The Annals of Applied Statistics
0, Vol. 0, No. 00, 1–20
https://doi.org/10.1214/24-AOAS1981

RICE-DISTRIBUTED AUTOREGRESSIVE TIME SERIES MODELING OF MAGNITUDE FUNCTIONAL MRI DATA

BY DANIEL W. ADRIAN[1,a], RAJAN MAITRA[2,b] AND DANIEL B. ROWE[3,c]

[1]*Department of Statistics, Grand Valley State University,* [a]*adriad1@gvsu.edu*
[2]*Department of Statistics, Iowa State University,* [b]*maitra@iastate.edu*
[3]*Department of Mathematical and Statistical Sciences, Marquette University,* [c]*daniel.rowe@marquette.edu*

Functional magnetic resonance imaging (fMRI) data generally consist of time series image volumes of the magnitude of complex-valued observations at each voxel. However, incorporating Gaussian-based time series models and the Rice distribution—a more accurate model for the data—in the time series have been separated by a distributional "mismatch." We bridge this gap by including pth-order autoregressive (AR) errors into the Gaussian model for the latent real and imaginary components underlying the Rice-distributed magnitude data. Parameter estimation is then done by augmenting the observed magnitude data with the missing phase data in an expectation-maximization (EM) algorithm framework and followed by AR order determination and computation of test statistics for activation detection. Using simulated and experimental low-SNR fMRI data, we compare the performance of this Ricean time series model with a Gaussian AR(p) model for the magnitude data and also with a complex Gaussian time series model for the entire complex-valued data. Our results show improved parameter estimation and activation detection under the Ricean AR(p) model for the magnitude data than its Gaussian counterpart. The model using the complex-valued data (which is rarely collected in practice) detects activation better than both magnitude-only models but only because it has more data. Thus, while our results here provide for the improved analysis of commonly-collected and archived magnitude-only fMRI datasets, they also argue strongly against the currently routine practice of discarding the phase of the complex-valued fMRI time series, advocating instead for their inclusion in the analysis.

1. Introduction.

Functional magnetic resonance imaging (fMRI) is a prominent noninvasive modality for studying human brain function. It is built upon the principle of the blood oxygen level dependent (BOLD) contrast (Ogawa et al. (1990), Belliveau et al. (1991), Kwong et al. (1992), Bandettini et al. (1993)), where firing neurons lead to changes in the blood oxygen levels of neighboring vessels, and the magnetic resonance (MR) signal fluctuates due to the differing magnetic susceptibilities of oxygenated and deoxygenated hemoglobin (Lazar (2008)). Scientists can gain insight on the functional structures of the brain by analyzing time courses of MR signals acquired while a subject performs a designed series of tasks.

The voxelwise MR signal at each time point is originally complex-valued, containing real and imaginary (or equivalently, magnitude and phase) components. This complex-valued attribute is a consequence of how the data are acquired: the originally measured, complex-valued k-space data (Brown, Kincaid and Ugurbil (1982), Tweig (1983), Ljunggren (1983)) consist of the different frequency contributions to the signal from each voxel resulting from magnetic field gradients (Jezzard and Clare (2001)). Then the application of the inverse Fourier transform (Jain (1989)), a complex-valued operation on the k-space data, separates

---

Received June 2024; revised October 2024.
Key words and phrases. Bilateral finger-tapping motor experiment, EM algorithm, empirical information matrix, hemodynamic response function, Monte Carlo integration, Rice distribution, signal-to-noise ratio.

these frequencies and localizes each voxel's measurements. However, despite the fact that the original signal is complex-valued, statistical analysis of fMRI data is almost always necessarily only on the magnitude data, because the concomitant phase measurements are discarded and irretrievable once the (magnitude) data are extracted from the scanner (where it is overwritten by subsequent scans). We refer to such analyses as "magnitude-only" (MO) statistical analyses and note that this approach likely arises as a consequence of the default output of MR scanners that does not routinely include phase images, even though they can easily be collected by simply changing a preset variable in an input file (Yu et al. (2018)). Consequently, most fMRI data and analyses are MO and, at least figuratively, do not use half of the originally available data.

One of the most common forms of MO analysis fits, at each voxel, a general linear model (Friston et al. (1995)) for the (preprocessed) time series observations in terms of a waveform representing the expected BOLD contrast. This waveform is the convolution of the stimulus time course with the hemodynamic response function (HRF), which gives the BOLD response to an instantaneous neuronal activation (Friston, Jezzard and Turner (1994), Glover (1999)). These general linear models for magnitude fMRI time series also incorporate autoregressive (AR) (Bullmore et al. (1996), Marchini and Ripley (2000), den Dekker et al. (2009)) or autoregressive moving average (ARMA) (Locascio et al. (1997)) errors, due to several reasons. For one, the hemodynamic response to a single neural activation takes between 15 and 20 seconds (Lazar (2008)), which is much longer than the sampling intervals of many fMRI techniques—for instance, of between 100 milliseconds and five seconds for echoplanar imaging (EPI) techniques (Friston, Jezzard and Turner (1994)). Additional sources of autocorrelation are also provided by the subject's cardiac and respiratory cycles (Friston et al. (2000)) and by the common preprocessing step of temporal smoothing. From these model fits, the time series at each voxel is aggregated to a test statistic that measures the degree of activation in the statistical parametric mapping (SPM) framework of Friston et al. (1990). Thresholding methods are then applied to the SPM to identify activated voxels (Genovese, Lazar and Nichols (2002), Logan and Rowe (2004), Worsley et al. (1996)).

The above MO approaches assume that the magnitude measurements follow a Gaussian distribution, but a more correct model utilizes the Rice distribution (Gudbjartsson and Patz (1995), Rice (1944)), as we show in the following. It is well known (Henkelman (1985), Macovski (1996), Sijbers et al. (1998)) that the complex k-space data are Gaussian distributed, and this distributional assumption is also preserved, by linearity, upon applying an inverse Fourier transform. Specifically, it is commonly assumed (Wang and Lei (1994)) that the real and imaginary measurements at a single point in space and time are independent normal random variables with the same variance and phase-coupled means; it follows that the magnitude has a Rice distribution. The Gaussian MO model is often justified by the fact that the Rice distribution approaches the Gaussian distribution for large signal-to-noise ratios (SNRs).

However, fMRI scans that are more detailed, acquired faster or on portable systems (Liu et al. (2021)) come at the loss of SNR. For magnitude fMRI time series, the SNR represents the ratio of the mean, that is, the nonactivation-related, baseline signal to the standard deviation (SD) of the noise time series.[1, 2] It is well known that SNR is proportional to voxel volume and inversely proportional to image acquisition time (Lazar (2008)). Thus, although scans with increased spatiotemporal resolution show promise in a clinical setting in

---

[1] There is also the contrast-to-noise ratio, or CNR, that is the ratio of the amplitude of the BOLD contrast to the noise SD.

[2] It should be noted that we are speaking of temporal SNR—comparing measurements across time at a single voxel—and not image SNR, which would compare measurements across voxels at a single time point; see Murphy, Bodurka and Bandettini (2007).

presurgical mapping and also in understanding human neuropsychology (Rosen and Savoy (2012)), they also come at the cost of decreased SNR. SNR is also proportional to magnetic field strength, so ultrahigh field (UHF) MRI (Balchandani and Naidich (2015)) makes high-resolution fMRI more feasible, but UHF scanners presently lack wide availability (Cosottini and Roccatagliata (2021)). There has also lately been the move to make MRI (and fMRI) more accessible through the use of ultralow field magnets (Liu et al. (2021)), which while providing more accessible, cost-effective and environmentally sustainable scanning equipment, however, yield images of relatively poor quality, lower SNR and limited spatial resolution (Islam et al. (2023)).

Out of the above concern that the Gaussian assumption may not be adequate for such low-SNR data, Zhu et al. (2009) developed Rice-distributed models that ignored temporal dependence in the voxelwise time series, with Bayesian methodology for this problem also developed in Wegmann, Eklund and Villani (2017). Also, Solo and Noh (2007) demonstrated that Gaussian-model-based maximum likelihood (ML) estimates of parameters for simulated Ricean data are biased for SNRs under 5, with the bias increasing as the SNR decreases. But a Ricean model for the observations that also incorporates temporal dependence has never previously been developed, and we address this lacuna in this paper.

1.1. *Main contributions of this paper.* We highlight the main contributions to the statistics and scientific community that we make through this paper. First, we provide methodology and analyze the use of autoregressive time series regression for data that are from the Rice distribution. In doing so, we address the gaps in Zhu et al. (2009) or in Wegmann, Eklund and Villani (2017) that ignored temporal structure while developing Rice-based regression models for the fMRI time series. Adrian, Maitra and Rowe (2018) showed substantial gains in using the complex-valued data over those using (Gaussian-distribution-assumed) MO-only data; however, all archived or acquired datasets from fMRI studies are MO, and, therefore, it is important to see if the analysis of such datasets can be improved by using more accurate Rice-distributed AR time series models, especially in low-SNR situations.

Second, we provide a thorough and detailed comparison of the CV-based analysis (Adrian, Maitra and Rowe (2018)) of fMRI data vis-a-vis that obtained using the methods developed in this paper. We show that while our Rice-distributed modeling of MO datasets improves parameter estimation and activation detection accuracy, over those done by Gaussian-based MO analysis, with the improvement very pronounced at low SNRs, both analyses approaches are outperformed by CV-based analysis. Therefore, we advocate to the fMRI and scientific community for the storage (and analysis) of CV (both magnitude and phase) fMRI data, especially because, as mentioned earlier, its acquisition is a simple matter of modifying a preset variable in an input file (Yu et al. (2018)). At the same time, it is important to reiterate, as in the previous paragraph, that archived datasets or those gathered under current practice, only have the magnitude values, with the phase having been discarded and, therefore, unretrievable, and so need the methods developed in this paper for their improved MO-based analysis.

Third, as will be discussed in Section 2, we provide to the community not only a complex-valued fMRI dataset of a finger-tapping experiment but also one that, with the innovative use of a radio-frequency (RF) coil, provides us with a low-SNR dataset of a well-studied experiment and allows for benchmarking of performance of our more accurate MO-based methods in lower-SNR frameworks. Ultralow field MRI is increasingly gaining popularity, however these systems are mostly unavailable publicly, and our framework fills this gap, providing a CV dataset with both magnitude and phase measurements.

Fourth, but also significant, our immediate application in this paper is in the context of analyzing fMRI time series, but similar problems also arise in the context of signal processing and communications (Abdi et al. (2001), Hajri, Youssef and Patzold (2009), Lindsey (1964), Bar-Shalom, Li and Kirubarajan (2001), Arafa and Messier (2010)) where it may be desirable to understand the characteristics and behavior of wireless, radar and sonar signals in relation to other factors. Another context is in meteorology (Best, Rayner and Thas (2010), Baïle, Muzy and Poggi (2011), Baggio and Muzy (2024)), where first the Rayleigh and then the Rice and the M-Rice distributions have lately been used to better model the windspeed and improving regression modeling of Rice-based time series data are important, for instance, for predicting windspeed and wind energy energy output (Wang et al. (2019)).

1.2. *Organization of this paper.* The remainder of the paper is structured as follows. Section 2 introduces a series of low-SNR images acquired from a finger-tapping experiment that is our motivating application. These images have a lower SNR than typical fMRI data, due to their being acquired with the body coil instead of the head coil, and their analysis here is motivated by the divergence of the Rice and Gaussian distributions at low SNRs. Section 3 illustrates the novel methodological development of the MO AR(p) Ricean model. Section 4 performs simulation experiments to study the validity of the methodological results and compares the results using different models under a known ground truth. We compare three models: the AR(p) Ricean and Gaussian models for MO voxel time series as well as a model that utilizes the entire complex-valued (CV) data. Our primary focus is on the two models for MO data, as the entire CV data is rarely collected, but we also present comparisons between the CV vs. MO data-based models to advocate for collecting the complete data. Section 5 presents the statistical analysis of the low-SNR dataset and its implications. We conclude with a discussion of the results and the paper. Our paper also has an online supplement (Adrian, Maitra and Rowe (2025)) containing additional details regarding methodology, the simulation experiments, and the dataset analysis. Sections, figures and equations in the Supplementary Material are referenced here with the prefix “S-.”

2. A low-SNR fMRI finger-tapping experiment dataset.

2.1. *Data acquisition.* We develop our methods in the backdrop of a sequential finger-tapping experiment, the type of which are applied to noninvasive neurosurgical preparation (Lee et al. (1999)). Experiments that use finger tapping, sponge squeezing or brushing of the palms as stimuli in block design are used to identify the location of hand function in candidate patients for resective surgical treatment for tumors and epilepsy (Lee, Jack and Riederer (1998)). While it is well known that the central sulcus in the sensorimotor cortex is the location of hand function for normal healthy adults (Rumeau et al. (1994)), fMRI allows the location of specific functional areas to be verified noninvasively in relation to the proposed surgical target (Lee et al. (1999)).

In our application, MR images were acquired with TR $= 1$ s during a block design experiment with an initial 16 s of rest followed by 19 epochs of 16 s of right-hand finger tapping alternating with 16 s of rest. Following standard practice, the first three images are excluded from our analysis, due to machine “warmup” effects, leaving us with a temporal sequence of $n = 621$ images. Each volume image was composed of seven 2.5 mm thick 128×128 axial slices with a 24.0 cm FOV. Due to the decussation of nerve fibers in the upper slices of the brain, crossing from one lateral side to the other (Carpenter (1991)), the right-hand finger tapping activates the left central sulcus (identified in Figure 1).

The data were acquired with the body coil (also known as the RF coil, as it produces the radiofrequency pulse used to produce the “resonance” condition of MRI). In general, the body coil has a large measurement field and thus has lower SNR than specialized coils such as the head coil usually used in fMRI studies. Health care sites in developing countries

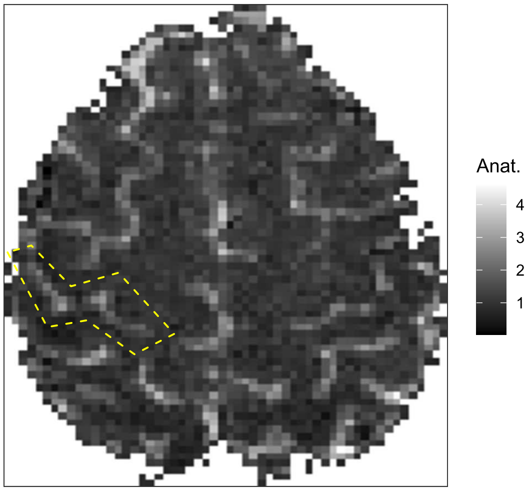

FIG. 1. *Anatomical image of the second slice identifying the left central sulcus.*

may use body coil acquisition because specialized coils represent additional costs beyond the MR scanner. Additionally, spinal (Powers, Ioachim and Stroman (2018)) and renal fMRI (Zhang et al. (2013)) use body coil acquisition. Using body (rather than head) coil acquisition provides us with a lower SNR dataset of a well-studied experiment to serve as a marker for performance in low-SNR settings, which is where CV, MO Ricean and MO Gaussian models have been shown to diverge previously. Further, as mentioned in Section 1.1, it is extremely rare for such a dataset to be publicly available, so we make it available to the scientific community at https://github.com/dadrian14/arp-rice-data.

2.2. *Data processing pipeline.* For this dataset the phase components of the time series images were not discarded but stored along with the magnitude images used in traditional fMRI analysis. The data processing flow included Nyquist ghost removal and correction for global zero-order off-resonance using three navigator echos (Jesmanowicz, Wong and Hyde (1993), Nencka, Hahn and Rowe (2008)), image reconstruction from k-space by inverse Fourier transform (Kumar, Welti and Ernst (1975), Rowe (2016)) and estimation and correction of the dynamic field using temporal off-resonance alignment of single-echo time-series (TOAST) (Hahn, Nencka and Rowe (2009, 2012)). A binary mask of voxels above 12% of the maximum voxel signal magnitude was generated from the first magnitude image of the dataset (before discarding the first three images) to represent the voxels within the brain subject to statistical activation detection. In addition, we used smoothing splines to detrend the voxel time series, after comparing several methods (see Section S-1.1).

After preprocessing we applied the CV and MO models to each voxel time series. For each model the design matrix X had $n = 621$ rows and $q = 2$ columns: one column was an intercept modeling the baseline MR signal, and the other was a zero-centered waveform modeling the expected BOLD response given by a convolution of the stimulus time course with the Glover (1999) hemodynamic response function. The bottom panel of Figure S-1 shows a superposition of the block design stimulus time course with this expected BOLD response waveform. The next section presents these CV and MO time series models in detail.

3. Methodological development.

We compare three models for fMRI time series: the AR(p) Ricean and Gaussian models for MO voxel time series as well as a model that utilizes the entire complex-valued (CV) data. However, the AR(p) Ricean model and its estimation needs development, so we use this section to do so.

3.1. *Statistical models for CV and MO time series.* We first introduce notation, focusing on a single voxel (and suppressing voxel-related subscripts). The CV measurement at time t can be denoted in real/imaginary form by $y_{Rt} + i y_{It}$ or in magnitude/phase form by $r_t \exp(i\phi_t) = r_t(\cos\phi_t + i \sin\phi_t)$. Trigonometric identities in the complex plane hold that $y_{Rt} = r_t \cos\phi_t$, $y_{It} = r_t \sin\phi_t$, $r_t = (y_{Rt}^2 + y_{It}^2)^{\frac{1}{2}}$, and $\phi_t = \arctan_4(y_{It}, y_{Rt})$, the four-quadrant arctangent (see Glisson ((2011), p. 348)) corresponding to $\arctan(y_{It}/y_{Rt})$. We denote the real, imaginary, magnitude and phase time series vectors by $\boldsymbol{y}_R = (y_{R1}, \dots, y_{Rn})'$, $\boldsymbol{y}_I = (y_{I1}, \dots, y_{In})'$, $\boldsymbol{r} = (r_1, \dots, r_n)'$ and $\boldsymbol{\phi} = (\phi_1, \dots, \phi_n)'$, with n denoting the number of MR scans. The Rowe and Logan (2004) model states that

$$\begin{pmatrix} \boldsymbol{y}_R \\ \boldsymbol{y}_I \end{pmatrix} = \begin{pmatrix} \boldsymbol{X} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{X} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta} \cos\theta \\ \boldsymbol{\beta} \sin\theta \end{pmatrix} + \begin{pmatrix} \boldsymbol{\eta}_R \\ \boldsymbol{\eta}_I \end{pmatrix}, \tag{1}$$

where the expected magnitude response $\boldsymbol{X}\boldsymbol{\beta}$ is coupled with the constant phase location parameter θ. The columns of $\boldsymbol{X}$ represent various components of the magnitude signal including the baseline level and the expected BOLD contrast. The errors $\boldsymbol{\eta} = (\boldsymbol{\eta}_R', \boldsymbol{\eta}_I')' \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma} \otimes \boldsymbol{\Phi})$, where $\boldsymbol{\Sigma}$ and $\boldsymbol{\Phi}$ are matrices of order 2 and n, specifying the real/imaginary and temporal covariances (the latter with an AR(p) structure), and the direct (Kronecker) product $\otimes$ implies separability of these covariances. Based on this framework, we present three models:[3]

1. *Complex-valued AR(p) (CV) model*: This model takes the form of (1) with an AR(p) structure for $\boldsymbol{\Phi}$ and $\boldsymbol{\Sigma} = \sigma^2 \boldsymbol{I}_2$. We denote the AR coefficients by $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_p)'$.
2. *Magnitude-only AR(p) Ricean (MOR) model*: The MOR model is the MO (marginal) counterpart of the CV model; in other words, under the MOR model the latent real and imaginary time series follow the CV model. Under the MOR model, the probability density function (PDF) of magnitude measurements r_t is

$$f(r_t; \mu_t, \gamma_0) = \frac{r_t}{\gamma_0} \exp\left[\frac{-(r_t^2 + \mu_t^2)}{2\gamma_0}\right] \mathbb{I}_0\left(\frac{\mu_t r_t}{\gamma_0}\right), \tag{2}$$

where $\mu_t = \boldsymbol{x}_t'\boldsymbol{\beta}$, $\boldsymbol{x}_t'$ is the tth row of $\boldsymbol{X}$, $\gamma_0 = \mathrm{Var}(\eta_{Rt}) = \mathrm{Var}(\eta_{It})$ is zeroth order autocovariance of the latent real and imaginary errors and $\mathbb{I}_0(\cdot)$ is the modified Bessel function of the first kind and the zeroth order.

3. *Magnitude-only AR(p) Gaussian (MOG) model*: This model assumes $\boldsymbol{r} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, where $\boldsymbol{\epsilon}$ follows an AR(p) structure.

These three models are summarized in Table 1, which lists the key features that differentiate them. These features are as follows:

TABLE 1
Summary of the three models compared in this paper for complex-valued (CV) and magnitude-only (MO) fMRI data and the features present in each model. These features are italicized whenever they appear in the text

		Model features	
Model	Model Abbrev.	Twice the Quantities	Ricean magnitudes
Complex-valued AR(p)	CV	✓	✓
Magnitude-only Ricean AR(p)	MOR		✓
Magnitude-only Gaussian AR(p)	MOG		

---

[3] For brevity, the distinction between the notations of corresponding parameters in different models is dropped.

1. *Twice the quantities*: The CV model uses the real and imaginary measurements at each voxel and time-point, while the MOR/MOG models use only the magnitudes, so the CV model uses *twice the quantities* of the MO data-based models. (This does not necessarily mean that twice the amount of useful information is present in the CV over the MO data, especially for models of magnitude-related activation.[4])
2. *Ricean magnitudes*: The MOR model assumes the magnitude measurements are Ricean distributed, while the MOG model uses a Gaussian approximation. The CV model also has Rice-distributed magnitudes.

Table 1 illustrates the hierarchical pattern of the features present across the three models. As a result, comparing results for the CV/MOR models allows us to isolate the influence of having *twice the quantities* in the CV data. Similarly, comparing the MOR/MOG models shows the result of the Gaussian approximation of the *Ricean magnitudes*.

3.2. *Parameter estimation.* Of the three models introduced in Section 3.1, we focus on the methodology of the MO Ricean AR(p) (or MOR) model due to its novelty. We relegate methodological discussions of the CV and MOG models to Section S-2.1.

The methodology for the Ricean AR(p) model fits nicely into the framework of the EM algorithm (Dempster, Laird and Rubin (1977), McLachlan and Krishnan (2008)) with $\boldsymbol{r}$, $\boldsymbol{\phi}$ and $(\boldsymbol{r}, \boldsymbol{\phi})$ playing the roles of the observed, missing and complete data, respectively. And because the EM algorithm and its extensions serve as our “workhorse” methodology for the Ricean AR(p) model, we review it briefly here. An iteration of the EM algorithm consists of the expectation (E-) and maximization (M-) steps. At the $(k+1)$th iteration, the E-step calculates the objective function $Q(\boldsymbol{\tau}; \boldsymbol{\tau}^{(k)}) = \mathbb{E}_{\boldsymbol{\phi}|\boldsymbol{r},\boldsymbol{\tau}^{(k)}}[\log f(\boldsymbol{r}, \boldsymbol{\phi}; \boldsymbol{\tau})]$, the expectation of the complete data log-likelihood with respect to the conditional distribution $\boldsymbol{\phi}|\boldsymbol{r}$ at the current parameter estimates $\boldsymbol{\tau}^{(k)}$. The M-step calculates the updated parameter values $\boldsymbol{\tau}^{(k+1)} = \operatorname{argmax}_{\boldsymbol{\tau}} Q(\boldsymbol{\tau}; \boldsymbol{\tau}^{(k)})$ by maximizing the objective function. We denote the vector of parameters by $\boldsymbol{\tau} = (\boldsymbol{\alpha}', \boldsymbol{\beta}', \sigma^2)'$. The EM algorithm has well-known favorable properties such as monotone increase of the likelihood for each iteration and reliable global convergence (McLachlan and Krishnan (2008)).

3.2.1. *EM algorithm for ML estimation under the MO AR(p) Ricean model.* In this estimation procedure, we assume a known AR order p, after following the methods described in Section 3.4. To compute starting values $\boldsymbol{\tau}^{(0)}$, we use the MOG model, as demonstrated in Section S-2.1.2. With the algorithm initialized, the E- and M-steps are as follows:

3.2.1.1. *E-step.* The complete data log-likelihood can be shown to be

$$\log f(\boldsymbol{r}, \boldsymbol{\phi}; \boldsymbol{\tau}) = -n \log \sigma^2 - \log |\boldsymbol{R}_n| - h/2\sigma^2 \tag{3}$$

(Miller (1995), Pourahmadi (2001)), where $h = \tilde{\boldsymbol{\alpha}}' \boldsymbol{D} \tilde{\boldsymbol{\alpha}}$, with $\tilde{\boldsymbol{\alpha}}$ as the $(p+1)$-vector $(1, -\alpha_1, \dots, -\alpha_p)$ and $\boldsymbol{D}$ the $(p+1) \times (p+1)$ symmetric matrix with (i, j)th element

$$d_{ij} = \sum_{t=1}^{n-i-j} \begin{bmatrix} r_{t+i} r_{t+j} \cos(\phi_{t+i} - \phi_{t+j}) - \mu_{t+i} r_{t+j} \cos(\phi_{t+j} - \theta) \\ -\mu_{t+j} r_{t+i} \cos(\phi_{t+i} - \theta) + \mu_{t+i} \mu_{t+j} \end{bmatrix}, \tag{4}$$

where $\mu_t = \boldsymbol{x}_t' \boldsymbol{\beta}$, $\boldsymbol{x}_t'$ is the tth row of $\boldsymbol{X}$. In view of (3) and (4), the E-step involves two kinds of expectations: the univariate expectations $\mathbb{E}[\cos(\phi_t - \theta) | r_t; \boldsymbol{\tau}^{(k)}]$, $t = 1, \dots, n$ and the bivariate expectations $\mathbb{E}[\cos(\phi_t - \phi_{t+j}) | r_t, r_{t+j}; \boldsymbol{\tau}^{(k)}]$, $j = 1, \dots, p$, $t = 1, \dots, n-j$. The univariate expectations can be shown (Section S-2.3.1) to be

$$\mathbb{E}_{\phi_t | r_t; \boldsymbol{\tau}^{(k)}}[\cos(\phi_t - \theta)] = \mathbb{A}(\mu_t^{(k)} r_t / \gamma_0^{(k)}), \quad t = 1, \dots, n,$$

---

[4] Rowe (2005a) introduces fMRI models that allow for activation in both the magnitude and phase data.

where $\mathbb{A}(\xi) = \mathbb{I}_1(\xi)/\mathbb{I}_0(\xi)$, with $\mathbb{I}_m(\xi)$ being the mth order modified Bessel function of the first kind (Abramowitz and Stegun (1966)) evaluated at ξ.

The bivariate expectations are more cumbersome to obtain. First, we reduce $\mathbb{E}_{\phi_t,\phi_{t+j}|r_t,r_{t+j};\boldsymbol{\tau}^{(k)}}[\cos(\phi_t - \phi_{t+j})]$ to the univariate expectation

$$\mathbb{E}_{\phi_t|r_t,\boldsymbol{\tau}^{(k)}}\left\{\frac{\mathbb{A}(K(\phi_t))}{K(\phi_t)}\left[\kappa\cos(\phi_t - \theta) + \delta\right]\right\}, \tag{5}$$

where $K(\phi_t) = [\kappa^2 + \delta^2 + 2\kappa\delta\cos(\phi_t - \theta)]^{1/2}$, $\kappa = r_{t+j}(\gamma_0^{(k)}\mu_{t+j}^{(k)} - \gamma_j^{(k)}\mu_t^{(k)})/b^{(k)}$ and $\delta = \gamma_j^{(k)} r_t r_{t+j}/b^{(k)}$, with $b^{(k)} = \gamma_0^{2(k)} - \gamma_j^{2(k)}$. (See Section S-2.3.2 for more details.) Because (5) cannot be evaluated analytically, we approximate it via the Delta Method (Casella and Berger (2002)): $\mathbb{E}[f(X)] \approx f[\mathbb{E}(X)]$. When applied to (5), the Delta Method substitutes $\mathbb{A}(r_t\mu_t^{(k)}/\gamma_0^{(k)})$ for each instance of $\cos(\phi_t - \theta)$, including those in the $K(\phi_t)$ terms.

3.2.1.2. *M-step*. The global maxima of the objective function is not of closed form, so we obtain $\boldsymbol{\tau}^{(k+1)}$ through three conditional maximization steps as in the ECM algorithm (Meng and Rubin (1993)). First, we calculate the updated estimate $\boldsymbol{\alpha}^{(k+1)}$ via the equations

$$\sum_{j=1}^{p}\left(d_{ij}^{(k)} + 2j\gamma_{|j-i|}^{(k)}\right)\alpha_j = d_{i0}^{(k)}, \quad i = 1, \dots, p, \tag{6}$$

where $d_{ij}^{(k)}$ is the E-step expectation of d_{ij} with μ_t evaluated at $\mu_t^{(k)}$ and $\gamma_j^{(k)} = d_{0j}^{(k)}/(2n)$. Next, we calculate

$$\boldsymbol{\beta}^{(k+1)} = (\boldsymbol{X}'\boldsymbol{R}_n^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{R}_n^{-1}\boldsymbol{u}^{(k)}, \tag{7}$$

where $\boldsymbol{R}_n^{-1}$ is obtained from $\boldsymbol{\alpha}^{(k+1)}$ (as in Pourahmadi (2001)) and $\boldsymbol{u}^{(k)}$ is a vector of n variables with tth element $u_t^{(k)} = r_t\mathbb{A}(r_t\mu_t^{(k)}/\gamma_0^{(k)})$. Note that it may be necessary to enforce the boundary conditions $\boldsymbol{X}\boldsymbol{\beta}^{(k+1)} \geq \boldsymbol{0}$, in which case (7) needs to be modified as discussed in Section S-2.3.3. Finally, we calculate $\sigma^{2(k+1)} = h^{(k+1)}/(2n)$, where $h^{(k+1)} = \tilde{\boldsymbol{\alpha}}'^{(k+1)}\boldsymbol{D}^{(k+1)}\tilde{\boldsymbol{\alpha}}^{(k+1)}$ and $\boldsymbol{D}^{(k+1)}$ is a matrix as before with terms $d_{ij}^{(k+1)}$ evaluated using $\mu_t^{(k+1)} = \boldsymbol{x}_t'\boldsymbol{\beta}^{(k+1)}$.

3.2.2. *Hybrid algorithm for ML estimation*. As the EM algorithm progresses through iterations, we monitor convergence using the maximum change in the parameter values across successive iterations. However, as is commonly known (McLachlan and Krishnan (2008)), the convergence of the EM algorithm is slow at low SNRs, especially when $\beta_0/\sigma < 2$. In order to speed up convergence, we employ the hybrid algorithm of Aitkin and Aitkin (1996), which alternates the EM iterations with those from a modified Newton–Raphson (NR) method where the Fisher information matrix is replaced by the empirical information matrix (Meilijson (1989)). The hybrid algorithm starts with five EM iterations before switching to the modified NR method until convergence of the parameters. In the latter case, we halve the NR step size up to five times.

Parameter updates from the modified Newton–Raphson method are given by

$$\boldsymbol{\tau}^{(k+1)} = \mathscr{I}_e^{-1}(\boldsymbol{\tau}^{(k)}; \boldsymbol{r})\boldsymbol{S}(\boldsymbol{r}; \boldsymbol{\tau}^{(k)}), \tag{8}$$

where $\mathscr{I}_e(\boldsymbol{\tau}; \boldsymbol{r})$ is the empirical information matrix and $\boldsymbol{S}(\boldsymbol{r}; \boldsymbol{\tau})$ is the score statistic. Both are constructed from the contributions to the score statistic at $t=p+1, p+2, \dots, n$, denoted by $\boldsymbol{s}(r_t; \boldsymbol{\tau}) = \frac{\partial}{\partial\boldsymbol{\tau}}\log f(r_t|r_{t-1}, \dots, r_{t-p}; \boldsymbol{\tau})$, which can be calculated from the complete data log-likelihood using the identity (adapted from McLachlan and Krishnan (2008))

$$\boldsymbol{s}(r_t; \boldsymbol{\tau}) = \mathbb{E}_{\boldsymbol{\phi}|\boldsymbol{r};\boldsymbol{\tau}}\left[\frac{\partial}{\partial\boldsymbol{\tau}}\log f\big((r_t, \phi_t)|(r_{t-1}, \phi_{t-1}), \dots, (r_{t-p}, \phi_{t-p}); \boldsymbol{\tau}\big)\right]. \tag{9}$$

These calculations, detailed in Section S-2.3.4, use quantities from the E-step. The empirical information matrix is calculated as

$$\mathscr{I}_e(\boldsymbol{\tau};\boldsymbol{r}) = \sum_{t=p+1}^{n} \boldsymbol{s}(r_t;\boldsymbol{\tau})\boldsymbol{s}'(r_t;\boldsymbol{\tau}) - \frac{1}{n-p}\boldsymbol{S}(\boldsymbol{r};\boldsymbol{\tau})\boldsymbol{S}'(\boldsymbol{r};\boldsymbol{\tau}), \tag{10}$$

where $\boldsymbol{S}(\boldsymbol{r};\boldsymbol{\tau}) = \sum_{t=p+1}^{n} \boldsymbol{s}(r_t;\boldsymbol{\tau})$.

3.3. *Calculation of test statistics under the MO AR*(p) *Ricean model.*

We illustrate the calculations of Wald and likelihood ratio test statistics for a general test for activation, which posits $H_0 : \boldsymbol{C}\boldsymbol{\beta} = \boldsymbol{0}$ against $H_a : \boldsymbol{C}\boldsymbol{\beta} \neq \boldsymbol{0}$. Each test statistic is based on the MLEs $\widehat{\boldsymbol{\tau}}$ calculated by the above EM/NR hybrid algorithm.

3.3.1. *Wald test.*

The Wald test statistic is given by

$$W = (\boldsymbol{C}\widehat{\boldsymbol{\beta}})'\big[\boldsymbol{C}\mathscr{I}_e^{-1}(\widehat{\boldsymbol{\tau}};\boldsymbol{r})\boldsymbol{C}'\big]^{-1}(\boldsymbol{C}\widehat{\boldsymbol{\beta}}) \tag{11}$$

and asymptotically follows a null χ^2_m-distribution, where m is the rank of $\boldsymbol{C}$. It utilizes the empirical information matrix $\mathscr{I}_e$ of (10). However, our simulation studies reported in Section S-2.4.1 indicate that the Wald test statistic does not follow this null distribution for low-SNR time series, that is, when $\beta_0/\sigma < 2$. This shortcoming of the Wald test motivates the derivation of the likelihood ratio test (LRT) statistic described below.

3.3.2. *Likelihood ratio test for Ricean AR*(1) *model.*

We derive an LRT for the Ricean AR(1) model, whose false positive rate better conforms with the significance level than the Wald test for low-SNR time series, as shown in Figure S-6. This LRT statistic is only for the Ricean AR(1) model, as the observed data loglikelihood is quite intractible for higher AR orders.[5] From standard results the LRT statistic $\Lambda = 2[\ell(\widehat{\boldsymbol{\tau}}) - \ell(\widetilde{\boldsymbol{\tau}})]$, where $\ell(\cdot)$ is the log-likelihood function $\log L(\cdot)$ and $\widehat{\boldsymbol{\tau}}$ and $\widetilde{\boldsymbol{\tau}}$ represent the MLEs of $\boldsymbol{\tau}$ under H_a and H_0, respectively. Like the Wald statistic, the LRT statistic asymptotically follows a null χ^2_m-distribution. To derive the likelihood function $L(\boldsymbol{\tau}) = f(\boldsymbol{r};\boldsymbol{\tau})$ for the Ricean AR(1) model, we note that $f(\boldsymbol{r};\boldsymbol{\tau})$ can be factored as $f(r_1;\boldsymbol{\tau})\prod_{t=2}^{n} f(r_t|r_{t-1};\boldsymbol{\tau})$, where $f(r_1;\boldsymbol{\tau})$ is the Ricean PDF of (2). It can be shown (see Section S-2.4.2) that the conditional PDF $f(r_t|r_{t-1};\boldsymbol{\tau})$ is equal to

$$\frac{r_t}{\sigma^2}e^{C_0}\left[\mathbb{I}_0\left(\frac{r_{t-1}\mu_{t-1}}{\gamma_0}\right)\right]^{-1}\sum_{m=0}^{\infty}\omega_m \mathbb{I}_m(C_1)\mathbb{I}_m(C_2)\mathbb{I}_m(C_{12}), \tag{12}$$

where $C_0 = -[r_t^2 + \mu_t^2 + \alpha^2(r_{t-1}^2 + \mu_{t-1}^2) - 2\alpha\mu_{t-1}\mu_t]/(2\sigma^2)$, $C_1 = r_{t-1}(\mu_{t-1} - \alpha\mu_t)/\sigma^2$, $C_2 = r_t(\mu_t - \alpha\mu_{t-1})/\sigma^2$ and $C_{12} = \alpha r_{t-1}r_t/\sigma^2$. Also, in (12), $\omega_m = 1$ for $m = 0$ and $\omega_m = 2$ for $m \geq 1$.

3.4. *Choosing the order of the AR model.*

We suggest a sequential testing approach for determining the AR order p. Starting with $k = 1$ and for increasing k, we posit $H_0 : p = k - 1$ vs. $H_a : p \geq k$ (or, in terms of the AR coefficients, $H_0 : \forall j \geq k, \alpha_j = 0$ vs. $H_a : \exists j \geq k : \alpha_j \neq 0$). The estimated AR order is then $\widehat{p} = k' - 1$, where k' is the first k in the sequence of tests for which H_0 is not rejected. An LRT statistic given by $2(\ell_k - \ell_{k-1})$, where ℓ_k is the optimized log-likelihood function for the AR(k) model, may be employed under the CVS and MOG models. For the MOR model, we use the Wald test statistic

---

[5] Recall that the Wald test statistic can derived for any AR order.

$\widehat{\alpha}_k^2/\mathscr{I}_e^{-1}(\widehat{\boldsymbol{\tau}}, \boldsymbol{r})_{\alpha_k\alpha_k}$, where the denominator is the diagonal entry of the inverse empirical information matrix corresponding to α_k. From standard results, each test statistic (whether LRT or Wald) is asymptotically χ_1^2-distributed under $H_0 : p = k - 1$.

It can be shown that the significance level δ applied to each test controls "over-detecting" the order (i.e., $\widehat{p} > p$) in the sense that $\delta = \mathbb{P}(\widehat{p} > p | \widehat{p} \geq p)$; see Adrian, Maitra and Rowe ((2018), Section S-2.4) for a justification. Section S-2.5 uses simulation to demonstrate that the sequential testing approach to detecting p gives similar results to approaches based on the AIC and BIC (Akaike (1973), Schwarz (1978)).

Our development in this section has laid the groundwork for our investigation of fMRI analyses using the three models. We now proceed with our evaluations.

4. Simulation experiments.

We perform simulation experiments to compare the AR(p) Ricean model to two other models in a setting of known truth; see Section 3.1 and Table 1 for a summary of these three models: the CV, MOR and MOG models. We simulated low-SNR complex-valued time series under the CV model—and, therefore, also magnitude time series under the MOR model—with the $\boldsymbol{X}$ matrix of the finger-tapping experiment (see Section 2). We used white noise variance $\sigma^2 = 1$, AR(1) temporal dependence with AR parameters $\alpha = 0.2, 0.4, 0.6$, and 0.8, baseline signal levels β_0 from 0.5 to 5.0 and activation levels $\beta_1 = 0.1, 0.2$ and 0.3. Thus, the SNRs and CNRs of the simulated data corresponded with these in the dataset (see Figure S-3). For each combination of parameter values, we generated 10,000 voxel time series and fit each of the three models under an assumed AR order of 1. Our model comparison analysis has two main components: properties of the parameter estimates and activation detection performance. To quantify simulation variability, we calculated standard errors for all quantities using the bootstrap method (Efron and Tibshirani (1986)) with 1000 replications.

4.1. *Properties of parameter estimates.*

Figure 2 summarizes the properties of the parameter estimates, displaying the biases, standard errors (SEs) and root mean squared errors (RMSEs) of $\widehat{\beta}_0$, $\widehat{\beta}_1$, $\widehat{\sigma}^2$ and $\widehat{\alpha}$. (It shows results for $\alpha = 0.4$; results for other αs are given in Figure S-9.) Focusing first on the biases, it is immediately evident that the MOG model produces the most biased parameter estimates due to its Gaussian approximation of the truly Ricean magnitudes. Specifically, the biases of $\widehat{\beta}_0$ and $\widehat{\sigma}^2$ result from the mean and variance of the Rice distribution (which are the quantities that the MOG model estimates) being above and below the Ricean location and scale parameters (Zhu et al. (2009)). The MOG model biases decrease with increasing β_0 because the Gaussian approximation to the Rice distribution improves with SNR. These results match those already observed for the temporally independent case (Adrian, Maitra and Rowe (2013), Solo and Noh (2007)), and additionally, we see here that the MOG model-based estimate of the AR coefficient α is the most biased as well. The MOR model-based parameter estimates also show some biases at the lowest SNR time series; the CV model-based estimates are unbiased in all cases.

Switching our focus to the SEs, we note that the CV model-based SEs of $\widehat{\sigma}^2$ and $\widehat{\alpha}$ are lower than those for the MO models. This corresponds to a similar result, observed in Rowe (2005b), about the sampling variances of $\widehat{\sigma}^2$ for the temporally independent case and suggests that the *twice the quantities* feature of the CV data is driving this difference. Overall, the RMSEs, which account for both bias and SE as $\mathrm{RMSE}^2(\cdot) = \mathrm{Bias}^2(\cdot) + \mathrm{SE}^2(\cdot)$, are lowest for the CV model; in addition, the CV model-based results are constant with β_0, that is, not related to the SNR. This suggests that, provided the CV data is collected, the CV model produces the most reliable parameter estimates (arguing against the current practice of discarding the phase data). However, given that archival datasets (of which there are massively many) do not have phase data stored, our results here also suggest the parameter estimation benefits of using the AR MOR model in place of the currently-used AR MOG model, especially in terms of bias for low SNR time series.

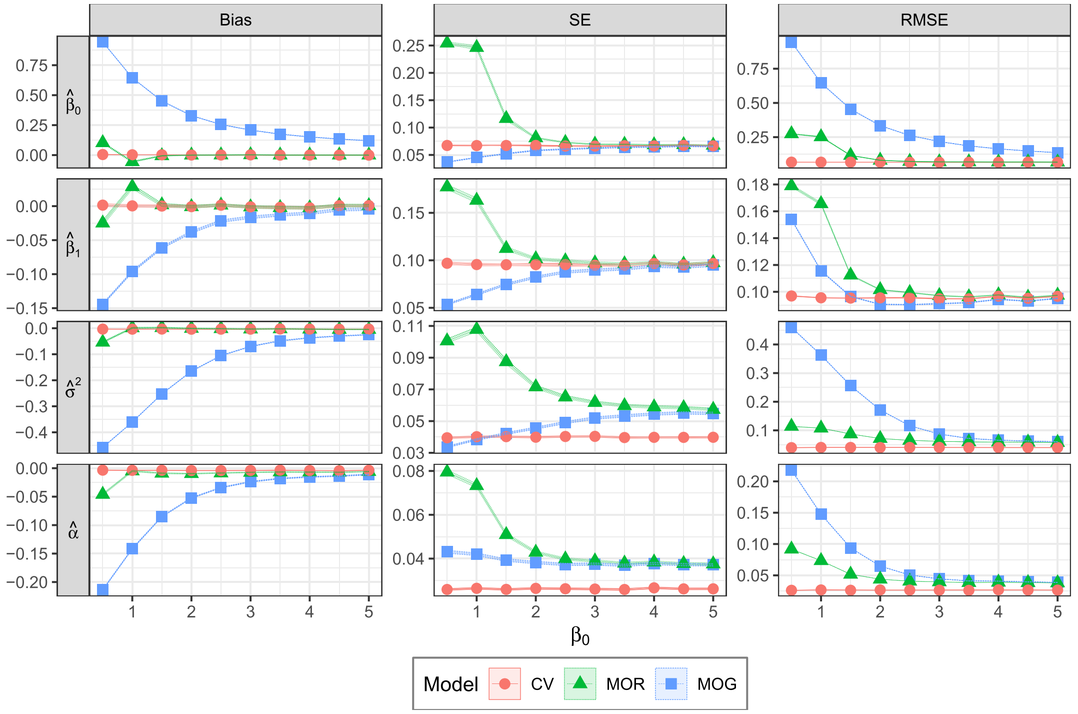

FIG. 2. *Biases, SEs (SE) and RMSEs (RMSE) of $\widehat{\beta}_0$, $\widehat{\beta}_1$, $\widehat{\sigma}^2$ and $\widehat{\alpha}$ under three models for simulated time series under $\beta_1 = 0.2$, $\alpha = 0.4$ and different values of β_0. The shaded areas cover the estimate $\pm$ 2 standard error intervals.*

4.2. *Activation detection performance.* Next, we compare activation detection performance, calculating LRT statistics for the activation test of $H_0 : \beta_1 = 0$ vs. $H_a : \beta_1 \neq 0$ for each simulated time series. To summarize the power of each LRT statistic, we calculated the partial area under the receiver operating characteristic curve or the pAUC (McClish (1989), Zhou, Obuchowski and McClish (2011)). The pAUC is the area under the ROC curve—where the ROC curve plots the true positive rate (TPR) against the false positive rate (FPR)—over a limited range of FPR values.[6] The rationale for using the pAUC rather than the (full) AUC, the area under the ROC curve for all FPR values from 0 to 1, is to exclude contributions to the curves from FPR values that are never used in practice and to highlight differences between the methods in the part of the ROC that are most likely to be used in practice. For instance, using FPR values greater than 0.1 would allow for far too many false positives than practically ever used; indeed, our significance level threshold used on the real data in Section 5 is 0.001. Therefore, we calculated the pAUCs over a FPR range of 0 to 0.05. We calculated the pAUCs (separately for each combination of parameters β_0, β_1 and α) as the average of the TPRs for the significance levels $\delta = 0.0001, 0.0002, 0.0003, \ldots, 0.0500$; each TPR is the proportion of the 10,000 simulated test statistics greater than the $\chi^2_{1-\delta,1}$ quantile.

As shown in Figure 3, the pAUCs of simulated LRT statistics are consistently in the order (highest to lowest) of CV, MOR and MOG models. While the figure shows between-panel differences in the pAUCs, due to the relationships between the CNR and the values of β_1 and α (positive and negative relationships, respectively), the within-panel patterns are quite similar. While the CV model-based pAUCs are relatively constant as a function of β_0, the pAUCs of the two MO model-based LRTSs decrease with β_0.

---

[6] There are also pAUC versions that limit the TPR range (or both the FPR and TPR ranges), but we use a FPR-limited pAUC here.

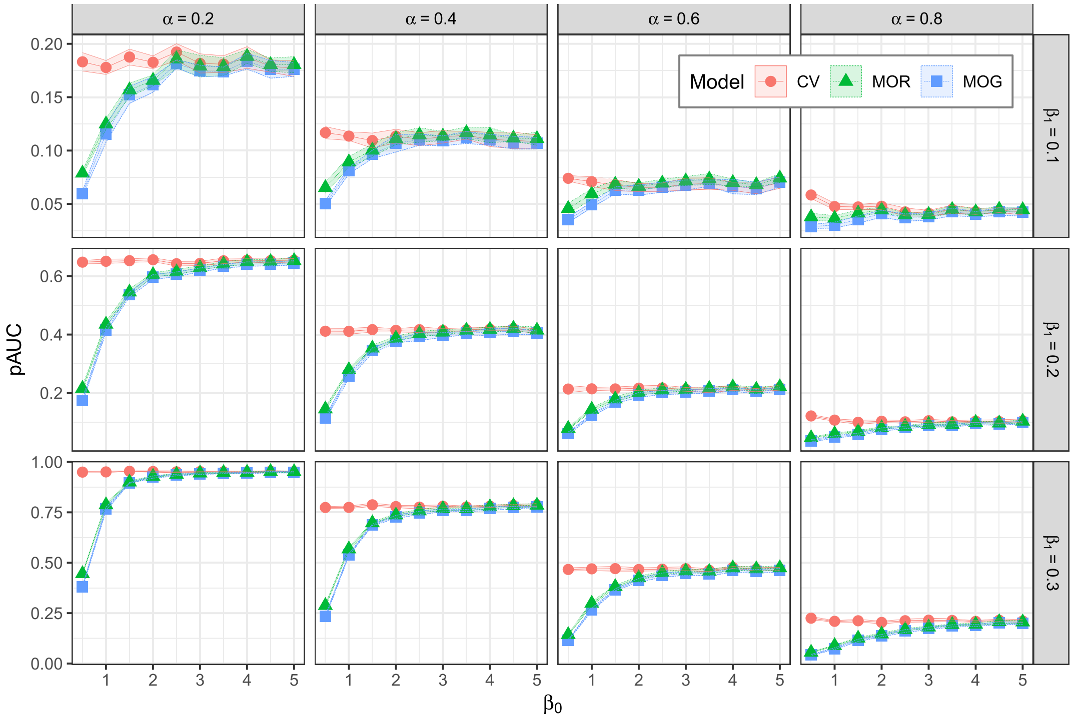

FIG. 3. *The partial AUCs of the magnitude-only* (*MO*) *data-based LRT statistics decrease at low SNRs relative to their complex-valued* (*CV*) *data-based counterparts. Also, the Ricean* (*MOR*) *model-based pAUCs show improvements over those from the Gaussian* (*MOG*) *model. The shaded regions show the pAUCs* ± 2 *standard errors.*

We attribute the increased MOR model-based pAUCs relative to the MOG model to proper modeling of the Rice-distributed magnitudes: the MOG model resorts to a Gaussian approximation of the Rice distribution. Figure 4 displays the percent pAUC increases of the MOR model-based LRTSs relative to the MOG model. We note that the sizes of MOR model improvements in pAUC increase as β_0 and β_1 decrease and α increases, which correspond to decreases in SNR and CNR. While it is not surprising that a lower SNR (with worse Gaussian approximation of the Ricean magnitudes) is associated with a larger MOR model improvement, it is worth noting that improvement is larger for lower activation levels (CNRs) as well.

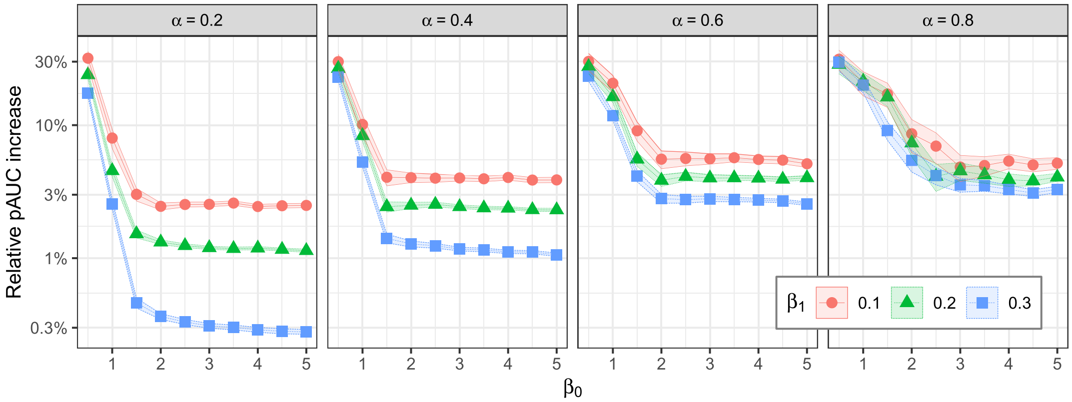

FIG. 4. *Relative improvement in MOG model-based pAUCs by MOR model, that is,* (*MOR–MOG*) / *MOG. Shaded areas show estimate* ± 2 *standard errors.*

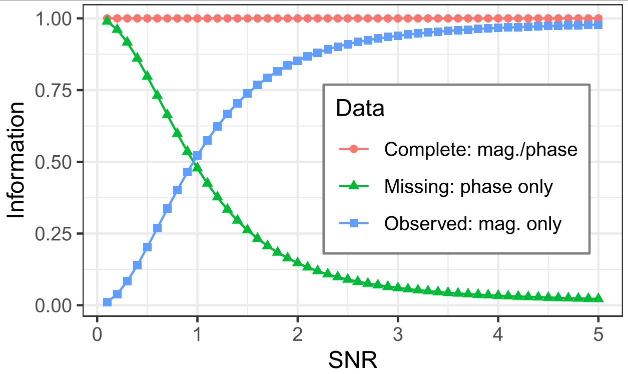

FIG. 5. *Plots of the information matrices as a function of the SNR.*

As for the CV vs. MO comparison, the *twice the quantities* feature seems to be driving the difference in detection performance. The additional information in the phase data (again, which is usually not collected) appears to be more valuable to the activation detection as β_0 (the SNR) decreases, as shown by the increasing discrepancy in CV/MO pAUCs. In fact, we may view this in terms of the "missing information principle" (Orchard and Woodbury (1972)). Section S-3.1.1 describes how to calculate the observed (MO), missing (phase only) and complete (CV) data-based the Fisher information matrices by applying the EM algorithm. Interestingly, the curves of the "observed information" (in the MO data) and the "complete information" (in the CV data) in Figure 5 look similar to the MO and CV model-based pAUC curves in Figure 3. Both figures suggest that the phase data contain useful information about the activation, even though the activation itself occurs in the magnitude signal $\boldsymbol{X\beta}$.

To summarize, the results of our experiments suggest that modeling the *Ricean magnitudes* produces less-biased parameter estimates and better activation detection performance than the Gaussian approximation. Also, using *twice the quantities* in CV rather than MO data-based models leads to parameter estimates with less variation and additionally improved activation detection. Although using the CV data produces the best estimation and activation detection, the CV data has historically not been collected; in this case, utilizing the Ricean model when only magnitude (archival) data is available produces sizeable gains over the Gaussian approximation.

5. Results on low-signal fMRI dataset.

In this section we identify activation in the low-SNR fMRI dataset introduced in Section 2 under the CV, MOR and MOG models. Working with each voxel time series of the dataset separately, we first detected the AR orders (see Figure S-10) and then tested for activation using $H_0 : \beta_1 = 0$ vs. $H_a : \beta_1 \neq 0$, obtaining p-values from the LR and Wald test statistics under the null χ^2_1-distribution. To determine activation, we used a significance level threshold of 0.001 (Woo, Krishnan and Wager (2014)).

5.1. *Activation detection in raw data.*

We display activation maps of the second slice[7] according to each of the three models in Figure 6. Each map shows a grayscale anatomical image (the magnitude image at the first time point) with the voxels having p-values less than 0.001 colored according to the legend. The activation is rather sparse, except for a region containing the left central sulcus (Figure 1)—recall, the site guiding right-hand function for normal healthy adults—so we focus on this region of interest (ROI) in the inset maps.

---

[7] We focus on the second slice (in the superior direction) because it shows the strongest activation. For comparison with the two neighboring slices, see Figure **??**.

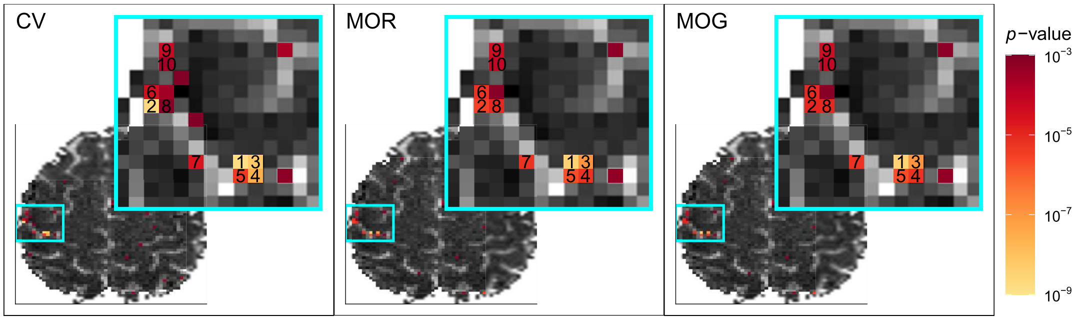

FIG. 6. *Activation maps with inset maps showing left central sulcus ROI; numbered voxels represent those having test statistics with the lowest p-values.*

Specifically, we focus on the 10 voxels in this ROI with test statistics that provide the lowest minimum p-value over the three models, which are identified by numerals (from $1 =$ lowest to $10 =$ highest p-value) in the inset maps of Figure 6. This numbering carries over to Table 2 where the size of the p-values can be compared more clearly across models. (See Figure S-12 for a graphical representation of Table 2.) Though this comparison varies of across voxels, we note that the average of the log base-10 p-values is lowest for the CV model, followed by the MOR model and then by the MOG model. As we saw in simulation experiments, the experimental data also indicates that if the CV data is available, the CV model produces better activation detection power than MO data-based models. However, in the common case of MO data, our AR(p) Ricean model shows more power than the the corresponding model based on a Gaussian approximation to the Rice distribution.

5.2. *Adding noise to the raw data.* To further investigate the effect that low SNR has on activation detection, we added more noise to the acquired CV data. So for the 10 voxels identified in Figure 6 with acquired CV time series (y_{Rt}, y_{It}), $t = 1, \dots, 621$, we obtained new synthetic data $y^*_{Rt} = y_{Rt} + w_{Rt}$ and $y^*_{It} = y_{It} + w_{It}$, with $w_{Rt}, w_{It} \sim$ i.i.d. $N(0, \sigma_a^2)$. Because a representative, data-based estimate of the noise SD is 0.15, we generated data using $\sigma_a = 0.15, 0.20, 0.25$ and 0.30. Adding noise to the original data in this manner reduces the SNR and CNR (see Figure S-13) as would occur with increasing spatial resolution or decreasing field strength of the MR scan. Table 3 shows the proportions of the 10,000 generated datasets in which each of the 10 voxels was detected under each model at the 0.001 significance level (also see Figure S-14). With the added noise, the detection power more consistently follows the order of CV (greatest), MOR and MOG (least) models across the 10 voxels than the p-values of the raw data in Figure 6.

TABLE 2
Negative logarithm (base 10) of the p-values for the voxels numbered 1–10 in the inset maps of Figure 6 for test statistics under the CV, MOR and MOG models. (For instance, the CV model-based p-value for voxel 1 is $10^{-8.82}$.) Also, the "avg" column represents the average across voxels

	Voxel numbers from Figure 6										
Model	1	2	3	4	5	6	7	8	9	10	avg
CV	8.82	8.70	8.18	7.53	5.31	4.42	4.65	3.14	4.24	3.61	5.86
MOR	8.48	5.52	6.93	5.20	5.71	5.21	5.20	5.00	4.37	4.13	5.57
MOG	8.49	4.96	6.94	5.20	5.71	5.03	5.21	5.00	4.37	4.13	5.50

TABLE 3

Proportions (and their standard errors) of simulated time series detected as detected when noise with standard deviation σ_a was added to the raw data for the 10 *voxels identified in Figure* 6. *The "avg" rows show average proportions over the* 10 *voxels*

		SD of added noise σ_a			
Voxel	Model	0.15	0.2	0.25	0.3
1	CV	0.951 (0.002)	0.731 (0.004)	0.476 (0.005)	0.304 (0.004)
1	MOR	0.931 (0.002)	0.670 (0.004)	0.390 (0.004)	0.216 (0.004)
1	MOG	0.903 (0.003)	0.625 (0.004)	0.363 (0.004)	0.200 (0.004)
2	CV	0.947 (0.002)	0.719 (0.004)	0.479 (0.005)	0.306 (0.004)
2	MOR	0.929 (0.002)	0.660 (0.004)	0.409 (0.004)	0.230 (0.004)
2	MOG	0.868 (0.003)	0.579 (0.005)	0.357 (0.004)	0.204 (0.004)
3	CV	0.896 (0.003)	0.633 (0.004)	0.399 (0.004)	0.238 (0.004)
3	MOR	0.717 (0.004)	0.398 (0.004)	0.212 (0.004)	0.103 (0.003)
3	MOG	0.718 (0.004)	0.405 (0.004)	0.209 (0.004)	0.102 (0.003)
4	CV	0.880 (0.003)	0.611 (0.004)	0.378 (0.004)	0.229 (0.004)
4	MOR	0.408 (0.004)	0.180 (0.004)	0.090 (0.003)	0.045 (0.002)
4	MOG	0.351 (0.004)	0.156 (0.003)	0.081 (0.002)	0.042 (0.002)
5	CV	0.490 (0.005)	0.261 (0.004)	0.152 (0.003)	0.090 (0.003)
5	MOR	0.513 (0.005)	0.244 (0.004)	0.121 (0.003)	0.057 (0.002)
5	MOG	0.442 (0.005)	0.210 (0.004)	0.106 (0.003)	0.052 (0.002)
6	CV	0.811 (0.004)	0.525 (0.005)	0.314 (0.004)	0.192 (0.004)
6	MOR	0.788 (0.004)	0.468 (0.005)	0.242 (0.004)	0.132 (0.003)
6	MOG	0.629 (0.004)	0.357 (0.004)	0.188 (0.004)	0.108 (0.003)
7	CV	0.361 (0.004)	0.187 (0.004)	0.109 (0.003)	0.068 (0.002)
7	MOR	0.412 (0.004)	0.201 (0.004)	0.108 (0.003)	0.060 (0.002)
7	MOG	0.372 (0.004)	0.187 (0.004)	0.102 (0.003)	0.057 (0.002)
8	CV	0.443 (0.005)	0.240 (0.004)	0.133 (0.003)	0.085 (0.003)
8	MOR	0.410 (0.004)	0.214 (0.004)	0.108 (0.003)	0.065 (0.002)
8	MOG	0.322 (0.004)	0.171 (0.003)	0.091 (0.003)	0.057 (0.002)
9	CV	0.293 (0.004)	0.165 (0.003)	0.085 (0.003)	0.058 (0.002)
9	MOR	0.282 (0.004)	0.143 (0.003)	0.068 (0.002)	0.038 (0.002)
9	MOG	0.283 (0.004)	0.142 (0.003)	0.067 (0.002)	0.039 (0.002)
10	CV	0.193 (0.004)	0.109 (0.003)	0.067 (0.002)	0.042 (0.002)
10	MOR	0.216 (0.004)	0.109 (0.003)	0.054 (0.002)	0.028 (0.002)
10	MOG	0.182 (0.004)	0.093 (0.003)	0.048 (0.002)	0.025 (0.001)
avg	CV	0.627 (0.001)	0.418 (0.001)	0.259 (0.001)	0.161 (0.001)
avg	MOR	0.561 (0.001)	0.329 (0.001)	0.180 (0.001)	0.097 (0.001)
avg	MOG	0.507 (0.001)	0.293 (0.001)	0.161 (0.001)	0.089 (0.001)

6. Discussion.

In this paper we developed a Ricean model for fMRI magnitude time series that incorporates autoregressive time dependence. Our approach applies AR(p) errors to the Gaussian-distributed real and imaginary components from which the magnitudes are computed. We estimated model parameters from the MO data using the EM algorithm with the phase portion of the latent complex-valued data playing the role of missing data. We then extended the EM algorithm to derive Wald and LRTs for activation and AR order detection.

We compared this AR(p) Ricean model to two other models: the CV model for complex-valued data and the MOG model which employs a Gaussian approximation for the truly

Rice-distributed magnitudes. As previously discussed, complex-valued fMRI data is rarely collected in practice (although we strongly advocate for it), so our main focus is how to make the best use of existing (i.e., archived) magnitude-only data through the MOR rather than the MOG model. We expected the CV model to perform better than the two MO data-based models simply due to the CV data having *twice the quantities* and the MOR model to perform better than the MOG model due to properly modeling the *Ricean magnitudes*.

Simulation-based results confirmed these expectations when we evaluated the quality of parameter estimates and the activation detection under each model. The parameter estimates of the CV model had lower RMSEs than the MO data-based models, and the MOR model-based parameter estimates were much less biased than those from the MOG model. Similarly, we demonstrated greater activation detection power (as measured by pAUC) for the CV model than the two MO data-based models, and the MOR model-based pAUCs were greater than those from the MOG model. For all comparisons the differences were greatest at low SNRs.

We then compared these models on a finger-tapping experiment, acquiring the data using the body coil to study a low-SNR dataset. For voxels in the expected activation region showing the most activation, the CV model produced the lowest average p-values, followed by the MOR model (and the MOG model last). Although the order of the p-values varied by voxel, we are able to more consistently superior activation detection of the CV model (and MOR model greater than MOG model) when we added extra simulated white noise to the experimental voxel time series, which decreased the SNR.

A reviewer has asked about the computational complexity of MO Ricean over CV data-based methods. We note that, while the CV-based analysis has a larger storage requirement, the CV-based and Gaussian MO models are of lower computs and can use direct methods to estimate parameters, the CV-based analysis is the highest since it deals with twice the amount of data; however, in terms of CPU usage, CV-based analysis (and Gaussian MO-only analysis) have lower computational complexity than the Rice-based model. However, the Rice-based MO analysis is more accurate, in terms of parameter estimation and activation detection than the Gaussian MO-only analysis, especially at lower SNR.

Future directions for this research can explore fMRI time series models that allow for activation in magnitude and phase (Rowe (2005a)). The models we have focused upon in this article (even the CV data-based ones) have assumed task-related changes in magnitude only, with no task-related phase changes. It may be worth exploring a single model that allows for both. Another area for future work lies in generalizing the modeling of magnitude time series beyond the AR(p) Rice model presented here. MR images collected by simultaneous acquisition from multiple independent coils (Tristán-Vega, Aja-Fernández and Westin (2012)) can be shown to follow the noncentral chi distribution, with degrees of freedom equal to twice the number of coils (Wegmann, Eklund and Villani (2017)). The Rice distribution is the special case for a single coil and two degrees of freedom.

Acknowledgments. We thank the Editor, an anonymous Associate Editor and an anonymous reviewer whose comments on a earlier version of this paper greatly improved its content and presentation.

Funding. The second author was supported in part by the the National Science Foundation CAREER Grant # DMS-0437555 and the National Institutes of Health (NIH) awards #R21EB016212 and #R21EB034184.

The third author was supported in part by the National Institutes of Health (NIH) award #R21NS087450.

SUPPLEMENTARY MATERIAL

Supplement to "Rice-distributed autoregressive time series modeling of magnitude functional MRI data" (DOI: 10.1214/24-AOAS1981SUPPA; .pdf). Section S-1 provides details of the low-signal finger-tapping dataset used in this paper. Section S-2 provides additional methodological details and derivations. Additional details on simulation experiments are in Section S-3, while more details on the analysis of the finger-tapping dataset are provided in Section S-4.

Dataset (DOI: 10.1214/24-AOAS1981SUPPB; .zip). Also available on GitHub.

REFERENCES

ABDI, A., TEPEDELENLIOGLU, C., KAVEH, M. and GIANNAKIS, G. (2001). On the estimation of the K parameter for the Rice fading distribution. *IEEE Commun. Lett.* **5** 92–94.

ABRAMOWITZ, M. and STEGUN, I. A. (1966). *Handbook of Mathematical Functions, with Formulas, Graphs, and Mathematical Tables*. Dover, New York. MR0208797

ADRIAN, D. W., MAITRA, R. and ROWE, D. B. (2013). Ricean over Gaussian modelling in magnitude fMRI analysis—added complexity with negligible practical benefits. *Stat* **2** 303–316. MR4027320 https://doi.org/10.1002/sta4.34

ADRIAN, D. W., MAITRA, R. and ROWE, D. B. (2018). Complex-valued time series modeling for improved activation detection in fMRI studies. *Ann. Appl. Stat.* **12** 1451–1478. MR3852684 https://doi.org/10.1214/17-AOAS1117

ADRIAN, D. W., MAITRA, R. and ROWE, D. B. (2025). Supplement to "Rice-distributed autoregressive time series modeling of magnitude functional MRI data." https://doi.org/10.1214/24-AOAS1981SUPPA, https://doi.org/10.1214/24-AOAS1981SUPPB

AITKIN, M. and AITKIN, I. (1996). A hybrid EM/Gauss-Newton algorithm for maximum likelihood in mixture distributions. *Stat. Comput.* **6** 127–130.

AKAIKE, H. (1973). Information theory and an extension of the maximum likelihood principle. In *Second International Symposium on Information Theory* (*Tsahkadsor*, 1971) 267–281. Akad. Kiadó, Budapest. MR0483125

ARAFA, A. and MESSIER, G. G. (2010). A Gaussian model for dead-reckoning mobile sensor position error. In 2010 *IEEE 72nd Vehicular Technology Conference—Fall* 1–5.

BAGGIO, R. and MUZY, J.-F. (2024). Improving probabilistic wind speed forecasting using M-Rice distribution and spatial data integration. *Appl. Energy* **360** 122840.

BAÏLE, R., MUZY, J. F. and POGGI, P. (2011). An M-Rice wind speed frequency distribution. *Wind Energy* **14** 735–748.

BALCHANDANI, P. and NAIDICH, T. P. (2015). Ultra-high-field MR neuroimaging. *Am. J. Neuroradiol.* **36** 1204–1215. https://doi.org/10.3174/ajnr.A4180

BANDETTINI, P. A., JESMANOWICZ, A., WONG, E. C. and HYDE, J. S. (1993). Processing strategies for time-course data sets in functional MRI of the human brain. *Magn. Reson. Med.* **30** 161–173. https://doi.org/10.1002/mrm.1910300204

BAR-SHALOM, Y., LI, X. R. and KIRUBARAJAN, T. (2001). *Estimation with Applications to Tracking and Navigation: Theory Algorithms and Software*. Wiley, New York.

BELLIVEAU, J. W., KENNEDY, JR., D. N., MCKINSTRY, R. C., BUCHBINDER, B. R., WEISSKOFF, R. M., COHEN, M. S., VEVEA, J. M., BRADY, T. J. and ROSEN, B. R. (1991). Functional mapping of the human visual cortex by magnetic resonance imaging. *Science* **254** 716–719. https://doi.org/10.1126/science.1948051

BEST, D. J., RAYNER, J. C. W. and THAS, O. (2010). Easily applied tests of fit for the Rayleigh distribution. *Sankhya B* **72** 254–263. MR2820134 https://doi.org/10.1007/s13571-011-0011-2

BROWN, T. R., KINCAID, B. M. and UGURBIL, K. (1982). NMR chemical shift imaging in three dimensions. In *Proceedings of the National Academy of Sciences, USA* **79** 3523–3526.

BULLMORE, E., BRAMMER, M., WILLIAMS, S. C. R., RABE-HESKETH, S., JANOT, N., DAVID, A., MELLERS, J., HOWARD, R. and SHAM, P. (1996). Statistical methods of estimation and inference for function MR image analysis. *Magn. Reson. Med.* **35** 261–277.

CARPENTER, M. B. (1991). *Core Text of Neuroanatomy*. Williams & Wilkins, Baltimore.

CASELLA, G. and BERGER, R. L. (2002). *Statistical Inference*, 2nd ed. Thomson Learning.

COSOTTINI, M. and ROCCATAGLIATA, L. (2021). Neuroimaging at 7 T: Are we ready for clinical transition? *Eur. Radiol. Exp.* **5** 37. https://doi.org/10.1186/s41747-021-00234-0

DEMPSTER, A. P., LAIRD, N. M. and RUBIN, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Statist. Soc. Ser. B* **39** 1–38. MR0501537

DEN DEKKER, A. J., POOT, D. H. J., BOS, R. and SIJBERS, J. (2009). Likelihood-based hypothesis tests for brain activation detection from MRI data disturbed by colored noise: A simulation study. *IEEE Trans. Med. Imag.* **28** 287–296. https://doi.org/10.1109/TMI.2008.2004427

EFRON, B. and TIBSHIRANI, R. (1986). Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statist. Sci.* **1** 54–77. MR0833275

FRISTON, K. J., FRITH, C. D., LIDDLE, P. F., DOLAN, R. J., LAMMERTSMA, A. A. and FRACKOWIAK, R. S. J. (1990). The relationship between global and local changes in PET scans. *J. Cereb. Blood Flow Metab.* **10** 458–466.

FRISTON, K. J., HOLMES, A. P., WORSLEY, K. J., POLINE, J. B., FRITH, C. D. and FRACKOWIAK, R. S. J. (1995). Statistical parametric maps in functional imaging: A general linear approach. *Hum. Brain Mapp.* **2** 189–210.

FRISTON, K. J., JEZZARD, P. and TURNER, R. (1994). Analysis of functional MRI time-series. *Hum. Brain Mapp.* **1** 153–171.

FRISTON, K. J., JOSEPHS, O., ZARAHN, E., HOLMES, A. P., ROUQUETTE, S. and POLINE, J. (2000). To smooth or not to smooth? Bias and efficiency in fMRI time-series analysis. *NeuroImage* **12** 196–208. https://doi.org/10.1006/nimg.2000.0609

GENOVESE, C. R., LAZAR, N. A. and NICHOLS, T. E. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage* **15** 870–878.

GLISSON, T. H. (2011). *Introduction to Circuit Analysis and Design*. Springer, The Netherlands.

GLOVER, G. H. (1999). Deconvolution of impulse response in event-related BOLD fMRI. *NeuroImage* **9** 416–429. https://doi.org/10.1006/nimg.1998.0419

GUDBJARTSSON, H. and PATZ, S. (1995). The Rician distribution of noisy MRI data. *Magn. Reson. Med.* **34** 910–914. https://doi.org/10.1002/mrm.1910340618

HAHN, A. D., NENCKA, A. S. and ROWE, D. B. (2009). Improving robustness and reliability of phase-sensitive fMRI analysis using temporal off-resonance alignment of single-echo timeseries (TOAST). *NeuroImage* **44** 742–752. https://doi.org/10.1016/j.neuroimage.2008.10.001

HAJRI, N., YOUSSEF, N. and PATZOLD, M. (2009). A study on the statistical properties of double Hoyt fading channels. In 2009 6*th International Symposium on Wireless Communication Systems* 201–205.

HENKELMAN, R. M. (1985). Measurement of signal intensities in the presence of noise in MR images. *Med. Phys.* **12** 232–233. https://doi.org/10.1118/1.595711

ISLAM, K. T., ZHONG, S., ZAKAVI, P., CHEN, Z., KAVNOUDIAS, H., FARQUHARSON, S., DURBRIDGE, G., BARTH, M., MCMAHON, K. L. et al. (2023). Improving portable low-field MRI image quality through image-to-image translation using paired low- and high-field images. *Sci. Rep.* **13**.

JAIN, A. K. (1989). *Fundaments of Digital Image Processing*. Prentice Hall, New York.

JESMANOWICZ, A., WONG, E. C. and HYDE, J. S. (1993). Phase correction for EPI using internal reference lines. In *Proceedings from the. International Society of Magnetic Resonance in Medicine* **12** 1239.

JEZZARD, P. and CLARE, S. (2001). Principles of nuclear magnetic resonance and MRI. In *Functional MRI: An Introduction to Methods* (P. Jezzard, P. M. Matthews and S. M. Smith, eds.) **3** 67–92. Oxford Univ. Press, London.

KUMAR, A., WELTI, D. and ERNST, R. R. (1975). NMR Fourier Zeugmatography. *J. Magn. Reson.* **18** 69–83.

KWONG, K. K., BELLIVEAU, J. W., CHESLER, D. A., GOLDBERG, I. E., WEISSKOFF, R. M., PONCELET, B. P., KENNEDY, D. N., HOPPEL, B. E., COHEN, M. S. et al. (1992). Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proc. Natl. Acad. Sci. USA* **89** 5675–5679.

LAZAR, N. A. (2008). *The Statistical Analysis of Functional MRI Data*. Springer, Berlin.

LEE, C. C., JACK, C. R. and RIEDERER, S. J. (1998). Mapping of the central sulcus with functional MR: Active versus passive activation tasks. *Am. J. Neuroradiol.* **19** 847–852.

LEE, C. C., WARD, H. A., SHARBROUGH, F. W., MEYER, F. B., MARSH, W. R., RAFFEL, C., SO, E. L., CASCINO, G. D., SHIN, C. et al. (1999). Assessment of functional MR imaging in neurosurgical planning. *Am. J. Neuroradiol.* **20** 1511–1519.

HAHN, A. D., NENCKA, A. S. and ROWE, D. B. (2012). Enhancing the utility of complex-valued functional magnetic resonance imaging detection of neurobiological processes through postacquisition estimation and correction of dynamic B_0 errors and motion. *Hum. Brain Mapp.* **33** 288–306.

LINDSEY, W. (1964). Error probabilities for Rician fading multichannel reception of binary and n-ary signals. *IEEE Trans. Inf. Theory* **10** 339–350.

LIU, Y., LEONG, A. T. L., ZHAO, Y., XIAO, L., MAK, H. K. F., TSANG, A. C. O., LAU, G. K. K., LEUNG, G. K. K. and WU, E. X. (2021). A low-cost and shielding-free ultra-low-field brain MRI scanner. *Nat. Commun.* **12** 1–14.

LJUNGGREN, S. (1983). A simple graphical representation of Fourier-based imaging methods. *J. Magn. Reson.* **54** 338–343.

LOCASCIO, J. J., JENNINGS, P. J., MOORE, C. I. and CORKIN, S. (1997). Time series analysis in the time domain and resampling methods for studies of functional magnetic resonance brain imaging. *Hum. Brain Mapp.* **5** 168–193. https://doi.org/10.1002/(SICI)1097-0193(1997)5:3<168::AID-HBM3>3.0.CO;2-1

LOGAN, B. R. and ROWE, D. B. (2004). An evaluation of thresholding techniques in fMRI analysis. *NeuroImage* **22** 95–108. https://doi.org/10.1016/j.neuroimage.2003.12.047

MACOVSKI, A. (1996). Noise in MRI. *Magn. Reson. Med.* **36** 494–497. https://doi.org/10.1002/mrm.1910360327

MARCHINI, J. L. and RIPLEY, B. D. (2000). A new statistical approach to detecting significant activation in functional MRI. *NeuroImage* **12** 366–380. https://doi.org/10.1006/nimg.2000.0628

MCCLISH, D. K. (1989). Analyzing a portion of the ROC curve. *Med. Decis. Mak.* **9** 190–195. https://doi.org/10.1177/0272989X8900900307

MCLACHLAN, G. J. and KRISHNAN, T. (2008). *The EM Algorithm and Extensions*, 2nd ed. *Wiley Series in Probability and Statistics*. Wiley Interscience, Hoboken, NJ. MR2392878 https://doi.org/10.1002/9780470191613

MEILIJSON, I. (1989). A fast improvement to the EM algorithm on its own terms. *J. Roy. Statist. Soc. Ser. B* **51** 127–138. MR0984999

MENG, X.-L. and RUBIN, D. B. (1993). Maximum likelihood estimation via the ECM algorithm: A general framework. *Biometrika* **80** 267–278. MR1243503 https://doi.org/10.1093/biomet/80.2.267

MILLER, J. W. (1995). Exact maximum likelihood estimation in autoregressive processes. *J. Time Series Anal.* **16** 607–615. MR1366802 https://doi.org/10.1111/j.1467-9892.1995.tb00258.x

MURPHY, K., BODURKA, J. and BANDETTINI, P. A. (2007). How long to scan? The relationship between fMRI temporal signal to noise ratio and necessary scan duration. *NeuroImage* **34** 565–574.

NENCKA, A. S., HAHN, A. D. and ROWE, D. B. (2008). The use of three navigator echoes in Cartesian EPI reconstruction reduces Nyquist ghosting. In *Proceedings from the*. *International Society of Magnetic Resonance in Medicine* **16** 3032.

OGAWA, S., LEE, T. M., NAYAK, A. S. and GLYNN, P. (1990). Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high magnetic fields. *Magn. Reson. Med.* **14** 68–78.

ORCHARD, T. and WOODBURY, M. A. (1972). A missing information principle: Theory and applications. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability* (*Univ. California*, *Berkeley*, *Calif.*, 1970/1971), *Vol. I*: *Theory of Statistics* 697–715. Univ. California Press, Berkeley, CA. MR0400516

POURAHMADI, M. (2001). *Foundations of Time Series Analysis and Prediction Theory*. *Wiley Series in Probability and Statistics*: *Applied Probability and Statistics*. Wiley Interscience, New York. MR1849562

POWERS, J. M., IOACHIM, G. and STROMAN, P. W. (2018). Ten key insights into the use of spinal cord fMRI. *Brain Sci.* **8**. https://doi.org/10.3390/brainsci8090173

RICE, S. O. (1944). Mathematical analysis of random noise. *Bell Syst. Tech. J.* **23** 282–332. MR0010932 https://doi.org/10.1002/j.1538-7305.1944.tb00874.x

ROSEN, B. R. and SAVOY, R. L. (2012). fMRI at 20: Has it changed the world? *NeuroImage* **62** 1316–1324. https://doi.org/10.1016/j.neuroimage.2012.03.004

ROWE, D. B. (2005a). Modeling both the magnitude and phase of complex-valued fMRI data. *NeuroImage* **25** 1310–1324.

ROWE, D. B. (2005b). Parameter estimation in the magnitude-only and complex-valued fMRI data models. *NeuroImage* **25** 1124–1132.

ROWE, D. B. (2016). Handbook of Neuroimaging Data Analysis Image Reconstruction in Functional MRI, 205–232. Chapman & Hall/CRC, London.

ROWE, D. B. and LOGAN, B. R. (2004). A complex way to compute fMRI activation. *NeuroImage* **23** 1078–1092. https://doi.org/10.1016/j.neuroimage.2004.06.042

RUMEAU, C., TZOURIO, N., MURAYAMA, N., PERETTI-VITON, P., LEVRIER, O., JOLIOT, M., MAZOYER, B. and SALAMON, G. (1994). Location of hand function in the sensorimotor cortex: MR and functional correlation. *Am. J. Neuroradiol.* **15** 567–572.

SCHWARZ, G. (1978). Estimating the dimension of a model. *Ann. Statist.* **6** 461–464. MR0468014

SIJBERS, J., DEN DEKKER, A. J., SCHEUNDERS, P. and DYCK, D. V. (1998). Maximum-likelihood estimation of Rician distribution parameters. *IEEE Trans. Med. Imag.* **17** 357–361.

SOLO, V. and NOH, J. (2007). An EM algorithm for Rician fMRI activation detection. In *ISBI* 464-467.

TRISTÁN-VEGA, A., AJA-FERNÁNDEZ, S. and WESTIN, C.-F. (2012). Least squares for diffusion tensor estimation revisited: Propagation of uncertainty with Rician and non-Rician signals. *NeuroImage* **59** 4032–4043.

TWEIG, D. B. (1983). The k-trajectory formulation of the NMR imaging process with applications in analysis and synthesis of imaging methods. *Me. Phys.* **10** 610–21.

WANG, T. and LEI, T. (1994). Statistical analysis of MR imaging and its applications in image modeling. In *Proceedings of the IEEE International Conference on Image Processing and Neural Networks* **1** 866–870.

WANG, Y., HU, Q., LI, L., FOLEY, A. M. and SRINIVASAN, D. (2019). Approaches to wind power curve modeling: A review and discussion. *Renew. Sustain. Energy Rev.* **116** 109422.

WEGMANN, B., EKLUND, A. and VILLANI, M. (2017). Bayesian Rician regression for neuroimaging. *Front. Neurosci.* **11** 586. https://doi.org/10.3389/fnins.2017.00586

WOO, C.-W., KRISHNAN, A. and WAGER, T. D. (2014). Cluster-extent based thresholding in fMRI analyses: Pitfalls and recommendations. *NeuroImage* **91** 412–419. https://doi.org/10.1016/j.neuroimage.2013.12.058

WORSLEY, K. J., MARRETT, S., NEELIN, P., VANDAL, A. C., FRISTON, K. J. and EVANS, A. C. (1996). A unified statistical approach for determining significant voxels in images of cerebral activation. *Hum. Brain Mapp.* **4** 58–73.

YU, C.-H., PRADO, R., OMBAO, H. and ROWE, D. (2018). A Bayesian variable selection approach yields improved detection of brain activation from complex-valued fMRI. *J. Amer. Statist. Assoc.* **113** 1395–1410. MR3902218 https://doi.org/10.1080/01621459.2018.1476244

ZHANG, J. L., RUSINEK, H., CHANDARANA, H. and LEE, V. S. (2013). Functional MRI of the kidneys. *J. Magn. Reson. Imaging* **37** 282–293.

ZHOU, X.-H., OBUCHOWSKI, N. A. and MCCLISH, D. K. (2011). *Statistical Methods in Diagnostic Medicine*, 2nd ed. *Wiley Series in Probability and Statistics*. Wiley, Hoboken, NJ. MR2816760 https://doi.org/10.1002/9780470906514

ZHU, H., LI, Y., IBRAHIM, J. G., SHI, X., AN, H., CHEN, Y., GAO, W., LIN, W., ROWE, D. B. et al. (2009). Regression models for identifying noise sources in magnetic resonance images. *J. Amer. Statist. Assoc.* **104** 623–637. MR2751443 https://doi.org/10.1198/jasa.2009.0029

THE ORIGINAL REFERENCE LIST

The list of entries below corresponds to the original Reference section of your article. The bibliography section on previous page was retrieved from MathSciNet applying an automated procedure.
Please check both lists and indicate those entries which lead to mistaken sources in automatically generated Reference list.

Abdi, A., Tepedelenlioglu, C., Kaveh, M. and Giannakis, G. (2001). On the estimation of the K parameter for the Rice fading distribution. IEEE Communications Letters 5 92-94.

Abramowitz, M. and Stegun, I. (1965). Handbook of Mathematical Functions. Dover Publications.

Adrian, D. W., Maitra, R. and Rowe, D. B. (2013). Ricean over Gaussian modeling in magnitude fMRI analysis – Added complexity with neglible practical benefits. Stat 2 303-316.

Adrian, D. W., Maitra, R. and Rowe, D. B. (2018). Complex-valued time series modeling for improved activation detection in fMRI studies. Annals of Applied Statistics 12 1451-1478.

Adrian, D. W., Maitra, R. and Rowe, D. B. (2025). Supplement to “Rice-distributed autoregressive time series modeling of magnitude functional MRI data”.

Aitkin, M. and Aitkin, I. (1996). A hybrid EM/Gauss-Newton algorithm for maximum likelihood in mixture distributions. Statistics and Computing 6 127-130.

Akaike, H. (1973). 2nd International Symposium on Information Theory Information theory and an extension of the maximum likelihood principle, 267-281. Akademiai Kiado.

Arafa, A. and Messier, G. G. (2010). A Gaussian Model for Dead-Reckoning Mobile Sensor Position Error. In 2010 IEEE 72nd Vehicular Technology Conference—Fall 1-5.

(2024). Improving probabilistic wind speed forecasting using M-Rice distribution and spatial data integration. Applied Energy 360 122840.

Baïle, R., Muzy, J. F. and Poggi, P. An M-Rice wind speed frequency distribution. Wind Energy 14 735-748.

Balchandani, P. and Naidich, T. P. (2015). Ultra-High-Field MR Neuroimaging. American Journal of Neuroradiology 36 1204-1215.

Bandettini, P. A., Jesmanowicz, A., Wong, E. C. and Hyde, J. S. (1993). Processing strategies for time-course data sets in functional MRI of the human brain. Magnetic Resonance in Medicine 30 161-173.

Bar-Shalom, Y., Li, X. R. and Kirubarajan, T. (2001). Estimation with applications to tracking and navigation: theory algorithms and software. John Wiley & Sons.

Belliveau, J. W., Kennedy, D. N., McKinstry, R. C., Buchbinder, B. R., Weisskoff, R. M., Cohen, M. S., Vevea, J. M., Brady, T. J. and Rosen, B. R. (1991). Functional mapping of the human visual cortex by magnetic resonance imaging. Science 254 716-719.

Best, D. J., Rayner, J. . C. W. and Thas, O. (2010). Easily applied tests of fit for the Rayleigh distribution. Sankhya, Series B 72 254–263.

Brown, T. R., Kincaid, B. M. and Ugurbil, K. (1982). NMR chemical shift imaging in three dimensions. In Proceedings of the National Academy of Sciences, USA 79 3523-3526.

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

Carpenter, M. B. (1991). Core Text of Neuroanatomy. Williams and Wilkins.

Casella, G. and Berger, R. L. (2002). Statistical Inference, Second edition ed. Thomson Learning.

Cosottini, M. and Roccatagliata, L. (2021). Neuroimaging at 7 T: are we ready for clinical transition? European Radiology Experimental 5 37.

Dempster, A. P., Laird, N. M. and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. Journal of Royal Statistical Society Series B 23 1-38.

den Dekker, A. J., Poot, D. H. J., Bos, R. and Sijbers, J. (2009). Likelihood-based hypothesis tests for brain activation detection from MRI data disturbed by colored noise: a simulation study. IEEE Transactions on Medical Imaging 28 287-296.

Efron, B. and Tibshirani, R. (1986). Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy. Statistical Science 1 54 – 75.

Friston, K. J., Frith, C. D., Liddle, P. F., Dolan, R. J., Lammertsma, A. A. and Frackowiak, R. S. J. (1990). The relationship between global and local changes in PET scans. Journal of Cerebral Blood Flow and Metabolism 10 458-466.

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J. B., Frith, C. D. and Frackowiak, R. S. J. (1995). Statistical parametric maps in functional imaging: A general linear approach. Human Brain Mapping 2 189-210.

Friston, K. J., Jezzard, P. and Turner, R. (1994). Analysis of functional MRI time-series. Human Brain Mapping 1 153-171.

Friston, K. J., Josephs, O., Zarahn, E., Holmes, A. P., Rouqette, S. and Poline, J. B. (2000). To smooth or not to smooth? Bias and efficiency in fMRI time-series analysis. NeuroImage 12 196-208.

Genovese, C. R., Lazar, N. A. and Nichols, T. E. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate:. NeuroImage 15 870-878.

Glisson, T. H. (2011). Introduction to Circuit Analysis and Design. Springer, The Netherlands.

Glover, G. H. (1999). Deconvolution of impulse response in event-related BOLD fMRI. NeuroImage 9 416-429.

Gudbjartsson, H. and Patz, S. (1995). The Rician distribution of noisy data. Magnetic Resonance in Medicine 34 910-914.

Hahn, A. D., Nencka, A. S. and Rowe, D. B. (2009). Improving robustness and reliability of phase-sensitive fMRI analysis using temporal off-resonance alignment of single-echo timeseries (TOAST). NeuroImage 44 742-752.

Hajri, N., Youssef, N. and Patzold, M. (2009). A study on the statistical properties of double Hoyt fading channels. In 2009 6th International Symposium on Wireless Communication Systems 201-205.

Henkelman, R. M. (1985). Measurement of signal intensities in the presence of noise in MR images. Medical Physics 12 232-233.

Islam, K. T., Zhong, S., Zakavi, P., Chen, Z., Kavnoudias, H., Farquharson, S., Durbridge, G., Barth, M., McMahon, K. L., Parizel, P. M., Dwyer, A., Egan, G. F., Law, M. and Chen, Z. (2023). Improving portable low-field MRI image quality through image-to-image translation using paired low- and high-field images. Scientific Reports 13.

Jain, A. K. (1989). Fundaments of Digital Image Processing. Prentice Hall.

Jesmanowicz, A., Wong, E. C. and Hyde, J. S. (1993). Phase Correction for EPI Using Internal Reference Lines. In Proceedings from the International Society of Magnetic Resonance in Medicine 12 1239.

Jezzard, P. and Clare, S. (2001). Principles of nuclear magnetic resonance and MRI. In Functional MRI: An Introduction to Methods (P. Jezzard, P. M. Matthews and S. M. Smith, eds.) 3, 67-92. Oxford University Press.

Kumar, A., Welti, D. and Ernst, R. R. (1975). NMR Fourier Zeugmatography. Journal of Magnetic Resonance 18 69-83.

Kwong, K. K., Belliveau, J. W., Chesler, D. A., Goldberg, I. E., Weisskoff, R. M., Poncelet, B. P., Kennedy, D. N., Hoppel, B. E., Cohen, M. S., Turner, R., Cheng, H.-M., Brady, T. J. and Rosen, B. R. (1992). Dynamic Magnetic Resonance Imaging of Human Brain Activity During Primary Sensory Stimulation. Proceedings of the National Academy of Sciences of the United States of America 89 5675-5679.

Lazar, N. A. (2008). The Statistical Analysis of Functional MRI Data. Springer.

Lee, C. C., Jack, C. R. and Riederer, S. J. (1998). Mapping of the Central Sulcus with Functional MR: Active versus Passive Activation Tasks. American Journal of Neuroradiology 19 847-852.

Lee, C. C., Ward, H. A., Sharbrough, F. W., Meyer, F. B., Marsh, W. R., Raffel, C., So, E. L., Cascino, G. D., Shin, C., Xu, Y., Riederer, S. J. and Jack, C. R. (1999). Assessment of Functional MR Imaging in Neurosurgical Planning. American Journal of Neuroradiology 20 1511-1519.

Hahn, A. D., Nencka, A. S. and Rowe, D. B. (2012). Enhancing the Utility of Complex-Valued Functional Magnetic Resonance Imaging Detection of Neurobiological Processes Through Postacquisition Estimation and Correction of Dynamic B_0 Errors and Motion. Human Brain Mapping 33 288-306.

Lindsey, W. (1964). Error probabilities for Rician fading multichannel reception of binary and n-ary signals. IEEE Transactions on Information Theory 10 339-350.

Liu, Y., Leong, A. T. L., Zhao, Y., Xiao, L., Mak, H. K. F., Tsang, A. C. O., Lau, G. K. K., Leung, G. K. K. and Wu, E. X. (2021). A low-cost and shielding-free ultra-low-field brain MRI scanner. Nature Communications 12 1-14.

Ljunggren, S. (1983). A Simple Graphical Representation of Fourier-Based Imaging Methods. Journal of Magnetic Resonance 54 338-343.

Locascio, J. J., Jennings, P. J., Moore, C. I. 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.

Logan, B. R. and Rowe, D. B. (2004). An evaluation of thresholding techniques in fMRI analysis. NeuroImage 22 95-108.

Macovski, A. (1996). Noise in MRI. Magnetic Resonance in Medicine 36 494-497.

Marchini, J. L. and Ripley, B. D. (2000). A new statistical approach to detecting significant activation in functional MRI. NeuroImage 12 366-380.

McClish, D. K. (1989). Analyzing a Portion of the ROC Curve. Medical Decision Making 9 190-195.

McLachlan, G. J. and Krishnan, T. (2008). The EM Algorithm and Extensions. Wiley.

Meilijson, I. (1989). A fast improvement to the EM algorithm on its own terms. Journal of the Royal Statistical Society. Series B (Methodological) 51 127-138.

Meng, X.-L. and Rubin, D. B. (1993). Maximum likelihood estimation via the ECM algorithm: a general framework. Biometrika 80 267-278.

Miller, J. W. (1995). Exact Maximum Likelihood Estimation in Autoregressive Processes. Journal of Time Series Analysis 16 607-615.

Murphy, K., Bodurka, J. and Bandettini, P. A. (2007). How long to scan? The relationship between fMRI temporal signal to noise ratio and necessary scan duration. Neuroimage 34 565-574.

Nencka, A. S., Hahn, A. D. and Rowe, D. B. (2008). The Use of Three Navigator Echoes in Cartesian EPI Reconstruction Reduces Nyquist Ghosting. In Proceedings from the International Society of Magnetic Resonance in Medicine 16 3032.

Ogawa, S., Lee, T. M., Nayak, A. S. and Glynn, P. (1990). Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high magnetic fields. Magnetic Resonance in Medicine 14 68-78.

Orchard, T. and Woodbury, M. A. (1972). A missing information principle: theory and applications. In Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability 1 697-715. University of California Press, Berkeley, California.

Pourahmadi, M. (2001). Foundations of Time Series Analysis and Prediction Theory. Wiley.

Powers, J. M., Ioachim, G. and Stroman, P. W. (2018). Ten Key Insights into the Use of Spinal Cord fMRI. Brain Sciences 8 173.

Rice, S. O. (1944). Mathematical analysis of random noise. Bell Systems Technical Journal 23 282.

Rosen, B. R. and Savoy, R. L. (2012). fMRI at 20: Has it changed the world? NeuroImage.

Rowe, D. B. (2005a). Modeling both the magnitude and phase of complex-valued fMRI data. NeuroImage 25 1310-1324.

Rowe, D. B. (2005b). Parameter estimation in the magnitude-only and complex-valued fMRI data models. NeuroImage 25 1124-1132.

Rowe, D. B. (2016). Handbook of Neuroimaging Data Analysis Image Reconstruction in Functional MRI, 205-232. Chapman & Hall/CRC.

Rowe, D. B. and Logan, B. R. (2004). A complex way to compute fMRI activation. NeuroImage 23 1078-1092.

Rumeau, C., Tzourio, N., Murayama, N., Peretti-Viton, P., Levrier, O., Joliot, M., Mazoyer, B. and Salamon, G. (1994). Location of Hand Function in the Sensorimotor Cortex: MR and Functional Correlation. American Journal of Neuroradiology 15 567-572.

Schwarz, G. E. (1978). Estimating the dimension of a model. Annals of Statistics 6 461-464.

Sijbers, J., den Dekker, A. J., Scheunders, P. and Dyck, D. V. (1998). Maximum-Likelihood Estimation of Rician Distribution Parameters. IEEE Transactions on Medical Imaging 17 357-361.

Solo, V. and Noh, J. (2007). An EM algorithm for Rician fMRI activation detection. In ISBI 464-467.

Tristán-Vega, A., Aja-Fernández, S. and Westin, C.-F. (2012). Least squares for diffusion tensor estimation revisited: Propagation of uncertainty with Rician and non-Rician signals. NeuroImage 59 4032-4043. Copyright—Copyright Elsevier Limited Feb 15, 2012; Last updated—2016-11-01.

Tweig, D. B. (1983). The k-trajectory formulation of the NMR imaging process with applications in analysis and synthesis of imaging methods. Medical Physics 10 610-21.

Wang, T. and Lei, T. (1994). Statistical analysis of MR imaging and its applications in image modeling. Proceedings of the IEEE International Conference on Image Processing and Neural Networks 1 866-870.

Wang, Y., Hu, Q., Li, L., Foley, A. M. and Srinivasan, D. (2019). Approaches to wind power curve modeling: A review and discussion. Renewable and Sustainable Energy Reviews 116 109422.

Wegmann, B., Eklund, A. and Villani, M. (2017). Bayesian Rician Regression for Neuroimaging. Frontiers in Neuroscience 11.

Woo, C.-W., Krishnan, A. and Wager, T. D. (2014). Cluster-extent based thresholding in fMRI analyses. NeuroImage 91 412-419.

Worsley, K. J., Marrett, S., Neelin, P., Vandal, A. C., Friston, K. J. and Evans, A. C. (1996). A unified statistical approach for determining significant voxels in images of cerebral activation. Human Brain Mapping 4 58-73.

Yu, C.-H., Prado, R., Ombao, H. and Rowe, D. (2018). A Bayesian Variable Selection Approach Yields Improved Detection of Brain Activation from Complex-Valued fMRI. Journal of the American Statistical Association 113 1395-1410.

Zhang, J. L., Rusinek, H., Chandarana, H. and Lee, V. S. (2013). Functional MRI of the kidneys. Journal of Magnetic Resonance Imaging 37 282-293.

Zhou, X. H., Obuchowski, N. A. and McClish, D. K. (2011). Statistical methods in diagnostic medicine, Second edition ed. Wiley.

Zhu, H., Li, Y., Ibrahim, J. G., Shi, X., An, H., Chen, Y., Gao, W., Lin, W., Rowe, D. B. and Peterson, B. S. (2009). Regression Models for Identifying Noise Sources in Magnetic Resonance Images. Journal of the American Statistical Association 104 623-637.

META DATA IN THE PDF FILE

Following information will be included as pdf file Document Properties:

```
Title   : Rice-distributed autoregressive time series modeling of magnitude
  functional MRI data
Author  : Daniel W. Adrian, Rajan Maitra, Daniel B. Rowe
Subject : The Annals of Applied Statistics, 0, Vol. 0, No. 00, 1-20
Keywords: Bilateral finger-tapping motor experiment, EM algorithm, empirical
  information matrix, hemodynamic response function, Monte Carlo integration,
  Rice distribution, signal-to-noise ratio
```

THE LIST OF URI ADDRESSES

Listed below are all uri addresses found in your paper. The non-active uri addresses, if any, are indicated as ERROR. Please check and update the list where necessary. The e-mail addresses are not checked – they are listed just for your information. More information can be found in the support page: http://www.e-publications.org/ims/support/urihelp.html.

```
200 https://imstat.org/journals-and-publications/annals-of-applied-
    statistics/ [1:p.1] OK
301 https://www.imstat.org [1:p.1] Moved Permanently //
    https://www.imstat.org/
--- mailto:adriad1@gvsu.edu [1:p.1] Check skip
--- mailto:maitra@iastate.edu [1:p.1] Check skip
--- mailto:daniel.rowe@marquette.edu [1:p.1] Check skip
200 https://github.com/dadrian14/arp-rice-data [1:p.5] OK
404 https://doi.org/10.1214/24-AOAS1981SUPPA [2:pp.17,17] Not Found
404 https://doi.org/10.1214/24-AOAS1981SUPPB [3:pp.17,17,17] Not Found
```