.A. Frank. Investigation of low frequency drift in fMRI signal. *NeuroImage*, 9(5):526–533, 1999.

[14] T.E. Wehrly and R.A. Johnson. Bivariate models for dependence of angular observations and a related markov process. *Biometrika*, 66:255–256, 1979.