An Efficient Fully Bayesian Approach to Brain Activity Mapping with Complex-Valued fMRI Data

Zhengxin Wang[1], Daniel B. Rowe[2], Xinyi Li[1], D. Andrew Brown[1]

JSM 2023 Toronto

Supported by United States National Science Foundation Grant DMS 2210686

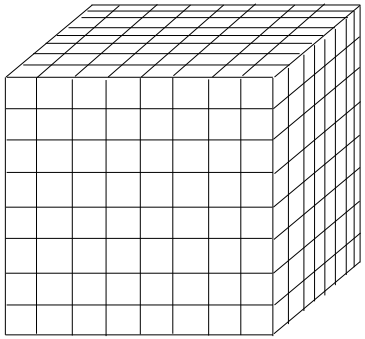
[1] School of Mathematical and Statistical Sciences, Clemson University, Clemson, SC, USA
[2] Department of Mathematical and Statistical Sciences, Marquette University, Milwaukee, WI, USA

CLEMSON
School of MATHEMATICAL AND STATISTICAL SCIENCES

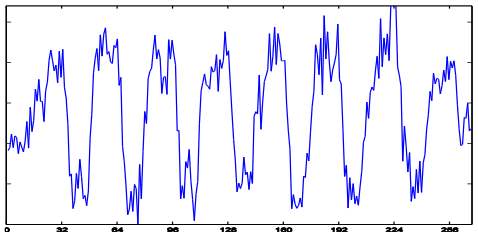
NSF

Functional MRI

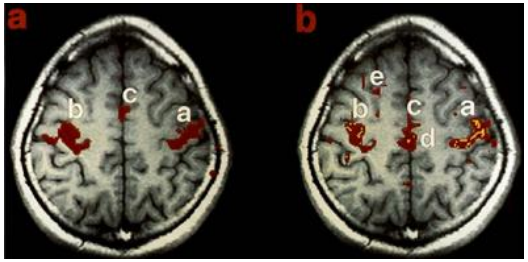
Voxels



Time Course

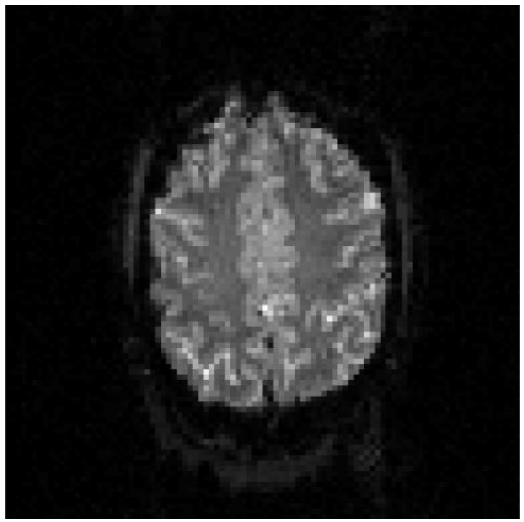


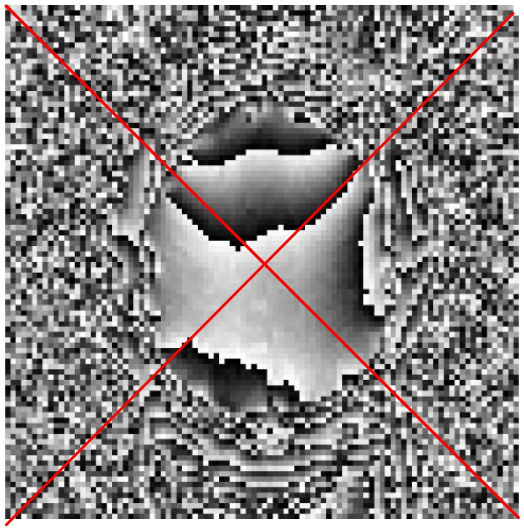
Activation or Connectivity



Courtesy of D. Rowe

- Raw signals are collected in spatial frequency space (Fourier, k-space)
- After inverse Fourier Transform, imperfect reconstruction yields complex-valued image
- Different reconstruction procedures (e.g., solving the inverse problem) can affect sensitivity to detect activation





Phase discarded!
(in nearly all fMRI)

Courtesy of D. Rowe

Brown et al. (2014); Sakitis, **B**, Rowe (2023+)

- Typically only the magnitude component of the data are retained, discarding phase
- Discarding phase can lead to inappropriate distributional assumptions in the statistical model (e.g., Gaussian vs. Ricean), and sacrifice power in the presence of low SNR
- Some information about neuronal activation may be contained in phase changes, not captured in magnitude alone
- Previous work has shown using and modeling both components can yield improved activation detection

Rowe and Logan (2004, 2005), Rowe (2009), Adrian et al. (2018), Yu et al. (2018), . . .

Contributions

- Currently, there are relatively few statistical approaches for complex-valued fMRI
  - Even fewer Bayesian approaches
- Bayesian approach offers flexibility / interpretability in capturing spatial/temporal associations in cv-fMRI
- **Our Contribution**: A Bayesian approach that
  1. Models full cv-fMRI data (not just magnitude)
  2. Captures spatial/temporal association
  3. Facilitates compuationally-efficient MCMC-based inference

Woolrich et al. (2004), Smith and Fahrmeir (2007), Musgrove et al. (2016), Bezener et al. (2018), Yu et al. (2023)

Model Formulation

- For each voxel v, we suppose that

$$\mathbf{y}^v = \mathbf{x}\beta^v + \mathbf{r}^v\rho^v + \boldsymbol{\varepsilon}^v; \;\; v = 1, \ldots, V$$

  where all terms are $\mathbb{C}$-valued except for $\mathbf{x} \in \mathbb{R}^T$.

- $\boldsymbol{\varepsilon}^v \sim \mathcal{CN}_T(\boldsymbol{\mu}^v = \mathbf{0}, \boldsymbol{\Gamma}^v = 2\sigma_v^2\mathbf{I}, \mathbf{C}^v = \mathbf{0})$.

- $\mathbf{r}^v \in \mathbb{C}^T$ is a vector of lag-1 prediction errors for AR(1) errors.

- Equivalent (easier?) representation:

$$\underbrace{\begin{pmatrix} \mathbf{y}^v_{Re} \\ \mathbf{y}^v_{Im} \end{pmatrix}}_{\mathbf{y}^v_r} = \underbrace{\begin{pmatrix} \mathbf{x} & \mathbf{0} \\ \mathbf{0} & \mathbf{x} \end{pmatrix}}_{\mathbf{X}_r} \underbrace{\begin{pmatrix} \beta^v_{Re} \\ \beta^v_{Im} \end{pmatrix}}_{\boldsymbol{\beta}^v_r} + \underbrace{\begin{pmatrix} \mathbf{r}^v_{Re} & -\mathbf{r}^v_{Im} \\ \mathbf{r}^v_{Im} & \mathbf{r}^v_{Re} \end{pmatrix}}_{\mathbf{R}^v_r} \underbrace{\begin{pmatrix} \rho^v_{Re} \\ \rho^v_{Im} \end{pmatrix}}_{\boldsymbol{\rho}^v_r} + \underbrace{\begin{pmatrix} \boldsymbol{\varepsilon}^v_{Re} \\ \boldsymbol{\varepsilon}^v_{Im} \end{pmatrix}}_{\boldsymbol{\varepsilon}^v_r}$$

with

$$\boldsymbol{\varepsilon}^v_r \sim \mathcal{N}_{2T}(\mathbf{0}, \boldsymbol{\Sigma}^v)$$

and

$$\boldsymbol{\Sigma}^v = \begin{pmatrix} \boldsymbol{\Sigma}^v_{Re,Re} & \boldsymbol{\Sigma}^v_{Re,Im} \\ \boldsymbol{\Sigma}^v_{Im,Re} & \boldsymbol{\Sigma}^v_{Im,Im} \end{pmatrix}.$$

Rowe and Logan (2004), Lee et al. (2007), Rowe (2009)

Brain Partitioning

- Musgrove et al. (2016) partition brain voxels into parcels, facilitating parcel-wise (magnitude-only) estimation/inference **in parallel**.
- Different strategies can be used to determine partitions; e.g:
  - Parcels with equal numbers of voxels
  - Partitions based on anatomical ROIs
- In this work, we partition 2- or 3D images into G parcels of approximately the same (geometric) size.
- Empirical work suggests that, regardless of partition strategy, edge effects within each parcel are minimal

Amunts et al. (2000) Tzourio-Mazoyer et al. (2002), Musgrove et al. (2016), Wang ... **B** (2023+)

Identifying Neuronal Activation

- Under the assumed model, voxel v ($v = 1, \ldots, V_g$) is classified as an active voxel under the stimulus if its complex-valued regression coefficient

$$\beta^v = \beta^v_{Re} + i\beta^v_{Im} \neq 0$$

.

- $\Rightarrow$ complex-valued variable selection problem. We take the "spike-and-slab" approach:

$$\beta^v \mid \gamma_v \sim \gamma_v \mathcal{CN}_1(0, 2\tau_g^2, 0) + (1 - \gamma_v)\mathcal{I}_0$$

- Real-valued equivalent:

$$\boldsymbol{\beta}^v_r = \begin{pmatrix} \beta^v_{Re} \\ \beta^v_{Im} \end{pmatrix} \mid \gamma_v \sim \mathcal{N}_2\left(\mathbf{0},\ \gamma_v \tau_g^2 \mathbf{I}\right)$$

Mitchell and Beauchamp (1988), Yu et al. (2018)

Spatial Dependence

- Spatial association may arise from several sources:
  - Noise structure of the data, unmodeled neuronal activities, spatial normalization, image reconstruction, and spatial smoothing, etc..

- Approach: sparse spatial generalized linear mixed model (sSGLMM):

$$\begin{aligned}
\gamma_v \mid \eta_v &\overset{iid}{\sim} \mathcal{B}ern\left\{\mathbf{\Phi}(\psi + \eta_v)\right\} \\
\eta_v \mid \boldsymbol{\delta}_g &\sim \mathcal{N}_1\left(\mathbf{m}_v'\boldsymbol{\delta}_g, 1\right) \\
\boldsymbol{\delta}_g \mid \kappa_g &\sim \mathcal{N}_q\left\{\mathbf{0}, (\kappa_g \mathbf{M}_g{}'\mathbf{Q}_g\mathbf{M}_g)^{-1}\right\} \\
\kappa_g &\sim \mathcal{G}amma\left(a_\kappa, b_\kappa\right)
\end{aligned}$$

where ψ is a fixed tuning parameter to control sparsity.

Friston et al. (1995), Krüger and Glover (2001), Reich et al. (2006), Mikl et al. (2008), Bianciardi et al. (2009), Rowe et al. (2009), Hughes and Haran (2013), Musgrove et al. (2016)

Graph-based Dimension Reduction

- Defining $\mathbf{A}_g \in \{0,1\}^{V_g \times V_g}$ to be the adjacency matrix associated with the assumed Markov graph, $\mathbf{M}_g \in \mathbb{R}^{V_g \times q}$ comprises the q principal eigenvectors of $\mathbf{A}_g$
  - Typically with $q \ll V_g$
- $\mathbf{m}'_v$ is a $1 \times q$ row vector of "synthetic" spatial predictors corresponding to the vth row of $\mathbf{M}_g$.
- $\mathbf{Q}_g = \text{diag}(\mathbf{A}_g \mathbf{1}_{V_g}) - \mathbf{A}_g$ is the graph Laplacian.
- This approach essentially projects spatial effects onto orthogonal spaces to correct for confounding, and captures spatial trends at multiple resolutions

Reich et al. (2006), Hughes and Haran (2013)

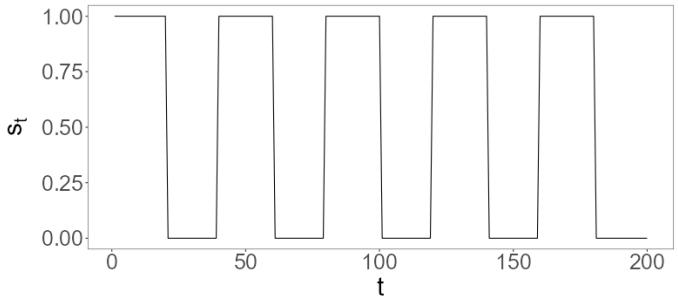
Numerical Experiments

- We simulate two-dimensional complex-valued time series of fMRI signals three ways:
  1. Data with non-AR noise,
  2. Data with AR(1) noise
  3. More realistic simulated non-AR data, imitating the human brain

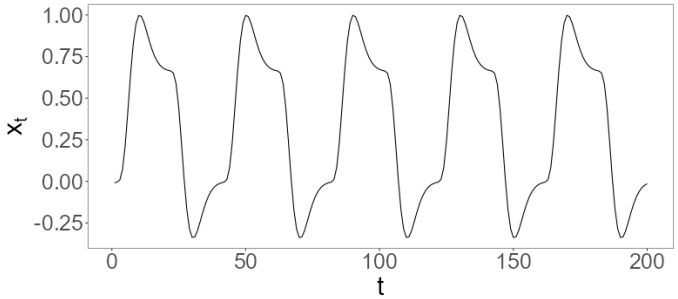
- Comparisons:
  - Musgove et al. (2016): magnitude-only, partitioned approach
  - Yu et al. (2018): complex-valued Bayesian variable selection, no spatial effect and no partitioning
  - Our proposed approach: complex-valued, partitioned, spatial (GMRF) effects

- Parameters / thresholds are tuned separately for each model to yield optimal performance

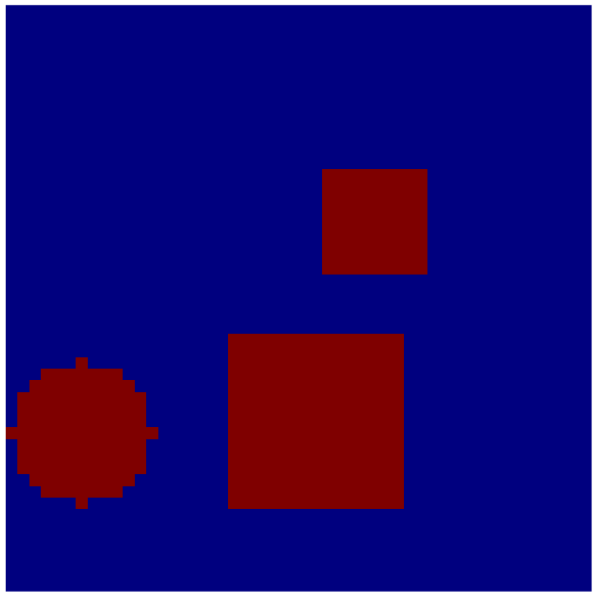
Simulated stimulus & BOLD response



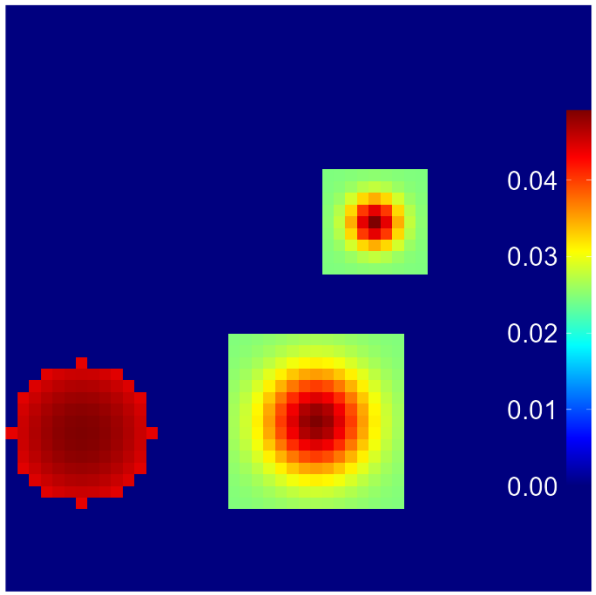
(a) Designed stimulus



(b) Expected BOLD response



(c) True activation map



(d) True magnitude map

Welvaert et al. (2011); Wang, Rowe, Li, and **B** (2023+)

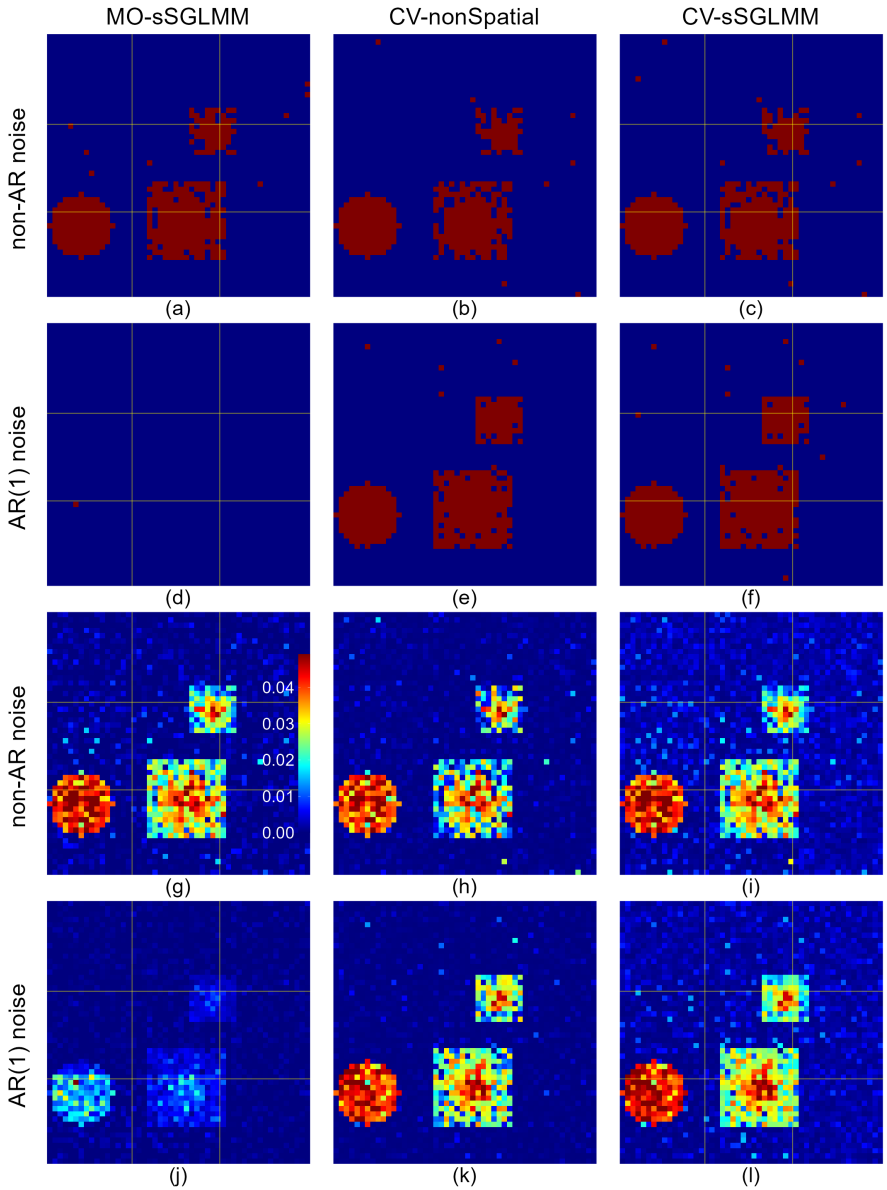
- We consider both IID and (time) auto-regressive errors.
- For the AR(1) simulation, we use

$$\begin{pmatrix} \varepsilon^v_{t,Re} \\ \varepsilon^v_{t,Im} \end{pmatrix} = \begin{pmatrix} 0.2 & -0.9 \\ 0.9 & 0.2 \end{pmatrix} \begin{pmatrix} \varepsilon^v_{t-1,Re} \\ \varepsilon^v_{t-1,Im} \end{pmatrix} + \begin{pmatrix} \xi^v_{Re} \\ \xi^v_{Im} \end{pmatrix}, \quad \begin{pmatrix} \xi^v_{Re} \\ \xi^v_{Im} \end{pmatrix} \sim \mathcal{N}_2\left(\mathbf{0}, \sigma^2 \mathbf{I}\right)$$

  which is the real-valued isomorphism of the complex-valued AR(1) errors:

$$\varepsilon^v_t = (0.2 + 0.9i)\varepsilon^v_{t-1} + \xi_v, \quad \xi_v \sim \mathcal{CN}_1(0, 2\sigma^2, 0)$$

Welvaert et al. (2011); Wang, Rowe, Li, and **B** (2023+)


MO-sSGLMM
CV-nonSpatial
CV-sSGLMM
non-AR noise
AR(1) noise
non-AR noise
AR(1) noise
0.04
0.03
0.02
0.01
0.00
(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
(j)
(k)
(l)

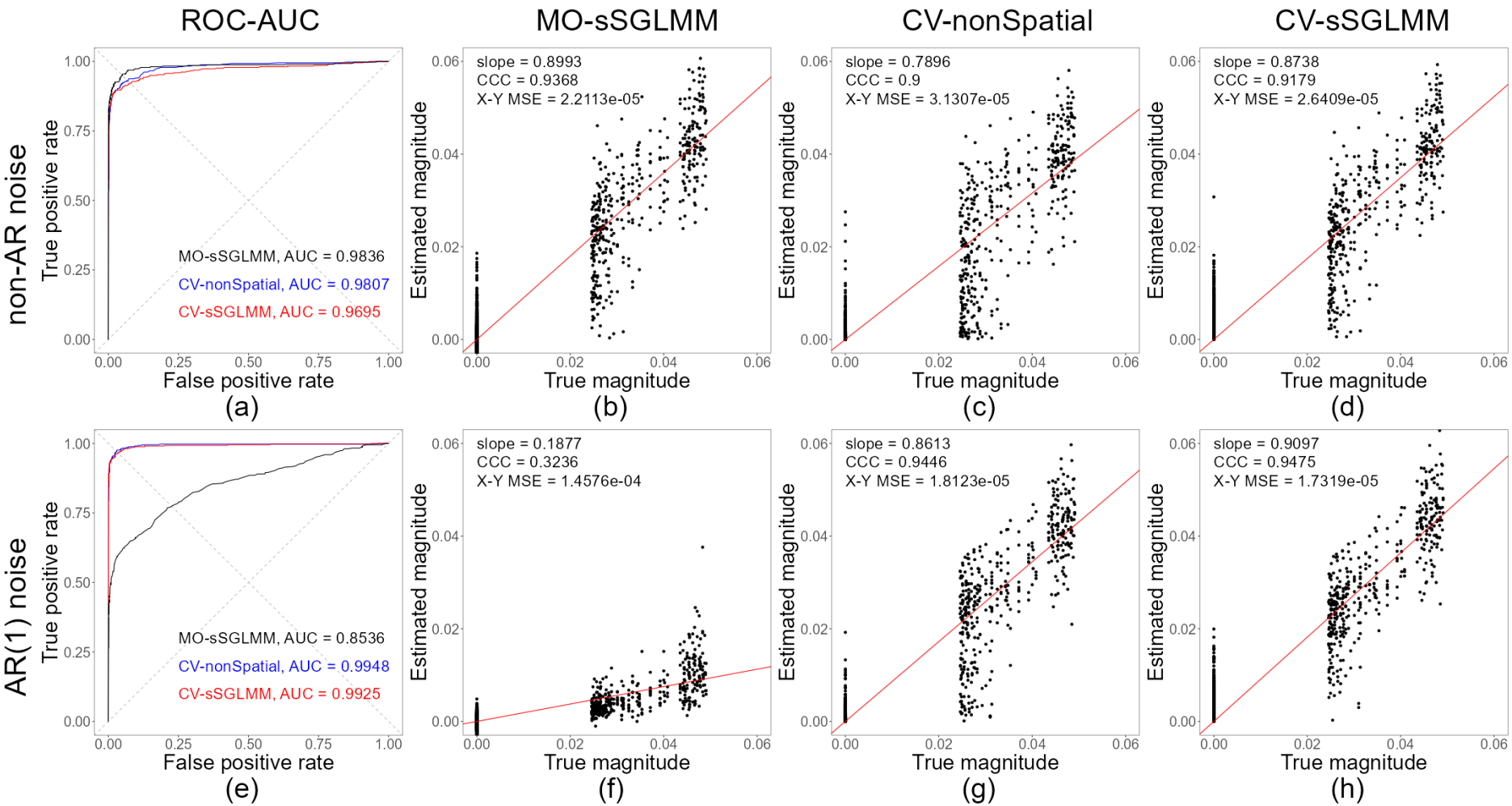


Figure: (a)-(d) are the ROC curve and plots comparing true versus estimated magnitudes for a non-AR dataset. (e)-(h) are analogous plots for an AR(1) dataset.

Performance Summary over Replications

AR type	Mode	Accuracy	Precision	Recall	F1 Score	AUC	Slope	CCC	X-Y MSE	Time (s)
non-AR	MO-sSGLMM	**0.9693**	0.9440	**0.8160**	**0.8741**	**0.9774**	**0.8586**	**0.9008**	**2.06e-5**	**2.4**
	CV-nonSpatial	0.9540	**0.9632**	0.6687	0.7853	0.9751	0.6771	0.8222	3.04e-5	41.9
	CV-sSGLMM	0.9622	0.9277	0.7742	0.8424	0.9625	0.8186	0.8627	2.54e-5	5.51
AR(1)	CV-nonSpatial	0.9765	**0.9733**	0.8407	0.9012	**0.9927**	0.8040	0.9096	1.69e-5	42.2
	CV-sSGLMM	**0.9797**	0.9381	**0.9039**	**0.9201**	0.9879	**0.8816**	**0.9145**	**1.60e-5**	**5.39**

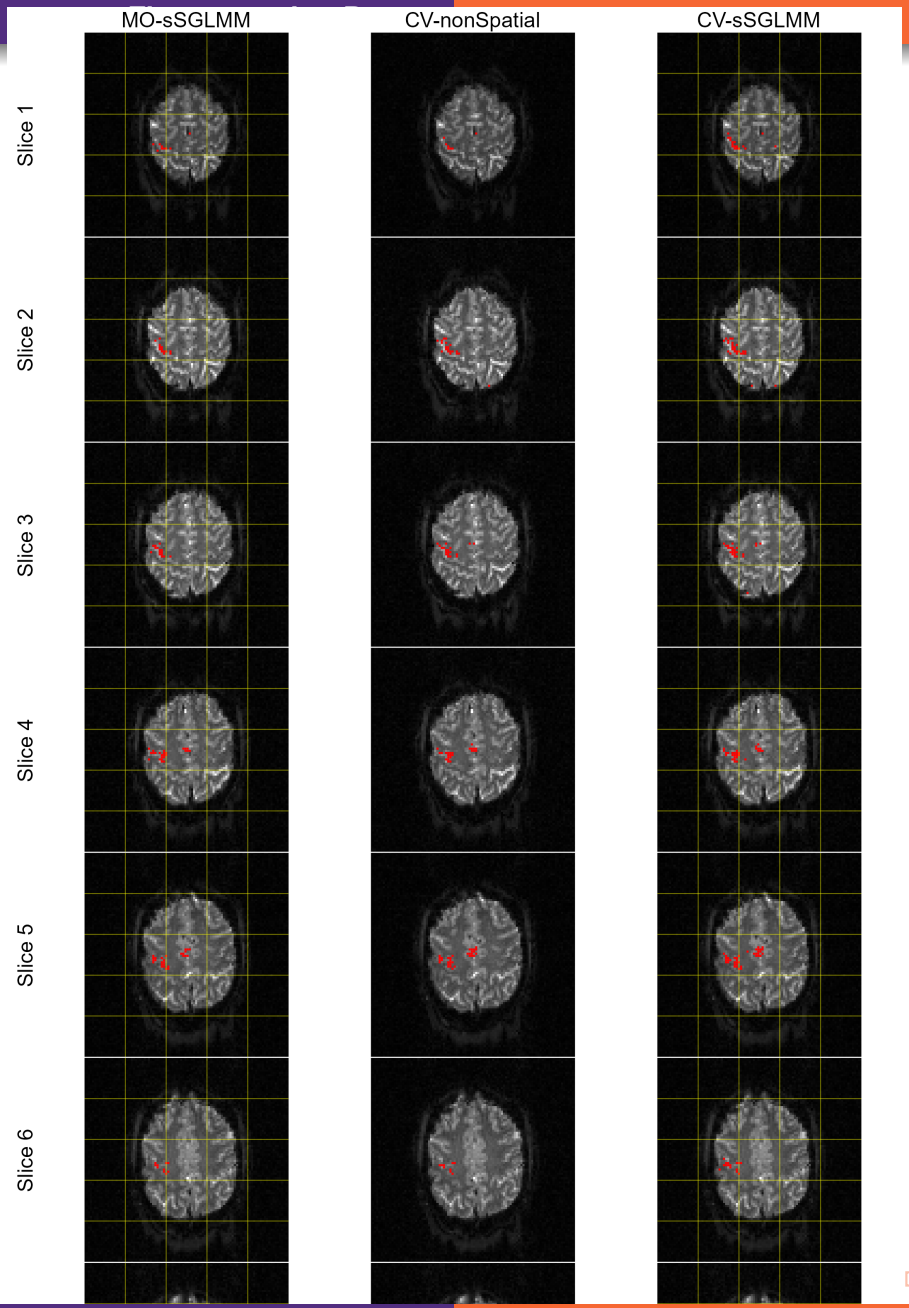
Table: Summary of average metrics across 100 non-AR and 100 AR(1) datasets produced by the MO-sSGLMM, CV-nonSpatial, and CV-sSGLMM models

Wang, Rowe, Li, and **B** (2023+)

Human cv-fMRI Data

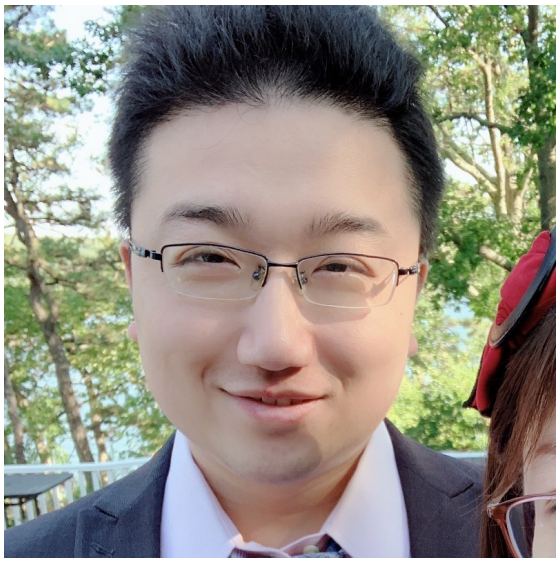
- We use the fMRI dataset analyzed by Yu et al. (2018), which was acquired during a unilateral finger-tapping experiment on a 3.0-T General Electric Signa LX MRI scanner.
- 16 epochs of alternating 15s on and 15s off periods $\Rightarrow T = 490$ (including warming-up period).
- For spatially-partitioned approaches, $G = 25$ partitions and a posterior probability threshold of 0.8722

Smith and Fahrmeir (2007); Yu et al. (2018); Wang, Rowe, Li, and **B** (2023+)


MO-sSGLMM
CV-nonSpatial
CV-sSGLMM
Slice 1
Slice 2
Slice 3
Slice 4
Slice 5
Slice 6

Collaborators

Zhengxin (Jason) Wang
(Clemson U. SMSS)



Daniel Rowe
(Marquette U. DMSS)

Xinyi Li
(Clemson U. SMSS)

References

- Wang, Z., Rowe, D. B., Li, X., and Brown, D. A. (2023+), “Efficient fully Bayesian approach to brain activity mapping with complex-valued fMRI data,” in preparation.
- Sakitis, C., Brown, D. A., and Rowe, D. B. (2023+), “A Formal Bayesian Approach to SENSE Image Reconstruction Produces Increased Accuracy in Detection of Task Activation,” submitted.

Thank you!