Bayesian Inference for Brain Activity from Dual-Resolution Functional Magnetic Resonance Imaging

Andrew Whiteman, Jian Kang, & Timothy D. Johnson

Department of Biostatistics, University of Michigan, Ann Arbor

BIOSTATISTICS

Introduction

- **Personalized medicine:** need for patient-specific functional mapping. Surgeons want to minimize damage to healthy, functional tissue
  - Variations in functional neuroanatomy [e.g. 1]
  - MR imaging methods used to aid presurgical planning and neuronavigation

Why multiple resolutions?

- Surgery requires high spatial precision
- Signal-to-noise ratio (SNR) decreases as spatial resolution increases

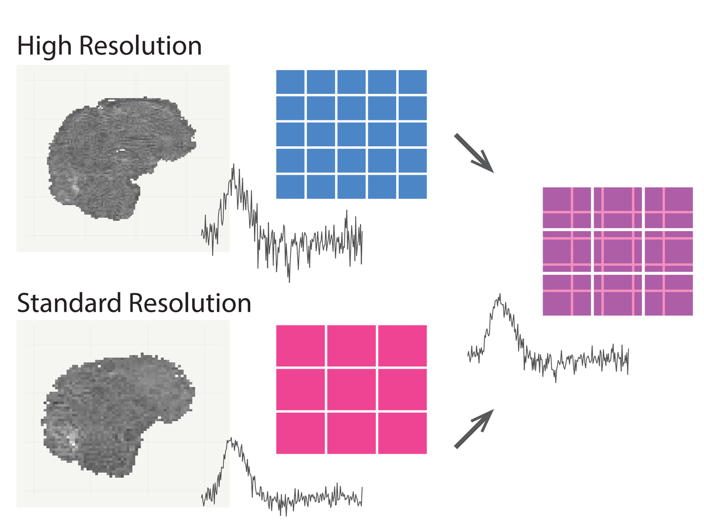


Fig 1: Unique blessing/curse of data collected at "high" and "standard" resolutions ($1.8 \times 1.8 \times 2.3$ mm^3 and $3 \times 3 \times 3.45$ mm^3 voxels; voxel–volumetric pixel)

Goal: to leverage spatial precision from high resolution data and SNR from standard resolution data

- Inference at high resolution voxel locations
- Utilize massive spatial information to conduct within-patient functional mapping
  - Fully Bayesian inference in data of this size can be computationally prohibitive

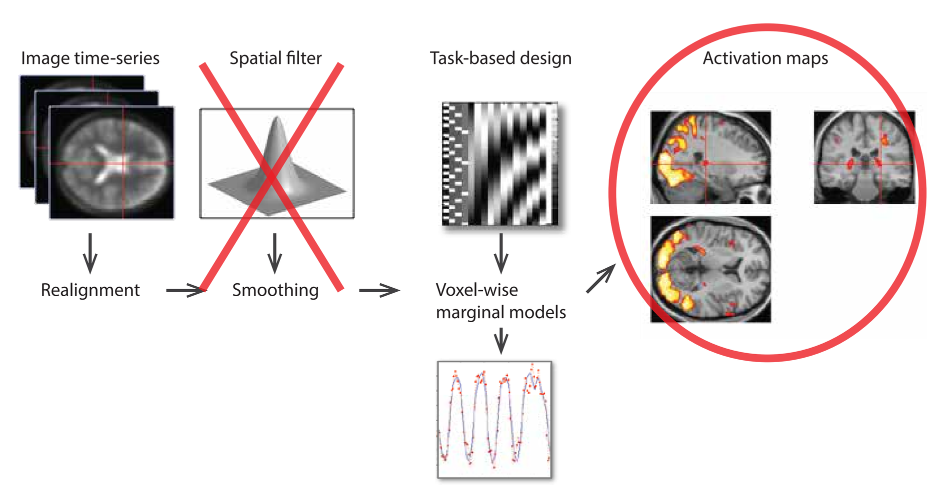


Fig 2: Typical fMRI pipeline. We received and used unsmoothed z-statistic maps as data, treating these as noisy summaries of spatial activation. Adapted from: https://www.fil.ion.ucl.ac.uk/spm/course/slides11/

Proposed Model for Dual-Resolution fMRI

$$y_h(\boldsymbol{v}_h) \sim \mathcal{N}(\mu(\boldsymbol{v}_h), \sigma_h^2) \qquad y_s(\boldsymbol{v}_s) \sim \mathcal{N}(\mu(\boldsymbol{v}_s), \sigma_s^2)$$
$$\pi(\sigma_h^2, \sigma_s^2) \propto \sigma_h^{-2}\sigma_s^{-2}\,\mathbf{1}(\sigma_h^{-2} < \sigma_s^{-2})$$

(h—high and s—standard resolution)

- Assume different images $\boldsymbol{y}_h$ and $\boldsymbol{y}_s$ are realizations of the same spatial activation process
- Independent noise: no smoothing during preprocessing

$$\mu(\boldsymbol{v}) \sim \mathcal{GP}(0, K) \qquad K(\boldsymbol{v}, \boldsymbol{v}') = \tau^2 \exp(-\psi \|\boldsymbol{v} - \boldsymbol{v}'\|^\nu), \quad \tau^2, \psi > 0; \nu \in (0, 2]$$

- Correlation between voxels and images induced by Gaussian Process prior on $\mu(\cdot)$
- Covariance parameters estimated via minimum contrast

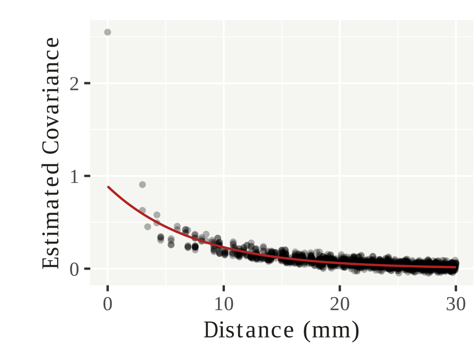
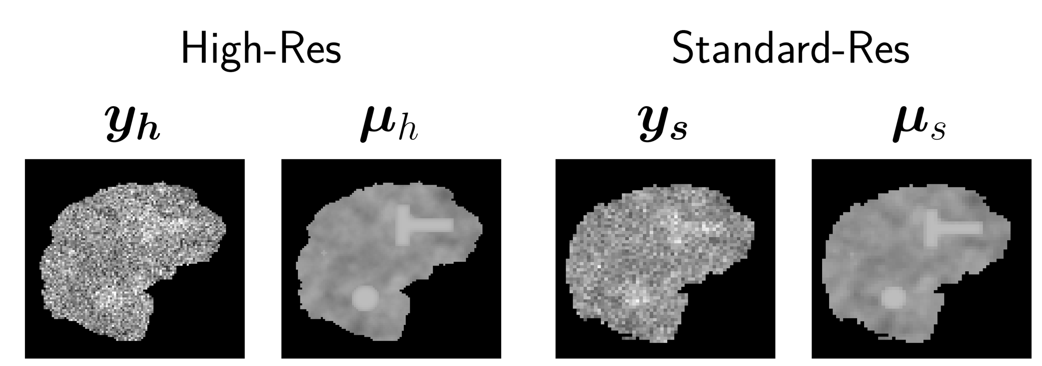


Fig 3: Empirical covariogram and fitted curve with exponential correlation function

Illustrative Example: Embedded Signal in 2D Slices



- Compare with single resolution methods
- **Naive** data combination: $\bar{\boldsymbol{y}}_{hs} = (\boldsymbol{y}_h + \boldsymbol{W}^\top \boldsymbol{y}_s)/2$

(Where $\boldsymbol{W}^\top \boldsymbol{y}_s$ is an ordinary kriging interpolation of $\boldsymbol{y}_s$)

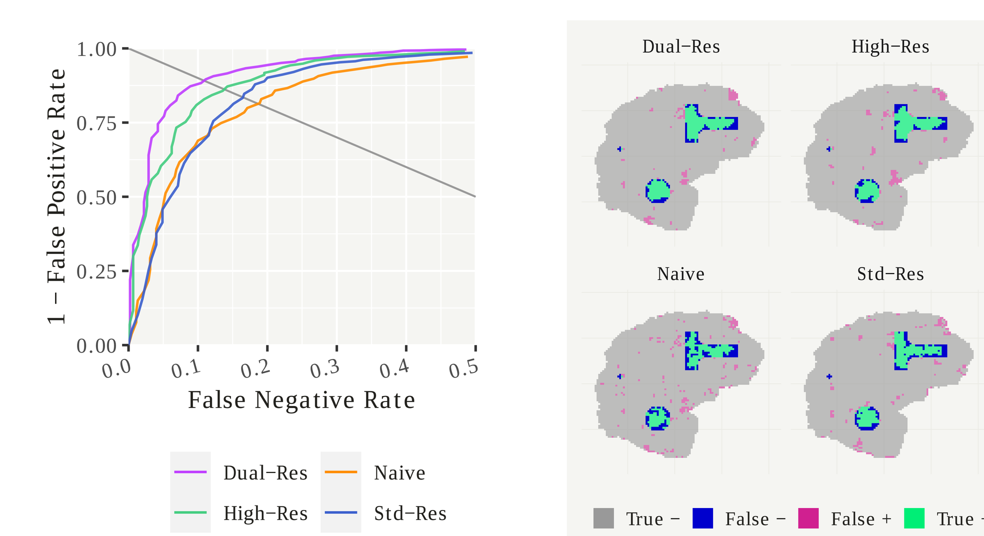


Fig 4: (*Left*) ROC curves show proposed method's improvement over single resolution methods. (*Right*) Panels show example inferential errors

Patient Data Analysis

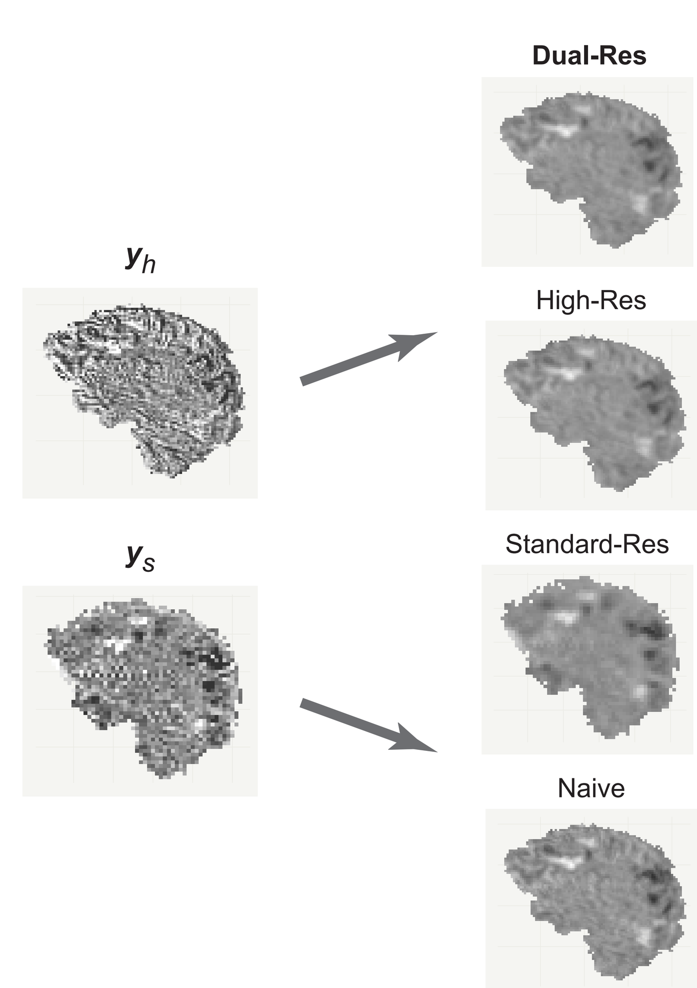


Fig 5: Representative slices from whole brain patient data. (*Left*) raw, unsmoothed data (z-statistic maps); (*right*) qualitative comparison of posterior mean images for the same slices

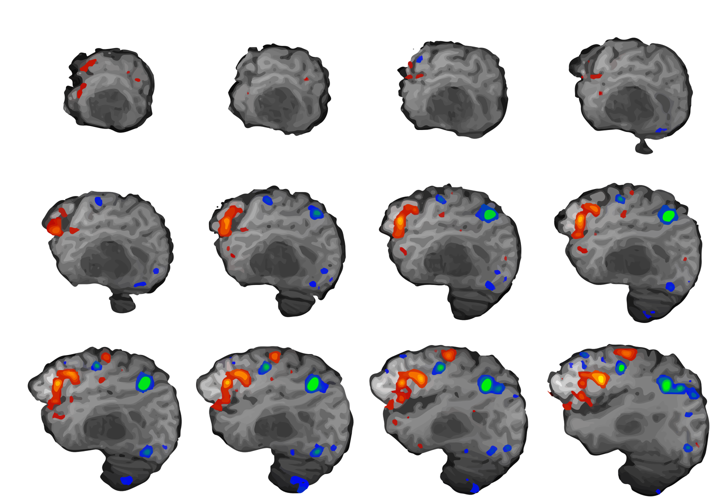


Fig 6: Example inference using the proposed model. Sagittal sections show inferred activation's proximity to a large glioblastoma in the left temporal lobe

Scanning Session Description

- Patient scanned while performing a word reading task (30 sec on/off block design)
- Two resolutions collected over separate sessions

(High-Res: $120 \times 120 \times 62$ grid size; Std-Res: $64 \times 64 \times 48$ grid size)

- Siemens 3 T scanner; 32 channel head coil

Posterior Inference

- Sampling via Riemann manifold HMC [2]
- Sparse approximation with high resolution voxel locations used as the "inducing set" to maximize spatial precision
  - Computational burden alleviated with circulant matrix embedding [5, 6]

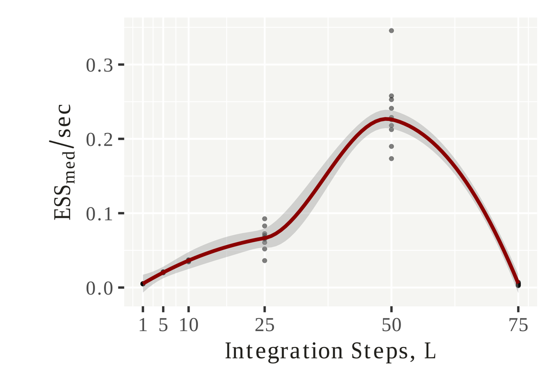


Fig 7: Computation time in analysis of real patient data. ESS_{med} denotes median effective sample size across all voxels

- Differently penalize false negative and false positive errors. Adapted from [4, 3]

$$L(\boldsymbol{m}, \boldsymbol{d}) = \sum_i \underbrace{-f(m_i)d_i - [1 - f(m_i)](1 - d_i)}_{\text{Gains for correct decisions}} \underbrace{+ k_1 f(m_i)(1 - d_i) + k_2[1 - f(m_i)]d_i + td_i}_{\text{Penalties for false } -/+}$$

- k_1, k_2, t tunable parameters, e.g. t penalizes the overall number of discoveries
- $m_i = |\mathbb{E}(\mu_{hi} | \boldsymbol{y}_h, \boldsymbol{y}_s)| / \sqrt{\text{var}(\mu_{hi} | \boldsymbol{y}_h, \boldsymbol{y}_s)}$
- $f(\cdot)$ some monotone function of signal strength: e.g. $f(m_i) = m_i / \max_i m_i$
  - Proxy measure for π("voxel i is active"$|\boldsymbol{y}_h, \boldsymbol{y}_s$)
- Collaborating radiologist's advice: penalize false negatives $11\times$ more heavily than false positives

References

[1] Belliveau, JW et al. *Science* 254.5032 (1991), pp. 716–719.
[2] Girolami, Mark and Calderhead, Ben. *JRSS B* 73.2 (2011), pp. 123–214.
[3] Liu, Zhuqing et al. *Bayesian Analysis* 11.2 (2016), p. 599.
[4] Muller, Peter, Parmigiani, Giovanni, and Rice, Kenneth. (2006).
[5] Rue, Havard and Held, Leonhard. Chapman and Hall/CRC, 2005.
[6] Wood, Andrew TA and Chan, Grace. *J Comp Graph Stat* 3.4 (1994), pp. 409–432.

Contact & Software

Andrew Whiteman: `awhitem@umich.edu`
Jian Kang: `jiankang@umich.edu`
Timothy D. Johnson: `tdjtdj@umich.edu`