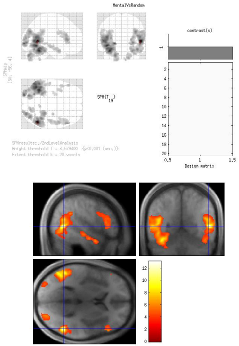
Spatiotemporal Mixed Modeling of task fMRI

Ben Risk, SAMSI and UNC-CH Biostatistics
David Matteson and David Ruppert, Cornell University

September 25, 2015

Scientific objective

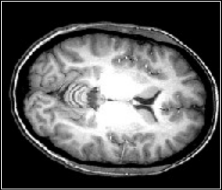
- A goal of neuroscience is to map the locations of brain activity that produce various thoughts and behaviors
- fMRI is used to localize patterns of activation in response to task stimuli
- Example: Theory of mind, or mentalizing, is the ability to infer another’s mental state



Overview of fMRI

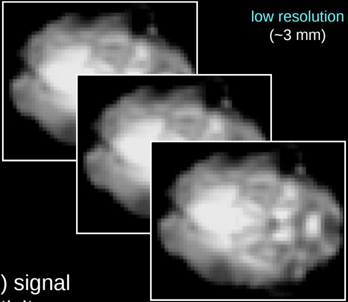
MRI

high resolution
(1 mm)



one image (no time)

fMRI



low resolution
(~3 mm)

many images
(e.g., every 2 sec for 5 mins)

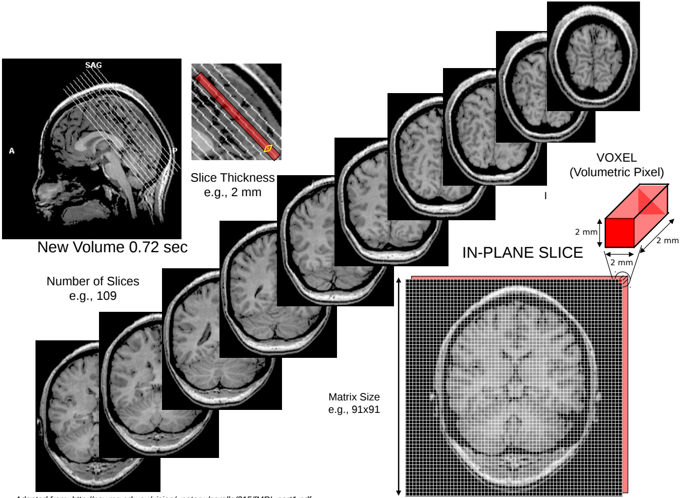
fMRI

Blood Oxygenation Level Dependent (BOLD) signal
-indirect, “sluggish” measure of neural activity

↑ neural activity ➔ ↑ blood oxygen ➔ ↑ fMRI signal

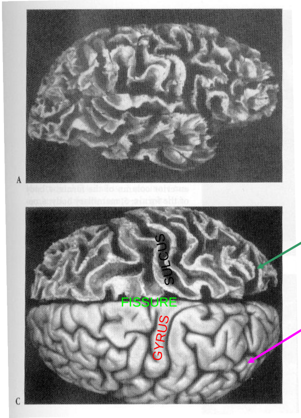
Adapted from: http://psy.mq.edu.au/vision/~peterw/corella/315/fMRI_part1.pdf

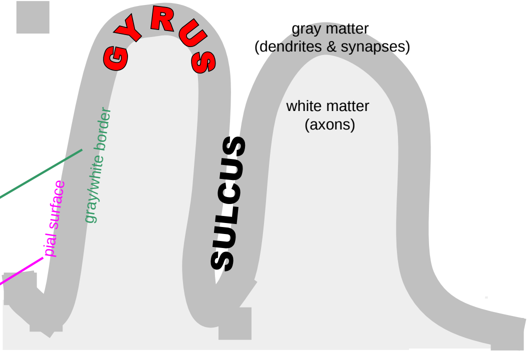
fMRI volumes



Adapted from: http://psy.mq.edu.au/vision/~peterw/corella/315/fMRI_part1.pdf

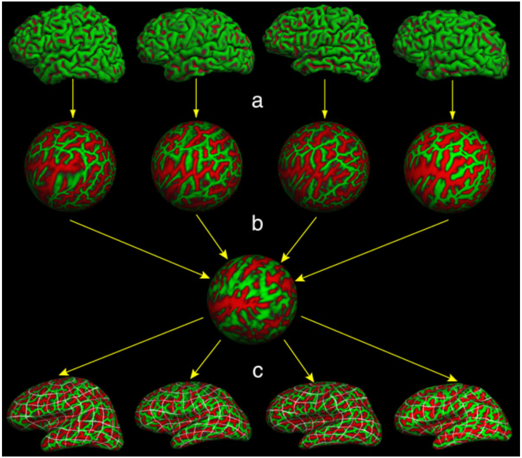
Topology of the brain





Adapted from: http://psy.mq.edu.au/vision/~peterw/corella/315/fMRI_part1.pdf

Surface registration: vertices



Challenges of fMRI

- Our example dataset:
  $N = 20$, 2 sessions, $T = 274$, Voxel dimensions = $91 \times 109 \times 91$;
  - 9.9 billion data points

Challenges of fMRI

- Our example dataset:
  $N = 20$, 2 sessions, $T = 274$, Voxel dimensions = $91 \times 109 \times 91$;
  - 9.9 billion data points
- After masking, we have
  $N = 20$, 2 sessions, $T = 274$, 230k voxels
  - 2.5 billion data points

Challenges of fMRI

- Our example dataset:
  $N = 20$, 2 sessions, $T = 274$, Voxel dimensions = $91 \times 109 \times 91$;
  - 9.9 billion data points
- After masking, we have
  $N = 20$, 2 sessions, $T = 274$, 230k voxels
  - 2.5 billion data points
- If we restrict ourselves to right cerebral cortex, $N = 20$, 2 session, $T = 274$, 30k vertices
  - 326 million data points

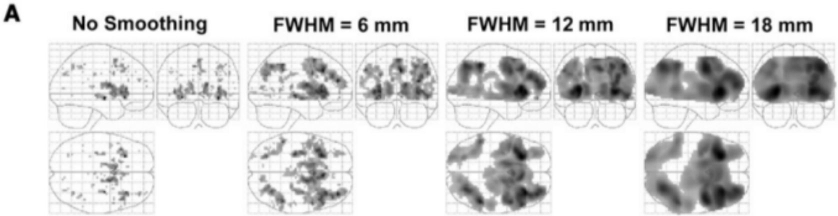
Motivation

Most popular approach: Massive Univariate Mixed Model (MUMM)

- Separate mixed model estimated for each voxel
- Spatial dependence can be accounted for post-estimation using random field theory
- Data are spatially smoothed with a Gaussian kernel during pre-processing to:
    1. Increase overlap of features between subjects
    2. Increase power and meet assumptions of random field theory
- Conflict: increase power but decrease precision of localization
- So how much smoothing? Not clear.

Smoothing affects inference

M. Mikl et al. / Magnetic Resonance Imaging 26 (2008) 490–503



Literature Review of Spatial Models

1. [Bowman, 2007] used subject-vertex random effects with exponential covariogram and functionally defined distance.

Literature Review of Spatial Models

1. [Bowman, 2007] used subject-vertex random effects with exponential covariogram and functionally defined distance.
   - Data application: $N = 12$, $V = 239$, $T = 4$ (PET data)

Literature Review of Spatial Models

1. [Bowman, 2007] used subject-vertex random effects with exponential covariogram and functionally defined distance.
   - Data application: $N = 12$, $V = 239$, $T = 4$ (PET data)
2. [Bernal-Rusiel et al., 2013] extended [Bowman, 2007] to cortical thickness using geodesic distances

Literature Review of Spatial Models

1. [Bowman, 2007] used subject-vertex random effects with exponential covariogram and functionally defined distance.
   - Data application: $N = 12$, $V = 239$, $T = 4$ (PET data)
2. [Bernal-Rusiel et al., 2013] extended [Bowman, 2007] to cortical thickness using geodesic distances
   - $N \approx 800$, $V = 149,000$, $T \approx 5$. Defined novel parcellation with 12 vertices/parcel, parcels independent (12,000 parcels!). FWHM=15mm.

Literature Review of Spatial Models

1. [Bowman, 2007] used subject-vertex random effects with exponential covariogram and functionally defined distance.
   - Data application: $N = 12$, $V = 239$, $T = 4$ (PET data)
2. [Bernal-Rusiel et al., 2013] extended [Bowman, 2007] to cortical thickness using geodesic distances
   - $N \approx 800$, $V = 149,000$, $T \approx 5$. Defined novel parcellation with 12 vertices/parcel, parcels independent (12,000 parcels!). FWHM=15mm.
3. [Derado et al., 2010] assumed constant correlation within Brodmann Area, independent BAs

Literature Review of Spatial Models

1. [Bowman, 2007] used subject-vertex random effects with exponential covariogram and functionally defined distance.
   - Data application: $N = 12$, $V = 239$, $T = 4$ (PET data)
2. [Bernal-Rusiel et al., 2013] extended [Bowman, 2007] to cortical thickness using geodesic distances
   - $N \approx 800$, $V = 149,000$, $T \approx 5$. Defined novel parcellation with 12 vertices/parcel, parcels independent (12,000 parcels!). FWHM=15mm.
3. [Derado et al., 2010] assumed constant correlation within Brodmann Area, independent BAs
   - Use coefficients from single-subject analysis, $N = 27$, 2 sessions, $53 \times 63 \times 46, T =$hundreds

Literature Review of Spatial Models

1. [Bowman, 2007] used subject-vertex random effects with exponential covariogram and functionally defined distance.
   - Data application: $N = 12$, $V = 239$, $T = 4$ (PET data)
2. [Bernal-Rusiel et al., 2013] extended [Bowman, 2007] to cortical thickness using geodesic distances
   - $N \approx 800$, $V = 149,000$, $T \approx 5$. Defined novel parcellation with 12 vertices/parcel, parcels independent (12,000 parcels!). FWHM=15mm.
3. [Derado et al., 2010] assumed constant correlation within Brodmann Area, independent BAs
   - Use coefficients from single-subject analysis, $N = 27$, 2 sessions, $53 \times 63 \times 46, T =$hundreds
4. [Kang et al., 2012] spatial modeling in spectral (time) domain allowing for correlated regions of interest (ROIs)

Literature Review of Spatial Models

1. [Bowman, 2007] used subject-vertex random effects with exponential covariogram and functionally defined distance.
    - Data application: $N = 12$, $V = 239$, $T = 4$ (PET data)
2. [Bernal-Rusiel et al., 2013] extended [Bowman, 2007] to cortical thickness using geodesic distances
    - $N \approx 800$, $V = 149,000$, $T \approx 5$. Defined novel parcellation with 12 vertices/parcel, parcels independent (12,000 parcels!). FWHM=15mm.
3. [Derado et al., 2010] assumed constant correlation within Brodmann Area, independent BAs
    - Use coefficients from single-subject analysis, $N = 27$, 2 sessions, $53 \times 63 \times 46$, $T =$hundreds
4. [Kang et al., 2012] spatial modeling in spectral (time) domain allowing for correlated regions of interest (ROIs)
    - $N = 1$, 4 sessions, $V = 20$/ROI, 3 ROIs, $T = 480$

SPM and FSL: Massive univariate analysis

- $n \in \{1, \ldots, N\}$ denote subject
- $v \in \{1, \ldots, V\}$ denote location (voxel for volume data, vertex for surface data)
- $t \in \{1, \ldots, T\}$ index time
- $q \in \{1, \ldots, Q\}$ index the task
- Task onsets and duration: $I_{nq}(t)$; shape of the BOLD response to a stimulus: $h(t)$.

$$x_{ntq} = \int_0^{f_t} h(u) I_{nq}(f_t - u) du$$

where f_t is the time in seconds corresponding to the tth timepoint and we assume the same $h(t)$ for all voxels.

Hierarchical MUMM

$$(1-\phi_{nv1}B-\phi_{nv2}B^2-\phi_{nv3}B^3)\left\{y_{nvt}-\mathbf{x}'_{nt}\left(\beta_v+\boldsymbol{b}_{nv}\right)-\mathbf{z}'_{nt}\gamma_{nv}\right\} = \epsilon_{nvt}$$

with $\boldsymbol{b}_{nv} \overset{iid}{\sim} \mathcal{N}\{\mathbf{0}, \mathrm{diag}(\sigma^2_{b_1}, \ldots, \sigma^2_{b_Q})\}$ and $\epsilon_{nvt} \overset{iid}{\sim} \mathcal{N}(0, \tau^2_{nv})$. All b_{nvq} and ϵ_{nvt} are mutually independent.
Let $\mathbf{B} = \mathrm{diag}(\sigma^2_{b_1}, \ldots, \sigma^2_{b_Q})$. Then,

$$\mathrm{Cov}\,\mathbf{Y}_{nv} = \mathbf{X}_{nt}\mathbf{B}\mathbf{X}'_{nt} + \xi^2_{nv}\boldsymbol{\Psi}_{nv}. \tag{1}$$

MUMM estimators

To simplify the exposition, assume $\mathbf{z}_{nt} = 0$. Define

$$\hat{\mathbf{e}}_{nv} = (\mathbf{X}_n'\mathbf{X}_n)^{-1}\mathbf{X}_n'\mathbf{Y}_{nv}. \tag{2}$$

Define the estimate of the population coefficients:

$$\hat{\beta}_v = \frac{1}{N}\sum_{n=1}^{N}\hat{\mathbf{e}}_{nv}. \tag{3}$$

Define the estimate of the variance of the fixed effect estimator:

$$\widehat{\mathrm{Cov}}\,\hat{\beta}_v = \frac{1}{N(N-1)}\sum_{n=1}^{N}\left(\hat{\mathbf{e}}_{nv} - \hat{\beta}_v\right)\left(\hat{\mathbf{e}}_{nv} - \hat{\beta}_v\right)'. \tag{4}$$

Spatiotemporal Mixed Model (STMM)

1. Automate smoothing using spatial random effects that capture population activation.

Spatiotemporal Mixed Model (STMM)

1. Automate smoothing using spatial random effects that capture population activation.
2. Incorporate subject-vertex random effects to allow subject-specific deviations in activation and/or alignment.

Spatiotemporal Mixed Model (STMM)

1. Automate smoothing using spatial random effects that capture population activation.
2. Incorporate subject-vertex random effects to allow subject-specific deviations in activation and/or alignment.
3. Unified model that includes subject- and vertex-specific autoregressive errors, which contrasts with previous methods that use the output from a first-level analysis.

Spatiotemporal Mixed Model (STMM)

1. Automate smoothing using spatial random effects that capture population activation.
2. Incorporate subject-vertex random effects to allow subject-specific deviations in activation and/or alignment.
3. Unified model that includes subject- and vertex-specific autoregressive errors, which contrasts with previous methods that use the output from a first-level analysis.
4. Leverage improvements in registration using geodesic distances and improvements in parcellation using the Gordon parcels.

Spatiotemporal Mixed Model (STMM)

1. Automate smoothing using spatial random effects that capture population activation.
2. Incorporate subject-vertex random effects to allow subject-specific deviations in activation and/or alignment.
3. Unified model that includes subject- and vertex-specific autoregressive errors, which contrasts with previous methods that use the output from a first-level analysis.
4. Leverage improvements in registration using geodesic distances and improvements in parcellation using the Gordon parcels.
5. Fast estimators of spatial dependence that can be used for whole-brain multi-subject studies.

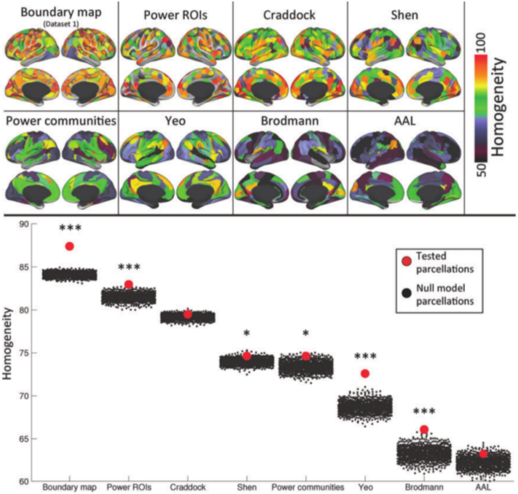
STMM

Define a separate model for each parcel.
To simplify the notation, let us consider a single region and ignore the index for region.
Let

$$\mathbf{Y}_n = [y_{n11}, \ldots, y_{n1T}, y_{n21}, \ldots, y_{n2T}, \ldots, y_{nVT}]'$$

denote the BOLD signal for the nth subject at the vertices in our region, where V is the number of vertices in the region.
Let $\mathbf{Y} = [\mathbf{Y}_1', \ldots, \mathbf{Y}_N']'$.

[Gordon et al., 2014]



STMM

$$y_{nvt} = \mathbf{x}'_{nt}(\beta + \boldsymbol{u}_v + \boldsymbol{s}_n + \boldsymbol{b}_{nv}) + \mathbf{z}'_{nt}\gamma_{nv} + a_{nvt}, \tag{5}$$

where

$$\boldsymbol{u}^q \sim \mathcal{N}(\mathbf{0}, \sigma^2_{u_q}\boldsymbol{\Gamma}_q);$$

$$\boldsymbol{s}_n \overset{iid}{\sim} \mathcal{N}(\mathbf{0}, \mathbf{S})$$

with $\mathbf{S} = \mathrm{diag}(\sigma^2_{s_1}, \ldots, \sigma^2_{s_Q})$;

$$\boldsymbol{b}^q_n \overset{iid}{\sim} \mathcal{N}(\mathbf{0}, \sigma^2_{b_q}\Omega_q);$$

and

$$\boldsymbol{a} \sim \mathcal{N}(\mathbf{0}, \oplus_{n=1}^{N} \oplus_{v=1}^{V} \xi^2_{nv}\boldsymbol{\Psi}_{nv}).$$

Spatial Correlation

We assume the spatial correlation structure defined by exponential covariogram:

$$\boldsymbol{\Gamma}_{q;vv'} = e^{-\theta_{bq}||v-v'||}$$

and

$$\Omega_{q;vv'} = e^{-\theta_{bq}||v-v'||}.$$

Covariance Model

$$\operatorname{Cov} y_{nvt}, y_{n'v't'} =$$

$$\begin{cases} \mathbf{x}'_{nt}(\mathbf{U}+\mathbf{S}+\mathbf{B})\mathbf{x}_{nt'} + \xi^2_{nv}\psi^{(nv)}_{tt'} & n = n';\, v = v';\ \text{any } t, t' \\ \mathbf{x}'_{nt}(\mathbf{U}\Gamma_{vv'} + \mathbf{S} + \mathbf{B}\Omega_{vv'})\mathbf{x}_{nt'} & n = n';\, v \neq v';\ \text{any } t, t' \\ \mathbf{x}'_{nt}\mathbf{U}\Gamma_{vv'}\mathbf{x}_{n't'} & n \neq n';\ \text{any } v, v';\ \text{any } t, t' \end{cases}$$

No independence anywhere.

ML and REML impossible

- We can not use maximum likelihood or restricted maximum likelihood methods to fit this model due to the enormous size of the covariance matrix, which is non-sparse for each parcel.

ML and REML impossible

- We can not use maximum likelihood or restricted maximum likelihood methods to fit this model due to the enormous size of the covariance matrix, which is non-sparse for each parcel.
- In our application, the largest parcel's covariance matrix is 11 million by 11 million.

ML and REML impossible

- We can not use maximum likelihood or restricted maximum likelihood methods to fit this model due to the enormous size of the covariance matrix, which is non-sparse for each parcel.
- In our application, the largest parcel's covariance matrix is 11 million by 11 million.
- The method we will propose involves a 39,440 $\times$ 39,440 covariance matrix, which is about as much as we can handle with 32GB of RAM and computational shortcuts such as Sherman-Morrison-Woodbury.

Second-level model

Transform the time dimension from T to the number of covariates Q:

$$\boldsymbol{d}_{nv} = \mathbf{K}_n{}'\mathbf{Y}_{nv}. \tag{6}$$

Define the second-level model

$$\begin{aligned}\boldsymbol{d} &= (\mathbf{1}_N \otimes \mathbf{1}_V \otimes \mathbf{I}_Q)\boldsymbol{\beta} + (\mathbf{1}_N \otimes \mathbf{I}_V \otimes \mathbf{I}_Q)\boldsymbol{u} + (\mathbf{I}_N \otimes \mathbf{1}_V \otimes \mathbf{I}_Q)\boldsymbol{s} \\ &+ (\mathbf{I}_N \otimes \mathbf{I}_V \otimes \mathbf{I}_Q)\boldsymbol{b} + \{_c\ \{_c\ \mathbf{K}_n'\boldsymbol{a}_{nv}\}_{v=1}^{V}\}_{n=1}^{N}.\end{aligned}$$

Estimation highlights

1. Bias-reduced estimators of AR parameters using OLS residuals and Yule-Walker equations (adapted from [Worsley et al., 2002])

Estimation highlights

1. Bias-reduced estimators of AR parameters using OLS residuals and Yule-Walker equations (adapted from [Worsley et al., 2002])
2. Assume $\mathbf{\Gamma}_q = \Omega_q$.

Estimation highlights

1. Bias-reduced estimators of AR parameters using OLS residuals and Yule-Walker equations (adapted from [Worsley et al., 2002])
2. Assume $\boldsymbol{\Gamma}_q = \Omega_q$.
3. Define a novel covariogram to allow vertex-specific temporal correlation

$$\delta_q(d_{nvq}, d_{nv'q}) = \begin{cases} \sigma^2_{u_q} + \sigma^2_{s_q} + \sigma^2_{b_q} + \beta^2_q & v = v', \\ (\sigma^2_{s_q} + \beta^2_q) + (\sigma^2_{u_q} + \sigma^2_{b_q})e^{-\theta_q ||v-v'||} & v \neq v'. \end{cases}$$

Estimation highlights

1. Bias-reduced estimators of AR parameters using OLS residuals and Yule-Walker equations (adapted from [Worsley et al., 2002])
2. Assume $\Gamma_q = \Omega_q$.
3. Define a novel covariogram to allow vertex-specific temporal correlation

$$\delta_q(d_{nvq}, d_{nv'q}) = \begin{cases} \sigma^2_{u_q} + \sigma^2_{s_q} + \sigma^2_{b_q} + \beta^2_q & v = v', \\ (\sigma^2_{s_q} + \beta^2_q) + (\sigma^2_{u_q} + \sigma^2_{b_q})e^{-\theta_q \|v - v'\|} & v \neq v'. \end{cases}$$

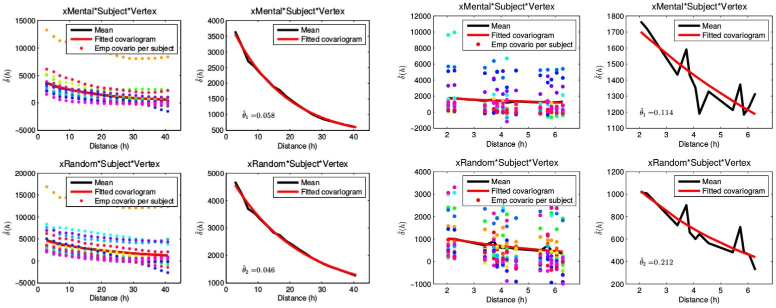
4. Define novel empirical covariogram for tolerance ν:

$$\hat{\delta}_n(h) = \frac{1}{N_h} \sum_{\{v,v'\}:\ \|v-v'\| \in (h-\nu, h+\nu]} d_{nvq} d_{nv'q} \quad (7)$$

We can estimate θ_q without knowing the other parameters!

Estimating the spatial correlation parameter

Figure : Empirical covariogram and fitted exponential covariogram for the spatial random effects for a large (777 vertices) and very small (29 vertices) parcel.



ANOVA estimators of variance components

Decompose the variance of the transformed data, $\boldsymbol{d}_{nv} = \mathbf{K}_n{}'\mathbf{Y}_{nv}$.

Source	DF	Sum of Squares
Vertex	V-1	$\sum_{n=1}^{N}\sum_{v=1}^{V}(\bar{\boldsymbol{d}}_{\cdot v} - \bar{\boldsymbol{d}}_{\cdot\cdot})(\bar{\boldsymbol{d}}_{\cdot v} - \bar{\boldsymbol{d}}_{\cdot\cdot})'$
Subject	N-1	$\sum_{n=1}^{N}\sum_{v=1}^{V}(\bar{\boldsymbol{d}}_{n\cdot} - \bar{\boldsymbol{d}}_{\cdot\cdot})(\bar{\boldsymbol{d}}_{n\cdot} - \bar{\boldsymbol{d}}_{\cdot\cdot})'$
Subject-Vertex	(N-1)(V-1)	$\sum_{n=1}^{N}\sum_{v=1}^{V}(\boldsymbol{d}_{nv} - \bar{\boldsymbol{d}}_{n\cdot} - \bar{\boldsymbol{d}}_{\cdot v} + \bar{\boldsymbol{d}}_{\cdot\cdot})(\boldsymbol{d}_{nv} - \bar{\boldsymbol{d}}_{n\cdot} - \bar{\boldsymbol{d}}_{\cdot v} + \bar{\boldsymbol{d}}_{\cdot\cdot})'$

ANOVA estimators of variance components

Let $g_q = \sum_{v=1}^{V} \sum_{v'=1}^{V} \mathbf{\Gamma}_{q;v,v'}$ and $\mathbf{G} = \mathrm{diag}(g_1, \ldots, g_Q)$.
Let $w_q = \sum_{v=1}^{V} \sum_{v'=1}^{V} \Omega_{q;v,v'}$ and $\mathbf{W} = \mathrm{diag}(w_1, \ldots, w_Q)$.

Source	EMS
Vertex	$\left(\frac{NV}{V-1}\mathbf{I}_Q - \frac{N}{V(V-1)}\mathbf{G}\right)\mathbf{U} + \left(\frac{V}{V-1}\mathbf{I}_Q - \frac{1}{V(V-1)}\mathbf{W}\right)\mathbf{B} + \frac{1}{NV}\sum_{n=1}^{N}\sum_{v'=1}^{V}\mathbf{K}_n\mathbf{\Psi}_{nv'}\mathbf{K}_n'$
Subject	$V\mathbf{S} + \frac{1}{V}\mathbf{W}\mathbf{B} + \frac{1}{NV}\sum_{n=1}^{N}\sum_{v=1}^{V}\xi_{nv}^2\mathbf{K}_n'\mathbf{\Psi}_{nv}\mathbf{K}_n$
Subject-Vertex	$\left(\frac{V}{V-1}\mathbf{I}_Q - \frac{1}{V(V-1)}\mathbf{W}\right)\mathbf{B} + \frac{1}{NV}\sum_{n=1}^{N}\sum_{v=1}^{V}\xi_{nv}^2\mathbf{K}_n'\mathbf{\Psi}_{nv}\mathbf{K}_n$

BLUEs, BLUPs, and eBLUPs

Define the BLUE for regional activation:

$$\hat{\beta} = \left\{(\mathbf{1}'_N \otimes \mathbf{1}'_V \otimes \mathbf{I}_Q)\boldsymbol{\Sigma}^{-1}(\mathbf{1}_N \otimes \mathbf{1}_V \otimes \mathbf{I}_Q)\right\}^{-1}(\mathbf{1}'_N \otimes \mathbf{1}'_V \otimes \mathbf{I}_Q)\boldsymbol{\Sigma}^{-1}\boldsymbol{d} \tag{8}$$

$$\mathrm{Cov}\,\hat{\beta} = \left\{(\mathbf{1}'_N \otimes \mathbf{1}'_V \otimes \mathbf{I}_Q)\boldsymbol{\Sigma}^{-1}(\mathbf{1}_N \otimes \mathbf{1}_V \otimes \mathbf{I}_Q)\right\}^{-1}.$$

Define the BLUP for vertex activation:

$$\hat{\boldsymbol{u}} = \Gamma(\mathbf{1}'_N \otimes \mathbf{I}_V \otimes \mathbf{I}_Q)\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{d} - \mathbf{1}_N \otimes \mathbf{1}_V \otimes \hat{\beta}\right). \tag{9}$$

We can also estimate $\mathrm{Cov}\,\hat{\boldsymbol{u}}$.

Approximate inference

$$[\boldsymbol{u}|\boldsymbol{d}, \widehat{\Sigma}, \hat{\beta}] \sim \mathcal{N}\left(\widehat{\Gamma}\left\{\mathbf{1}_N' \otimes \mathbf{I}_V \otimes \mathbf{I}_Q\right\}\widehat{\Sigma}^{-1}\left(\boldsymbol{d} - \mathbf{1}_N \otimes \mathbf{1}_V \otimes \hat{\beta}\right),\right.$$
$$\left.\widehat{\Gamma} - \widehat{\Gamma}\left\{\mathbf{1}_N' \otimes \mathbf{I}_V \otimes \mathbf{I}_Q\right\}\widehat{\Sigma}^{-1}\left\{\mathbf{1}_N \otimes \mathbf{I}_V \otimes \mathbf{I}_Q\right\}\widehat{\Gamma}\right).$$

For an arbitrary contrast vector $\mathbf{c}$,

$$H_0 : \mathbf{c}'[\beta', \mathrm{E}(\boldsymbol{u}|\boldsymbol{d})']' = 0$$
$$H_A : \mathbf{c}'[\beta', \mathrm{E}(\boldsymbol{u}|\boldsymbol{d})']' \neq 0.$$

We use $\hat{\boldsymbol{u}}$ as our estimate of $\mathrm{E}(\boldsymbol{u}|\boldsymbol{d})$. Approximate t-statistic:

$$t = \frac{\mathbf{c}'[\hat{\beta}', \hat{\boldsymbol{u}}']'}{\left\{\mathbf{c}'(\widehat{\mathrm{Cov}}\,\hat{\beta} \oplus \widehat{\mathrm{Cov}}\,\hat{\boldsymbol{u}})\mathbf{c}\right\}^{1/2}}.$$

Estimators are approximately unbiased

	Truth	df	Mean STMM	Var STMM	Bias^2 STMM	MSE STMM
τ^2_{nv}	30000	-	29811	3434698	35882	3470580
ϕ_{nv1}	0.2000	-	0.1988	0.0020	0.0000	0.0020
ϕ_{nv2}	0.1000	-	0.0969	0.0020	0.0000	0.0021
ϕ_{nv3}	0.0000	-	-0.0008	0.0020	0.0000	0.0020
β_1	20.0	-	19.9	16.2	0.0	16.2
$\sigma^2_{u_1}$	75.0	214	77.0	776.0	4.1	780.2
$\sigma^2_{s_1}$	250	29	241	8521	84	8605
$\sigma^2_{b_1}$	750	6206	750	3517	0	3517
θ_{b_1}	0.200	-	0.201	0.002	0.000	0.002
β_2	-20.0	-	-20.0	17.5	0.0	17.5
$\sigma^2_{u_2}$	75.0	214	75.9	808.2	0.9	809.0
$\sigma^2_{s_2}$	250	29	244	7947	38	7985
$\sigma^2_{b_2}$	750	6206	749	3747	2	3749
θ_{b_2}	0.200	-	0.203	0.002	0.000	0.002

Table : Accuracy of estimators for 1000 simulations with 30 subjects each in which 215 vertices were located according to a Gordon Parcel (ID 82).

STMM: more power with no **U**; less power with **U**

		Region			Vertex		
Scenario	Model	$\beta_1 = 20$ (Power)	$\beta_2 = 0$ (Type 1)	Contrast (Power)	$\beta_1 + u_{v1}$ (Power)	$\beta_2 + u_{v2}$ (Type 1*)	Contrast (Power)
$\theta_1 = \theta_2 = 5$ $\sigma^2_{u_1} = \sigma^2_{u_2} = 0$	MUMM	1.00	0.72	1.00	0.61	0.05	0.38
	ROIMM	1.00	0.05	0.99	-	-	-
	STMM	1.00	0.06	0.99	1.00	0.02	0.99
$\theta_1 = \theta_2 = 5$ $\sigma^2_{u_1} = \sigma^2_{u_2} = 100$	MUMM	1.00	0.70	1.00	0.57	0.19	0.41
	ROIMM	1.00	0.05	1.00	-	-	-
	STMM	1.00	0.06	1.00	0.63	0.02	0.35
$\theta_1 = \theta_2 = 0.2$ $\sigma^2_{u_1} = \sigma^2_{u_2} = 0$	MUMM	1.00	0.72	1.00	0.61	0.05	0.39
	ROIMM	1.00	0.05	0.99	-	-	-
	STMM	1.00	0.04	0.99	1.00	0.02	0.98
$\theta_1 = \theta_2 = 0.2$ $\sigma^2_{u_1} = \sigma^2_{u_2} = 100$	MUMM	1.00	0.79	1.00	0.56	0.19	0.42
	ROIMM	1.00	0.14	0.95	-	-	-
	STMM	0.99	0.03	0.86	0.57	0.02	0.31

Table : Power and type 1 error rates based on 300 simulations for each scenario.

(xMental - xRandom) from HCP ToM

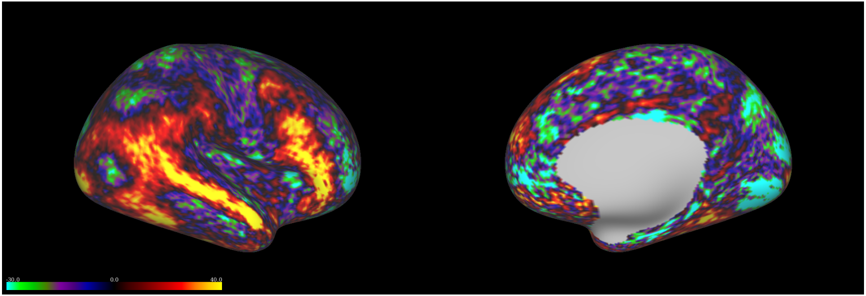


Figure : Contrast in activation from the mentalizing versus random tasks from the MUMM.

(xMental - xRandom) from HCP ToM

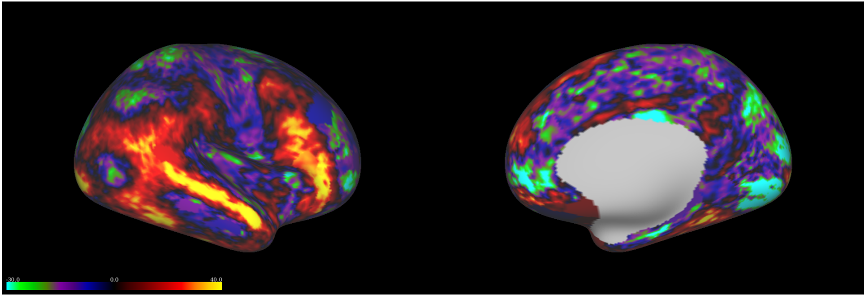


Figure : Contrast in activation from the mentalizing versus random tasks from the STMM.

t-statistic (xMental-xRandom) from HCP ToM

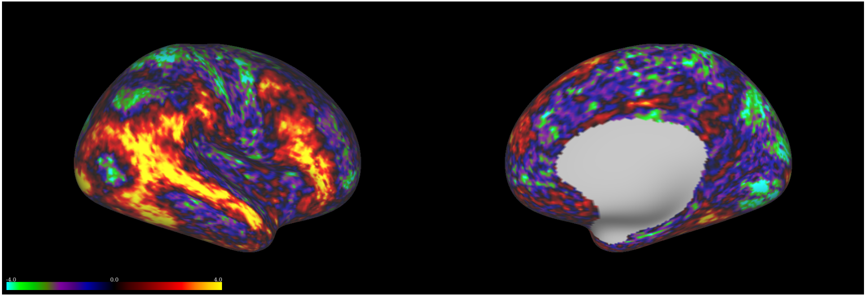


Figure : The contrast t-statistic from the MUMM

t-statistic (xMental-xRandom) from HCP ToM

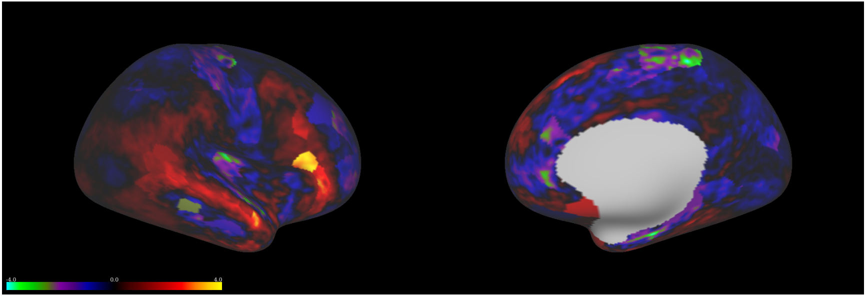


Figure : The contrast t-statistic from the STMM

Summary of Spatiotemporal Mixed Model (STMM)

1. Automate smoothing.

Summary of Spatiotemporal Mixed Model (STMM)

1. Automate smoothing.
2. Subject-vertex random effects allow subject-specific deviations in activation and/or alignment.

Summary of Spatiotemporal Mixed Model (STMM)

1. Automate smoothing.
2. Subject-vertex random effects allow subject-specific deviations in activation and/or alignment.
3. Unified model that includes subject- and vertex-specific autoregressive errors, which contrasts with previous methods that use the output from a first-level analysis.

Summary of Spatiotemporal Mixed Model (STMM)

1. Automate smoothing.
2. Subject-vertex random effects allow subject-specific deviations in activation and/or alignment.
3. Unified model that includes subject- and vertex-specific autoregressive errors, which contrasts with previous methods that use the output from a first-level analysis.
4. Leverage improvements in cortical registration and improvements in parcellation by using the geodesic distances between vertices within a Gordon parcel.

Summary of Spatiotemporal Mixed Model (STMM)

1. Automate smoothing.
2. Subject-vertex random effects allow subject-specific deviations in activation and/or alignment.
3. Unified model that includes subject- and vertex-specific autoregressive errors, which contrasts with previous methods that use the output from a first-level analysis.
4. Leverage improvements in cortical registration and improvements in parcellation by using the geodesic distances between vertices within a Gordon parcel.
5. Fast estimators of spatial dependence.

Discussion and the future

1. Preliminary research: a simpler model model with vertex fixed effects is more powerful method than MUMM or STMM, but don't get population effects.

Discussion and the future

1. Preliminary research: a simpler model model with vertex fixed effects is more powerful method than MUMM or STMM, but don't get population effects.
2. STMM could be used to improve pre-surgical mapping in an individual by shrinking towards population effects.

Discussion and the future

1. Preliminary research: a simpler model model with vertex fixed effects is more powerful method than MUMM or STMM, but don't get population effects.
2. STMM could be used to improve pre-surgical mapping in an individual by shrinking towards population effects.
3. Inference for any spatial domain: space is continuous in STMM, easy extension.

Discussion and the future

1. Preliminary research: a simpler model model with vertex fixed effects is more powerful method than MUMM or STMM, but don't get population effects.
2. STMM could be used to improve pre-surgical mapping in an individual by shrinking towards population effects.
3. Inference for any spatial domain: space is continuous in STMM, easy extension.
4. Examine residuals for mid/long range dependencies and model these correlations.

Discussion and the future

1. Preliminary research: a simpler model model with vertex fixed effects is more powerful method than MUMM or STMM, but don't get population effects.
2. STMM could be used to improve pre-surgical mapping in an individual by shrinking towards population effects.
3. Inference for any spatial domain: space is continuous in STMM, easy extension.
4. Examine residuals for mid/long range dependencies and model these correlations.
5. STMM could be used to examine how spatial correlation varies by ROI.

Acknowledgements

Thank you to my advisors, David Matteson and David Ruppert. Thank you to Nathan Spreng for scientific input. Data were provided (in part) by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

References I

Bernal-Rusiel, J. L., Reuter, M., Greve, D. N., Fischl, B., and Sabuncu, M. R. (2013).
Spatiotemporal linear mixed effects modeling for the mass-univariate analysis of longitudinal neuroimage data.
NeuroImage, 81:358–370.

Bowman, F. D. (2007).
Spatiotemporal models for region of interest analyses of functional neuroimaging data.
Journal of the American Statistical Association, 102(478):442–453.

Derado, G., Bowman, F. D., and Kilts, C. D. (2010).
Modeling the spatial and temporal dependence in fMRI data.
Biometrics, 66(3):949–957.

Gordon, E. M., Laumann, T. O., Adeyemo, B., Huckins, J. F., Kelley, W. M., and Petersen, S. E. (2014).
Generation and evaluation of a cortical area parcellation from resting-state correlations.
Cerebral Cortex, page bhu239.

Kang, H., Ombao, H., Linkletter, C., Long, N., and Badre, D. (2012).
Spatio-spectral mixed-effects model for functional magnetic resonance imaging data.
Journal of the American Statistical Association, 107(498):568–577.

Worsley, K. J., Liao, C., Aston, J., Petre, V., Duncan, G., Morales, F., and Evans, A. (2002).
A general statistical analysis for fmri data.
Neuroimage, 15(1):1–15.