SGPP: spatial Gaussian predictive process models for neuroimaging data

J. Won Hyun, Y. Li, J. H. Gilmore, Z. Lu, M. Styner, and H. Zhu

Neuroimage, vol. 89, pp. 70-80, 2014

February 5, 2015

Outline

Voxel-wise analysis

- Widely used to establish association between imaging data and covariates
- Two major steps:
  - Gaussian smoothing the imaging data
  - Fitting a statistical model at each voxel
- Drawbacks:
  - Gaussian smoothing may introduce bias in the statistical results
  - Does not take into account spatial correlations and dependence across different voxels
  - Generally not optimal in power
  - Not optimal in prediction

Modelling the spatial dependence

- A relatively simple covariance model has to be considered to model all voxels
  - A large unstructured variance-covariance matrix (and its functions) is computationally prohibitive to compute
- Under the Bayesian framework, spatial correlations in imaging data have been modelled through various spatial priors
  - Conditional autoregressive (CAR)
  - Markov random field (MRF)
  - Gaussian process (GP)
- Drawbacks:
  - Somehow restrictive to assume a specific type of correlation structure (CAR & MRF)
  - Several tuning parameters that need to be estimated

Scientific goals

- *Goal:* Develop a spatial Gaussian predictive process (SGPP) modelling framework for predicting neuroimaging data by using
  - A set of covariates of interest, such as age and diagnostic status
  - Existing imaging data (same & different modalities)
- To achieve a better prediction, the authors characterise both
  - Local & global spatial dependence (or variability) of imaging data
  - Spatial association of imaging data with a set of covariates of interest

Notation

- $n =$ # of subjects
- $\mathcal{D} =$ compact set in $\mathbb{R}^3$
- $d =$ centre of a voxel (or vertex) in $\mathcal{D}$
- $M =$ total # of voxels in $\mathcal{D}$
- $\mathbf{x}_i = (x_{i1}, \ldots, x_{ip})^\top = p \times 1$ vector of covariates for the ith subject (e.g., age, gender, and height)
- $\mathbf{y}_i(d_m) = (y_{i,1}(d_m), \ldots, y_{i,J}(d_m))^\top = J \times 1$ vector of neuroimaging measures (e.g., cortical thickness) at voxel d_m, $m = 1, \ldots, M$

SGPP

The SGPP is given by

$$y_{i,j}(d) = \mathbf{x}_i^\top \boldsymbol{\beta}_j(d) + \eta_{i,j}(d) + \epsilon_{i,j}(d)$$

for $i = 1, \ldots, n$ and $j = 1, \ldots J$

- $\boldsymbol{\beta}_j(d) = (\beta_{j1}(d), \ldots, \beta_{jp}(d))^\top = p \times 1$ vector of regression coefficients at d
- $\eta_{i,j}(d)$ characterises individual image variations from $\mathbf{x}_i^\top \boldsymbol{\beta}_j(d)$ & medium-to-long-range dependence of imaging data between $y_{i,j}(d)$ and $y_{i,j}(d')$ for any $d \neq d'$
- $\epsilon_{i,j}(d) =$ spatially correlated errors, capture local dependence
- $\boldsymbol{\eta}_i(d) = (\eta_{i,1}(d), \ldots, \eta_{i,J}(d))^\top$ & $\boldsymbol{\epsilon}_i(d) = (\epsilon_{i,1}(d), \ldots, \epsilon_{i,J}(d))^\top$ are mutually independent
- $\boldsymbol{\eta}_i \overset{iid}{\sim} \text{GP}(\mathbf{0}, \boldsymbol{\Sigma}_\eta)$, $\boldsymbol{\epsilon}_i \overset{iid}{\sim} \text{GP}(\mathbf{0}, \boldsymbol{\Sigma}_\epsilon)$

Functional principal component analysis (fPCA)

Consider an fPCA model for spatial process $\boldsymbol{\eta}_i(d)$:

- Spectral decomposition of $\boldsymbol{\Sigma}_\eta(d, d') = [\Sigma_{\eta,jj'}(d, d')]$:

$$\Sigma_{\eta,jj}(d, d') = \sum_{l=1}^{\infty} \lambda_{j,l}\psi_{j,l}(d)\psi_{j,l}(d')$$

  with $\{\lambda_{j,l\geqslant 0}\} \geqslant 0$ are the ordered eigenvalues, $\sum_{l=1}^{\infty} \lambda_{j,l} < \infty$, and $\psi_{j,l}(d)$'s are the corresponding orthonormal eigenfunctions

- Karhunen-Loéve expansion of $\eta_{i,j}(d)$:

$$\eta_{i,j}(d) = \sum_{l=1}^{\infty} \xi_{ij,l}\psi_{j,l}(d) \approx \sum_{l=1}^{L_0} \xi_{ij,l}\psi_{j,l}(d)$$

  where $\xi_{ij,l} = \int_{s\in\mathcal{D}} \eta_{i,j}(d)\psi_{j,l}(s)dL(s) = (j, l)$th functional principal component score of the ith subject. For each fixed (i, j), the $\xi_{ij,l}$'s are uncorrelated r.v.'s with $\mathbb{E}(\xi_{ij,l}) = 0$ and $\mathbb{E}(\xi_{ij,l}^2) = \lambda_{j,l}$

Multivariate simultaneous autoregressive (SAR) model

Assume a SAR model for $\epsilon_i(d)$:

$$\epsilon_{i,j}(d) = \rho \frac{1}{|N(d)|} \sum_{d' \in N(d)} \epsilon_{i,j}(d') + e_{i,j}(d)$$

- ρ = autocorrelation parameter, controls the strength of the local positive spatial dependence
- $N(d) =$ closest neighbouring voxels of d
- $|N(d)|$ = cardinality of $N(d)$
- $\mathbf{e}_i(d) = (e_{i,1}(d), \ldots, e_{i,J}(d))^\top \overset{iid}{\sim} \mathrm{GP}(\mathbf{0}, \mathbf{\Sigma}_e)$ with $\mathbf{\Sigma}_e(d, d') = \mathbf{0}$ for $d \neq d'$ and $\mathbf{\Sigma}_e(d, d) = \mathbf{\Sigma}_e(\theta(d))$
- $\theta(d) =$ vector of unknown parameters

SGPP model

Combining fPCA & SAR models:

$$y_{i,j}(d) \approx \mathbf{x}_i^\top \boldsymbol{\beta}_j(d) + \sum_{l=1}^{L_0} \xi_{ij,l}\psi_{j,l}(d)$$
$$+ \rho \frac{1}{|N(d)|} \sum_{d' \in N(d)} \left(y_{i,j}(d') - \mathbf{x}_i^\top \boldsymbol{\beta}_j(d') - \sum_{l=1}^{L_0} \xi_{ij,l}\psi_{j,l}(d') \right) + e_{i,j}(d)$$

Obtain a simple approximation to

$$\mathrm{Cov}(\mathbf{y}_i(d), \mathbf{y}_i(d')) = \boldsymbol{\Sigma}_y(d, d') = \boldsymbol{\Sigma}_\eta(d, d') + \boldsymbol{\Sigma}_\epsilon(d, d')$$

Estimation procedure

The estimation procedure follows three steps:

- Stage (I): the least squares estimate of the regression coefficients $\beta(d) = [\beta_1(d), \ldots, \beta_J(d)]$, denoted by $\hat{\beta}(d)$, across all voxels in $\mathcal{D}$
- Stage (II): a nonparametric estimate of $\mathbf{\Sigma}_\eta$ and its associated eigenvalues and eigenfunctions
- Stage (III): the restricted maximum likelihood estimation of ρ and $\boldsymbol{\theta} = \boldsymbol{\theta}(d)$

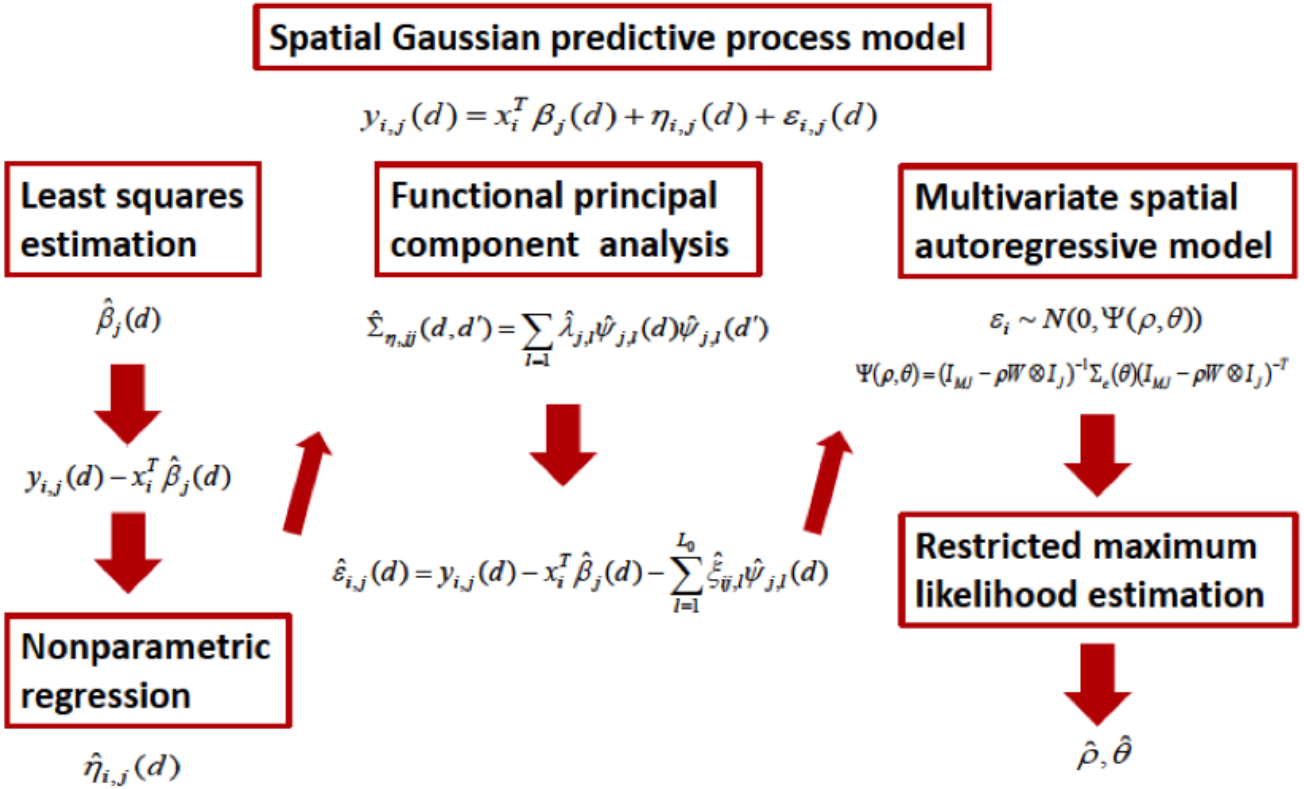

Fig. 1: A diagram for the SGPP model with three components including a general linear model (GLM) for characterizing the association between imaging measure and covariates of interest, a functional principal component model (fPCA) to capture the global spatial dependence, and a multivariate spatial autoregressive model (SAR) to capture the local spatial dependence. The first stage of the estimation procedure is the least squares estimation of the regression coefficients $\beta(d) = [\beta_1(d), \ldots, \beta_J(d)]$, the second stage is the nonparametric estimation of Σ_η and its associated eigenvalues and eigenfunctions, and the third stage is the restricted maximum likelihood estimation of all the parameters in the spatial autoregressive model.

Simulation study

- Simulated data at all 900 pixels on a 30×30 image for $n = 50$ subjects
- Data generated from a bivariate spatial Gaussian process model according to

$$y_{i,j}(d_m) = \beta_{j1}(d_m) + x_{i2}\beta_{j2}(d_m) + \eta_{i,j}(d_m) + \epsilon_{i,j}(d_m)$$

  and $j = 1, 2$; $x_{i2} \overset{iid}{\sim}$ Uniform$[1, 2], \forall i$
- $\eta_{i,j}(d_m) = \sum_{l=1}^{2} \xi_{ij,l}\psi_{j,l}(d_m)$, where the $\xi_{ij,l}$ are independently generated according to

$$\xi_{i1,1} \sim N(0, 14^2), \quad \xi_{i1,2}, \xi_{i2,2} \sim N(0, 7^2), \quad \xi_{i2,1} \sim N(0, 15^2)$$

- $\boldsymbol{\epsilon}_i = (\epsilon_i(d_1), \ldots, \epsilon_i(d_{900}))^\top$ generated from a GRF

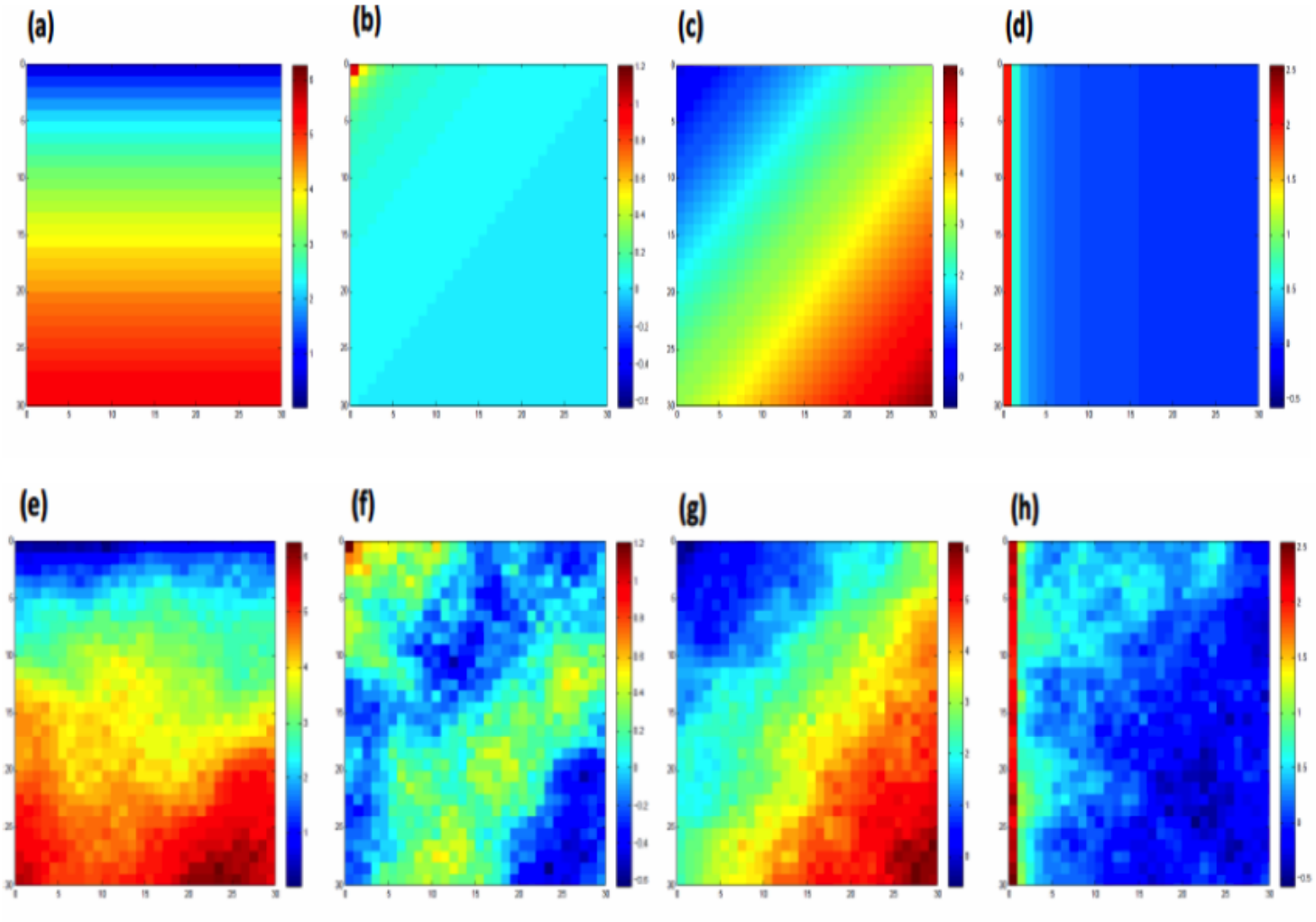

Fig. 2: Simulation results for the Gaussian random field: (a) true $\beta_{11}(d)$; (b) true $\beta_{12}(d)$; (c) true $\beta_{21}(d)$; (d) true $\beta_{22}(d)$; (e) $\hat{\beta}_{11}(d)$; (f) $\hat{\beta}_{12}(d)$; (g) $\hat{\beta}_{21}(d)$; (h) $\hat{\beta}_{22}(d)$.

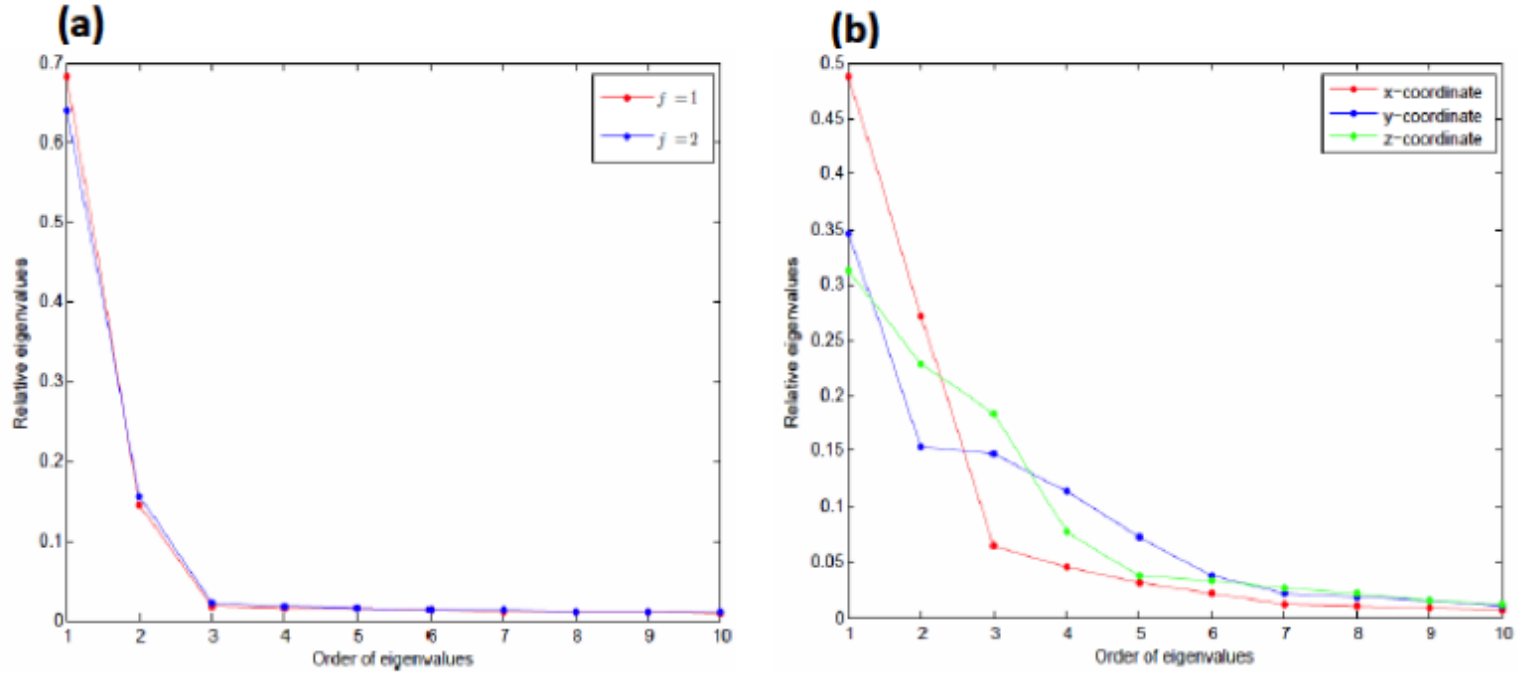

Fig. 3: The first 10 relative eigenvalues of $\hat{\Sigma}_{\eta,jj}(d,d')$ for (a) simulation results for the Gaussian random field and (b) the surface data of the left lateral ventricle.

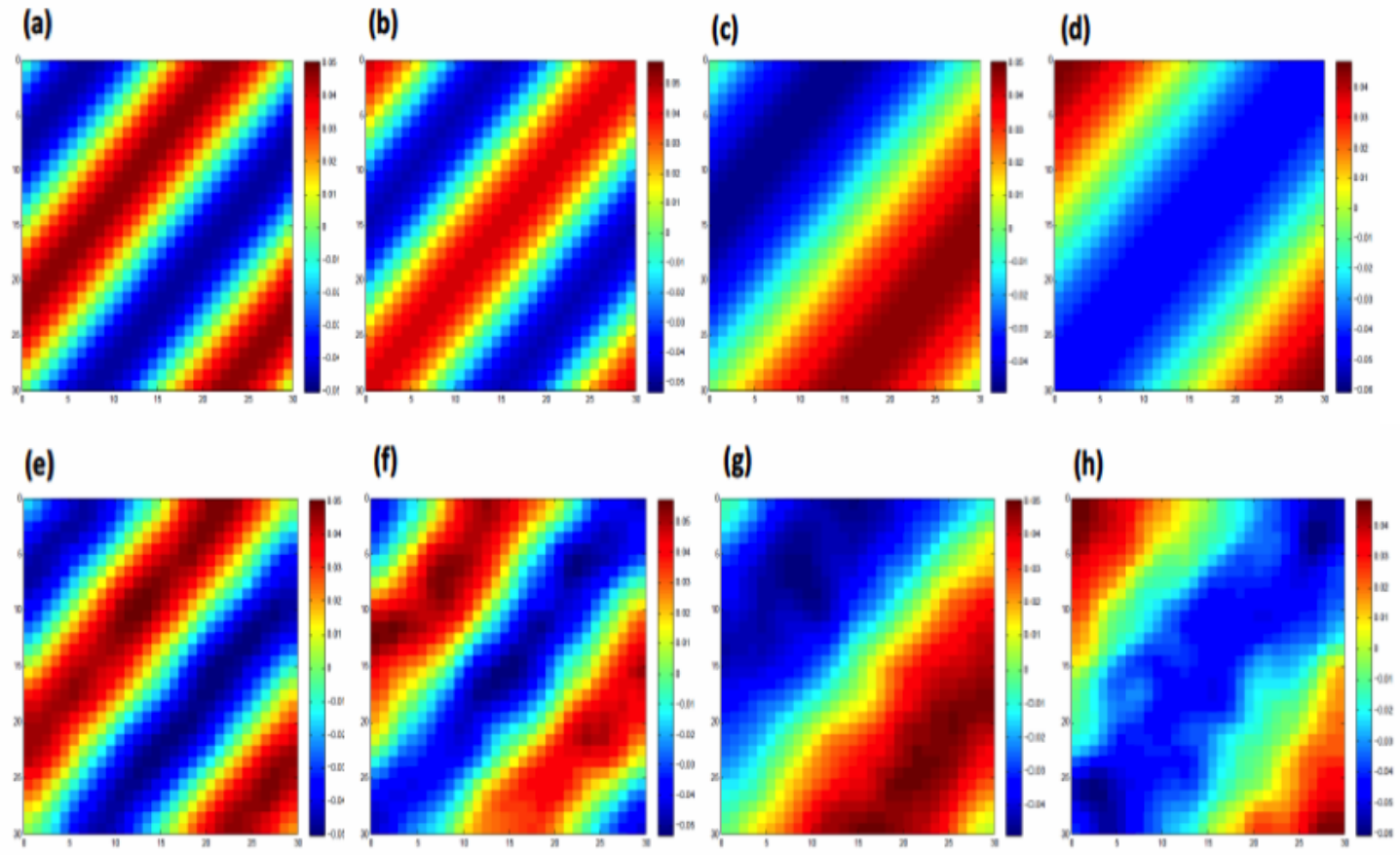

Fig. 4: Simulation results for the Gaussian random field: (a) true $\psi_{1,1}(d)$; (b) true $\psi_{1,2}(d)$; (c) true $\psi_{2,1}(d)$; (d) true $\psi_{2,2}(d)$; (e) $\hat{\psi}_{1,1}(d)$; (f) $\hat{\psi}_{1,2}(d)$; (g) $\hat{\psi}_{2,1}(d)$; and (h) $\hat{\psi}_{2,2}(d)$.

Table 1: rtMSPE for the simulated data with a Gaussian error process

Missingness		VWLM	GLM+fPCA	GLM+SAR	SGPP
10%	$j = 1$	0.5617	0.3203	0.4843	0.1707
	$j = 2$	0.6162	0.3611	0.5342	0.1966
30%	$j = 1$	0.5552	0.3189	0.4749	0.1736
	$j = 2$	0.6219	0.3700	0.5458	0.2094
50%	$j = 1$	0.5606	0.3205	0.4862	0.1837
	$j = 2$	0.6212	0.3707	0.5424	0.2181

Lateral ventricle surfaces

- Applied SGPP to the surface data of the left lateral ventricle
- 43 infants (23 males and 20 females) at age 1
- $\mathbf{x}_i = (1, \mathrm{G}_i, \mathrm{Gage}_i)^\top$; G_i denotes the gender (1 for female and 0 for male); Gage_i denotes the gestational age of the ith infant
- $\mathrm{Gage}_i \in [234, 295]$ days with mean Gage of 263 days and standard deviation of 12.8 days
- Responses based on the SPHARM-PDM representation of the lateral ventricle surfaces
- Ventricle represented by 1002 location vectors with each location vector consisting of the spatial x, y, z coordinates of the corresponding vertex on the SPHARM-PDM surface

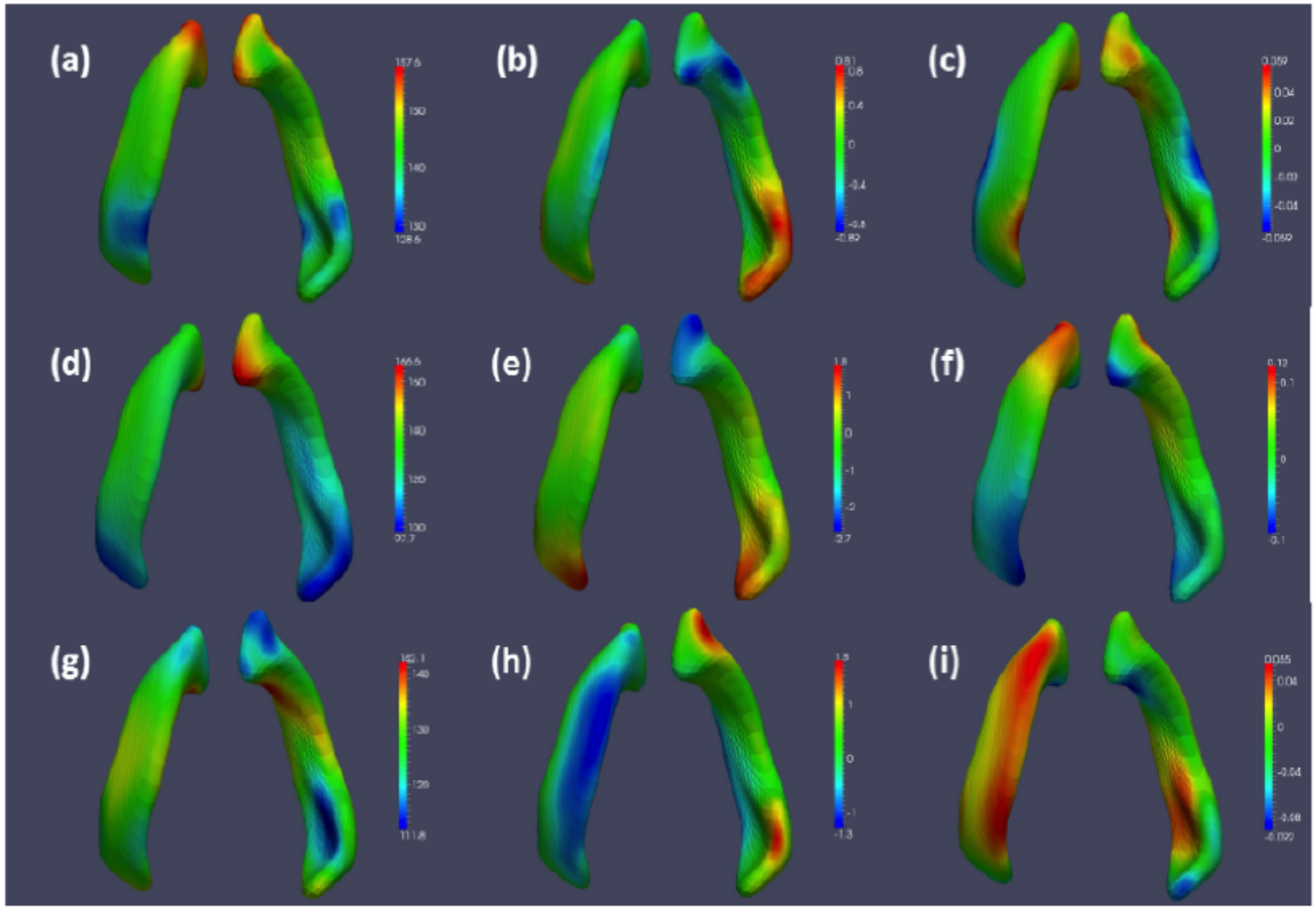

Fig. 5: Results from the surface data of the left lateral ventricle: (a) and (b) $\hat{\beta}_{11}(d)$, $\hat{\beta}_{12}(d)$, and $\hat{\beta}_{13}(d)$ (from left to right); (c) and (d) $\hat{\beta}_{21}(d)$, $\hat{\beta}_{22}(d)$, and $\hat{\beta}_{23}(d)$ (from left to right); (e) and (f) $\hat{\beta}_{31}(d)$, $\hat{\beta}_{32}(d)$, and $\hat{\beta}_{33}(d)$ (from left to right).

Hypothesis testing

Tested the effects of gender and gestational age on the x, y, z coordinates of the left lateral ventricle surface:

$$H_0 : \beta_{j2}(d) = 0 \quad \text{against} \quad \beta_{j2}(d) \neq 0$$

for gender effect and

$$H_0 : \beta_{j3}(d) = 0 \quad \text{against} \quad \beta_{j3}(d) \neq 0$$

for the gestational age across all voxels for $j = 1, 2, 3$.

(Adjusted) $-\log_{10}$(p-values) greater than 1.3 indicate a significant effect at 5% significance level; $-\log_{10}$(p-values) greater than 2 indicate a significant effect at 1% significance level

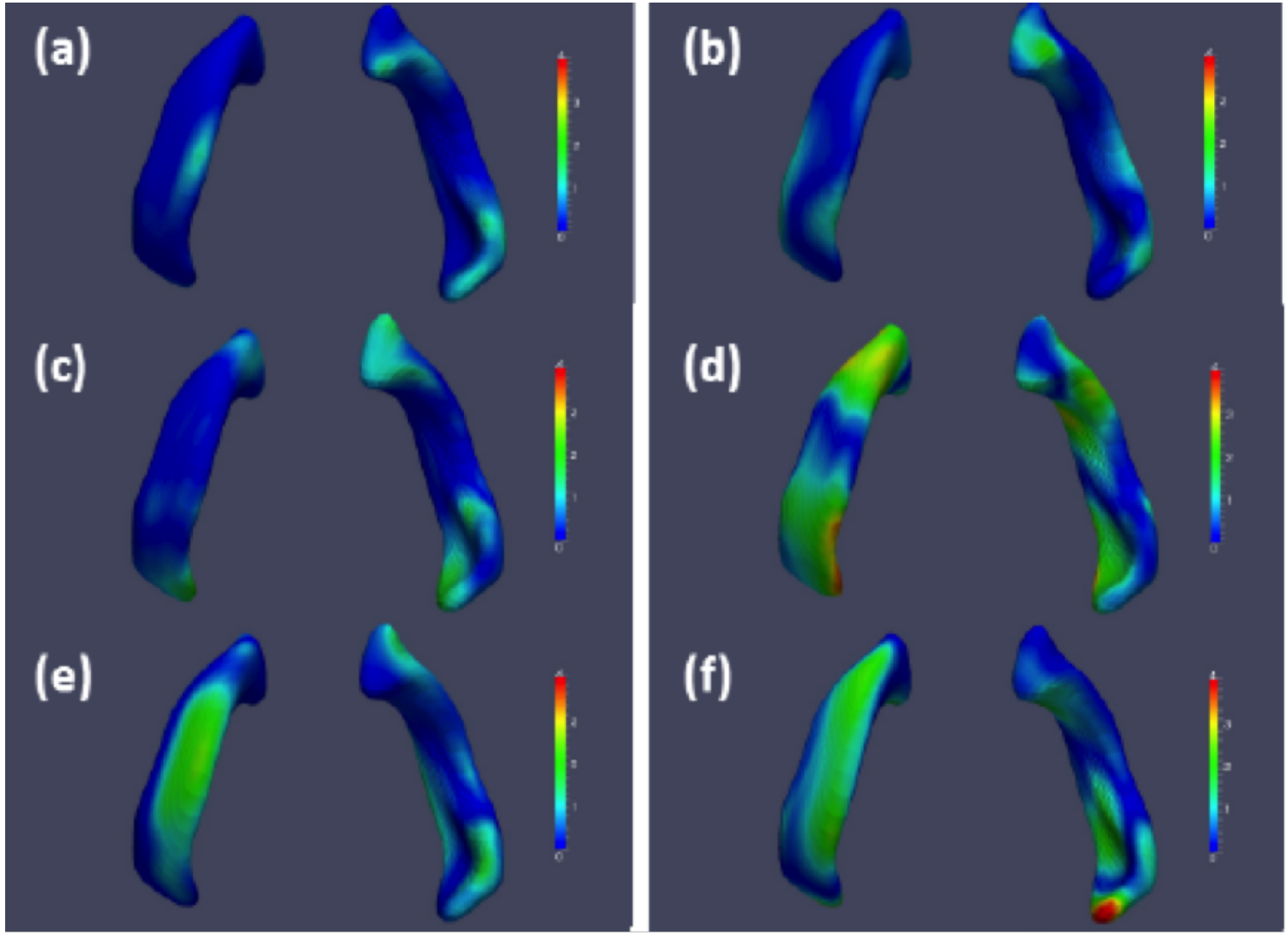

Fig. 6: Raw $-\log_{10}(p)$ maps for testing (a) $H_0 : \beta_{12}(d) = 0$; (b) $H_0 : \beta_{13}(d) = 0$; (c) $H_0 : \beta_{22}(d) = 0$; (d) $H_0 : \beta_{23}(d) = 0$; (e) $H_0 : \beta_{32}(d) = 0$; (f) $H_0 : \beta_{33}(d) = 0$.

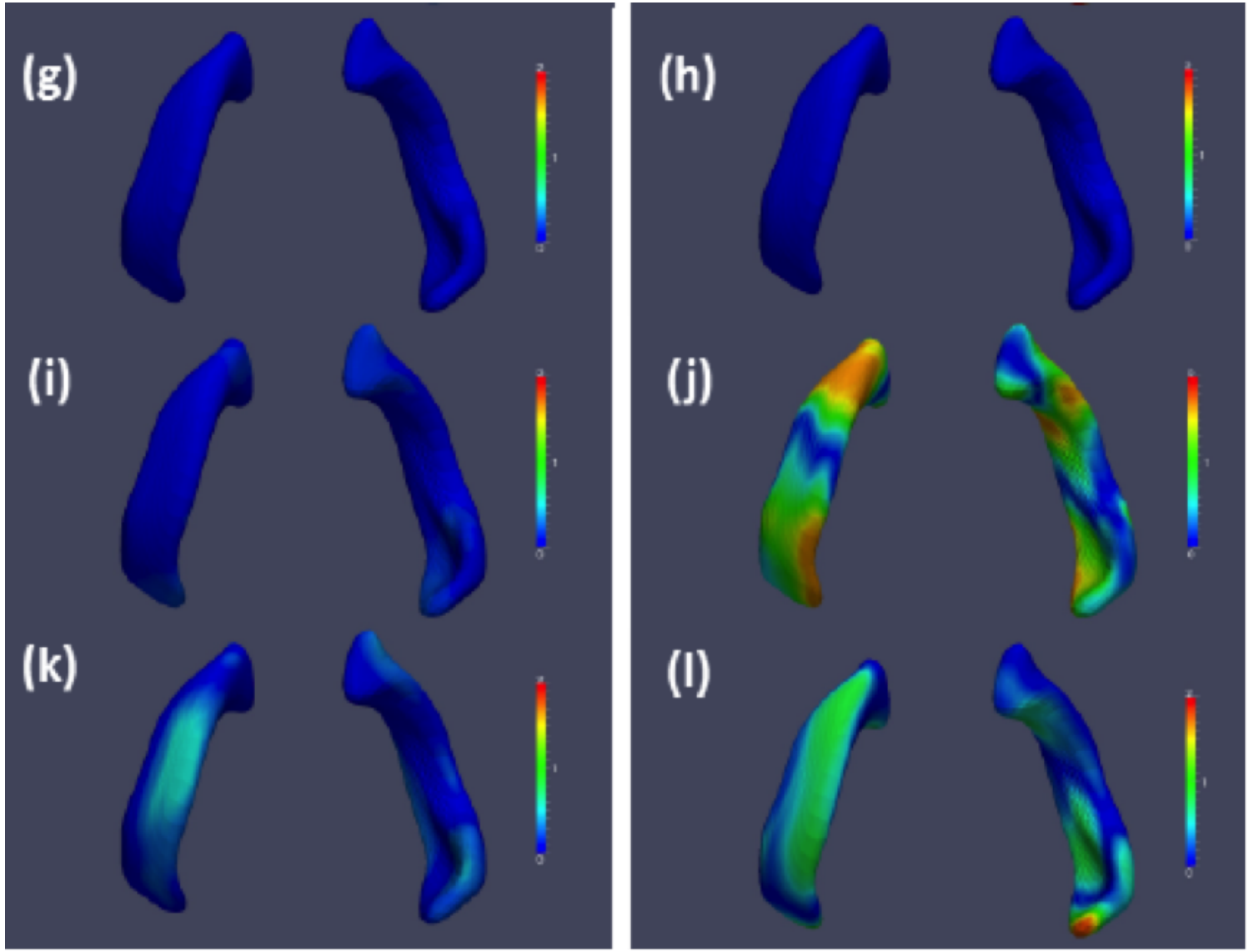

Fig. 7: Corrected $-\log_{10}(p)$ maps for testing (g) $H_0 : \beta_{12}(d) = 0$; (h) $H_0 : \beta_{13}(d) = 0$; (i) $H_0 : \beta_{22}(d) = 0$; (j) $H_0 : \beta_{23}(d) = 0$; (k) $H_0 : \beta_{32}(d) = 0$; (l) $H_0 : \beta_{33}(d) = 0$.

Table 3: rtMSPE for the surface data of the left lateral ventricle

Missingness		VWLM	GLM+fPCA	SGPP
10%	x-coordinate	1.9272	0.9810	0.0738
	y-coordinate	2.2448	1.3455	0.1067
	z-coordinate	2.1554	1.1753	0.0926
30%	x-coordinate	1.9337	1.0197	0.1156
	y-coordinate	2.2655	1.3827	0.1657
	z-coordinate	2.1906	1.2069	0.1446
50%	x-coordinate	1.9263	1.0294	0.1615
	y-coordinate	2.2012	1.3471	0.2204
	z-coordinate	2.1862	1.1830	0.1924

Conclusions

- SGPP essentially an extension of spatial mixed effects models for the analysis of geostatistical data
  - Uses fPCA to estimate spatial basis functions
  - Allows varying regression coefficients across the brain
- Possible extensions to the modelling of longitudinal neuroimaging data & to predict clinical outcomes
- Drawbacks:
  - Estimation procedure is not iterative; the authors should go back to stage (I) after stage (III), but this would likely kill the computation in the fPCA part
  - Real data application is not clear; not easy to interpret what the response is; not clear whether multiplicity adjustment is for voxels, or voxels and coordinate dimension