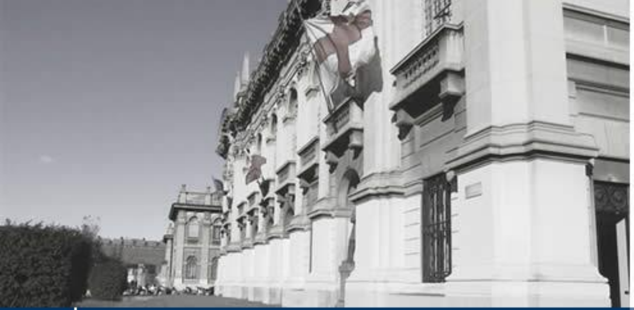


Dote Ricercatore

RegioneLombardia - POLIMI

FIRB 2008

FUTURO IN RICERCA

SNAPLE

POLITECNICO DI MILANO

POLITECNICO DI MILANO 150°

KAUST

8-12 March 2014

Spatial Statistics for Environmental and Energy Challenges

Spatial regression with PDE regularization

Laura M. Sangalli (1)

joint with Laura Azzimonti (1,4), Bree Ettinger (1,5), Fabio Nobile (2), Simona Perotto (1), James O. Ramsay (3), Piercesare Secchi (1)

(1) MOX - Dipartimento di Matematica, Politecnico di Milano, Italy

(2) EPFL École Polytechnique Fédérale de Lausanne

(3) Department of Psychology, McGill University, Montreal, Canada

(4) MOX-Off

(5) Emory University, Atlanta, U.S.A.

MODELLISTICA E CALCOLO SCIENTIFICO

MOX

MODELING AND SCIENTIFIC COMPUTING

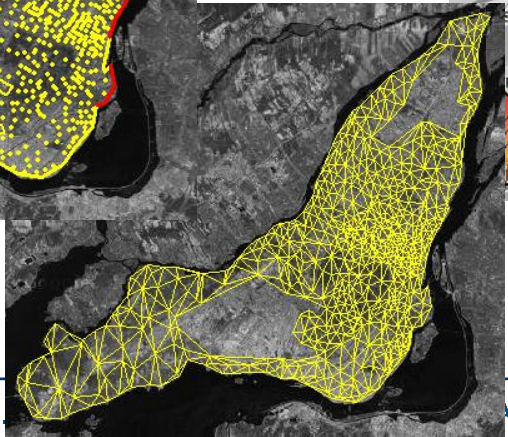
Spatial regression with differential regularization

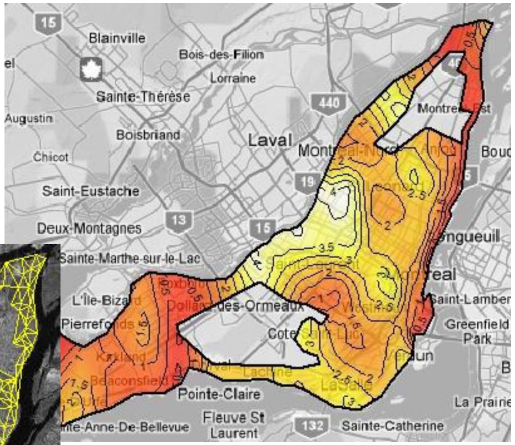
- Problem: *surface estimation* and *spatial field estimation* (spatial regression)

- We interface statistical methodology and numerical analysis techniques and propose

  spatial regression models with partial differential regularization

  ⟶ estimation problem solved via *Finite Elements*





- Can handle data distributed over irregular domains

- Can comply with general conditions at domain boundaries

⟶ Sangalli, Ramsay, Ramsay, 2013, JRSSB

► Can incorporate a piori knowledge about phenomenon under study allowing for very flexible modelling of space variation (anysotropy and non-stationarity)

⟶ Azzimonti, Sangalli, Secchi, Nobile, Domanin, 2013, TechRep

⟶ Azzimonti, Nobile, Sangalli, Secchi, 2013, TechRep

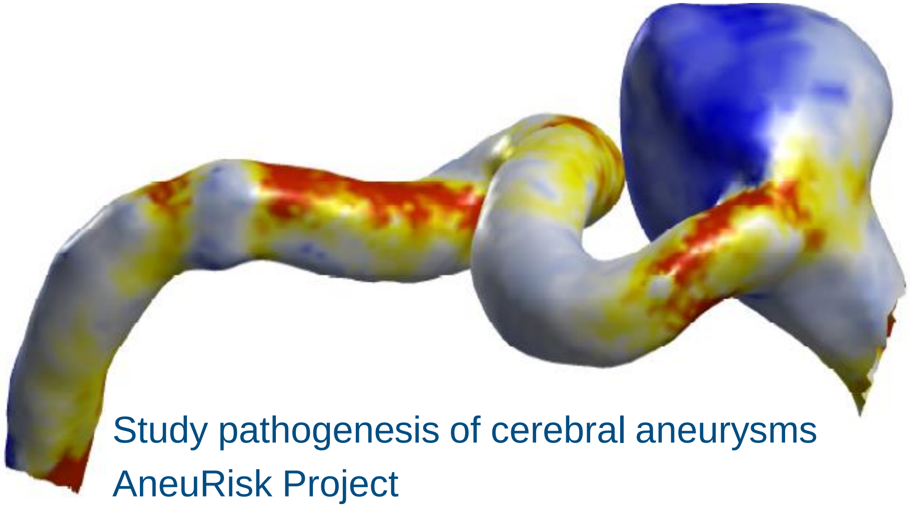
Study pathogenesis of atherosclerotic plaques

MAthematichs for CARotid ENdarterectomy @ MOX

► Can deal with data over bi-dimensional Riemannian manifolds

⟶ Ettinger, Perotto, Sangalli, 2012, TechRep

⟶ Dassi, Ettinger, Perotto, Sangalli, 2013, TechRep



Study pathogenesis of cerebral aneurysms

AneuRisk Project

► Can incorporate a piori knowledge about phenomenon under study allowing for very flexible modelling of space variation (anysotropy and non-stationarity)

⟶ Azzimonti, Sangalli, Secchi, Nobile, Domanin, 2013, TechRep

⟶ Azzimonti, Nobile, Sangalli, Secchi, 2013, TechRep

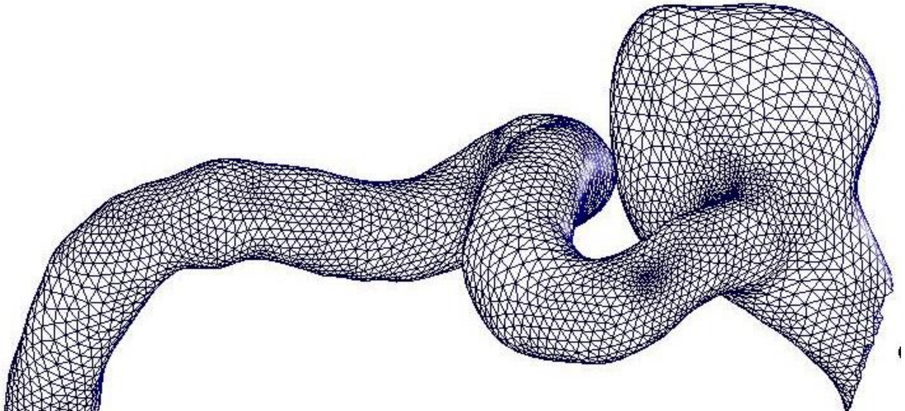
Study pathogenesis of atherosclerotic plaques

MAthematichs for CARotid ENdarterectomy @ MOX

► Can deal with data over bi-dimensional Riemannian manifolds

⟶ Ettinger, Perotto, Sangalli, 2012, TechRep

⟶ Dassi, Ettinger, Perotto, Sangalli, 2013, TechRep



Study pathogenesis of cerebral aneurysms

AneuRisk Project

- Data: $\Omega \subset \mathbb{R}^2$: a region of interest, bounded, with $\partial\Omega \in \mathcal{C}^2$

  for $i = 1, \dots, n$

  - $\mathbf{p}_i = (x_i, y_i) \in \Omega$
  - z_i: a real valued variable of interest observed $\mathbf{p}_i$
  - $\mathbf{w}_i = (w_{i1}, \dots, w_{iq})^t$: a q-vector of covariates associated to z_i

- Generalized Additive Model:

$$z_i = \mathbf{w}_i^t \boldsymbol{\beta} + f(\mathbf{p}_i) + \epsilon_i \qquad i = 1, \dots, n$$

  - $\epsilon_i, i = 1, \dots, n$, i.i.d. mean 0 and variance σ^2
  - $\boldsymbol{\beta} \in \mathbb{R}^q$
  - $f : \Omega \to \mathbb{R}$

Sangalli et al. 2013, JRSSB

- Estimate $\boldsymbol{\beta}$ and f minimizing

$$J_\lambda(\boldsymbol{\beta}, f) = \sum_{i=1}^{n} \left(z_i - \mathbf{w}_i^t \boldsymbol{\beta} - f(\mathbf{p}_i)\right)^2 + \lambda \int_\Omega (\Delta f)^2 d\Omega$$

$$\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

Inclusion of simple Partial Differential Equations (PDE) in statistical models:

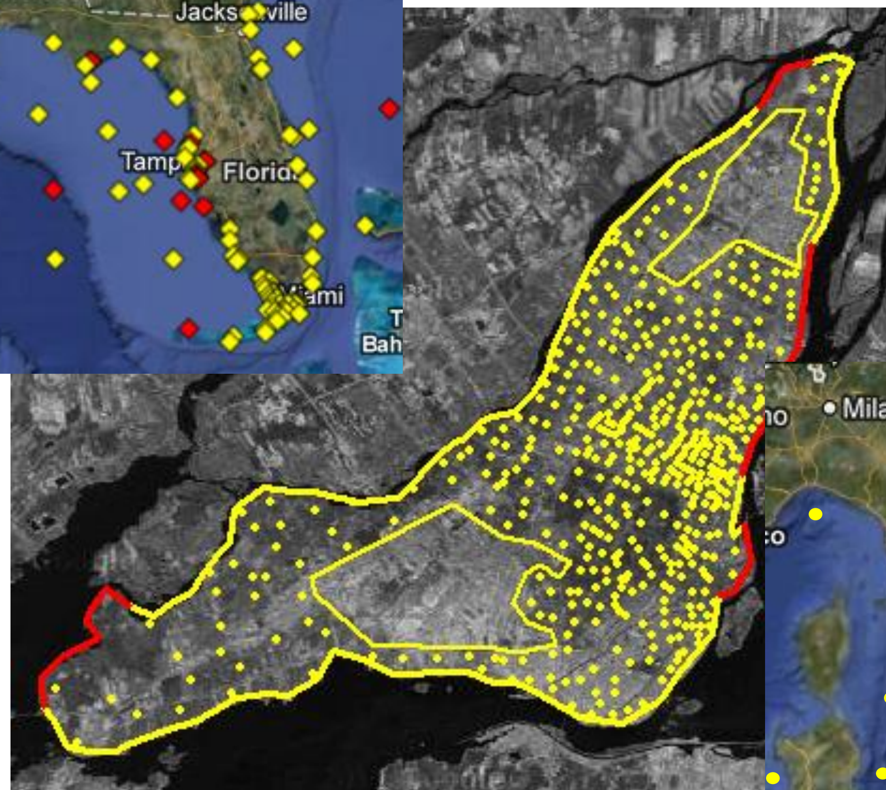
- Thin-plate splines (Wahba,1990; Stone, 1988)

$$\sum_{i=1}^{n} \left(z_i - f(\mathbf{p}_i)\right)^2 + \int_{\mathbb{R}^2} \left(\frac{\partial^2 f}{\partial x^2}\right)^2 + 2\frac{\partial^2 f}{\partial x \partial y} + \left(\frac{\partial^2 f}{\partial y^2}\right)^2$$

- Bivariate Splines (Guillas and Lai, 2010)
- FEL-splines (Ramsay, 2002), Soap-film smoothing (Wood *et al.*, 2008): *irregular domains*
- Stochastic PDE: Lindgren *et al.* (2011), Bayesian inverse problems: Stuart (2010)
- Data Assimilation in Inverse Problems

Irregularly shaped domains

Buoy data
(National Oceanic and Atmospheric Administration www.ndbc.noaa.gov)
→ Parnigoni Master thesis 2013

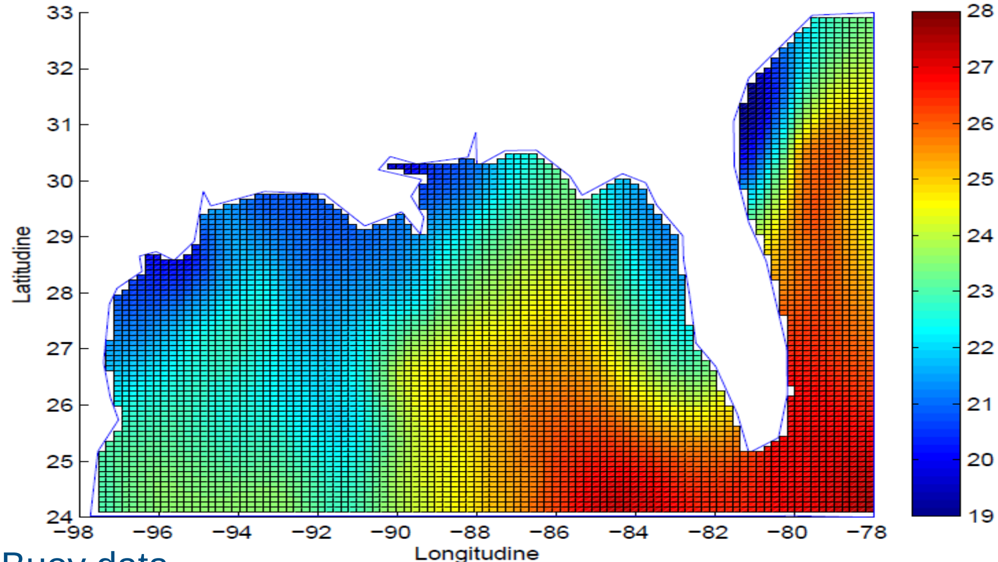


Census Canada data

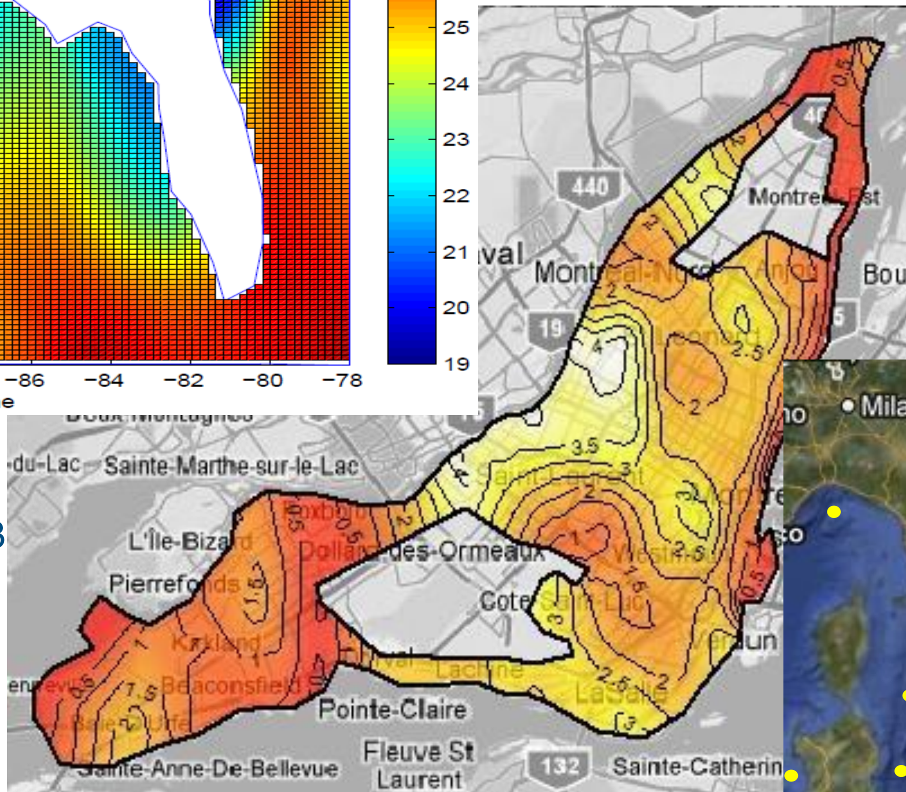
Fisheries data (NOAA)

Classical methods cannot handle irregular shaped domains
Global basis functions, Covariance (stationarity, isotropy)

Irregularly shaped domains



Buoy data
(National Oceanic and Atmospheric Administration www.ndbc.noaa.gov)
→ Parnigoni Master thesis 2013



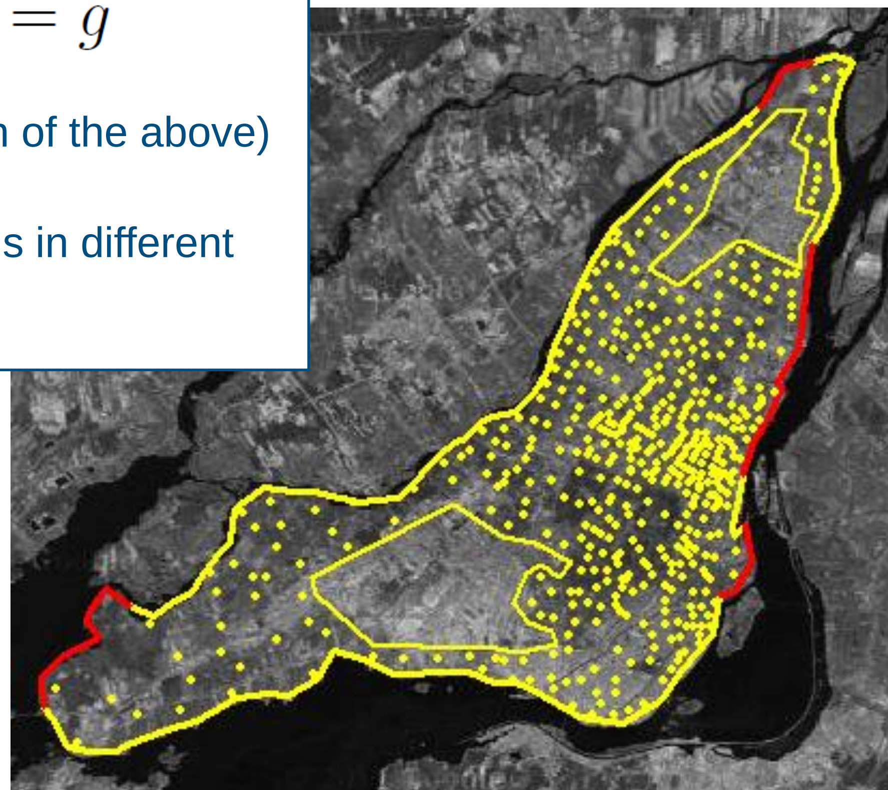
Census Canada data

Fisheries data (NOAA)

Classical methods cannot handle irregular shaped domains
Global basis functions, Covariance (stationarity, isotropy)

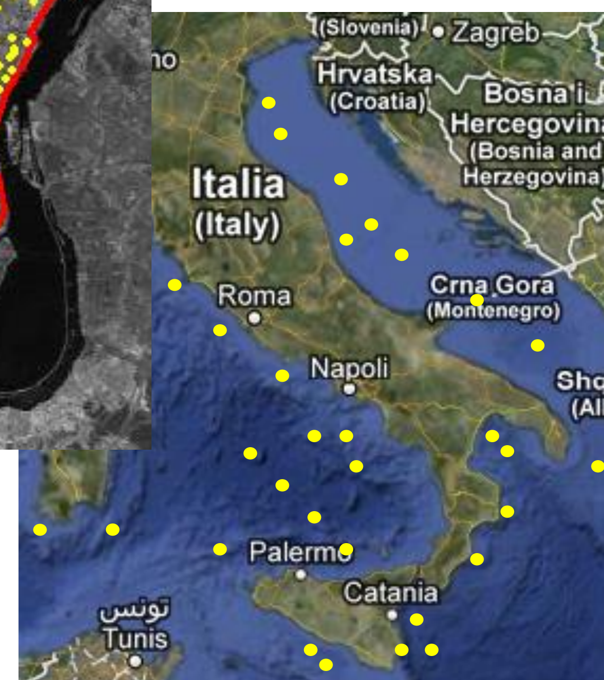
Boundary conditions

- Dirichlet $f|_{\partial\Omega} = g$
- Neumann $\partial_\nu f|_{\partial\Omega} = g$
- Robin (linear combination of the above)
- Mixed (different conditions in different parts of the boundary)



Census Canada data

Fisheries data (NOAA)



A priori information

Azzimonti et al., 2013a, TechRep
Azzimonti et al., 2013b, TechRep

Incorporating a priori information:
using PDE to model space variation
of the phenomenon

log-likelihood
data fidelity

prior
model fidelity

$$J_\lambda(\boldsymbol{\beta}, f) = \sum_{i=1}^{n} \left(z_i - \mathbf{w}_i^t \boldsymbol{\beta} - f(\mathbf{p}_i)\right)^2 + \lambda \int_\Omega (Lf - u)^2 d\Omega$$

more complex partial differential operator
(linear second order elliptic PDE)

$$Lf = -div(K \nabla f) + \mathbf{b} \cdot \nabla f + cf$$

Problem specific a priori information →
(physics, mechanics, chemistry, morphology)

spatially varying

- ► Diffusion tensor field: *non-stationary anisotropic diffusion*
- ► Transport vector field: *non-stationary directional smoothing*
- ► Reaction term: *non-stationary shrinking effect*

- ► PDEs are commonly used to describe *complex phenomena behaviors* in many fields of engineering and sciences
- ► model space variation

A priori information

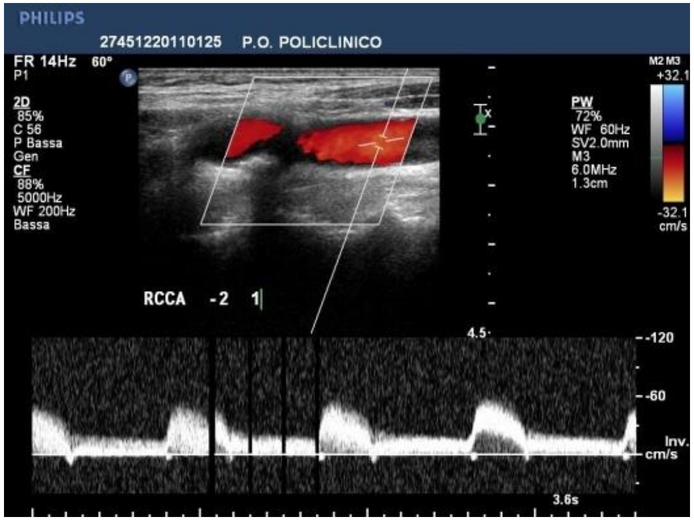
$$J_\lambda(\boldsymbol{\beta}, f) = \sum_{i=1}^{n} \left(z_i - \mathbf{w}_i^t \boldsymbol{\beta} - f(\mathbf{p}_i)\right)^2 + \lambda \int_\Omega (Lf - u)^2 d\Omega$$

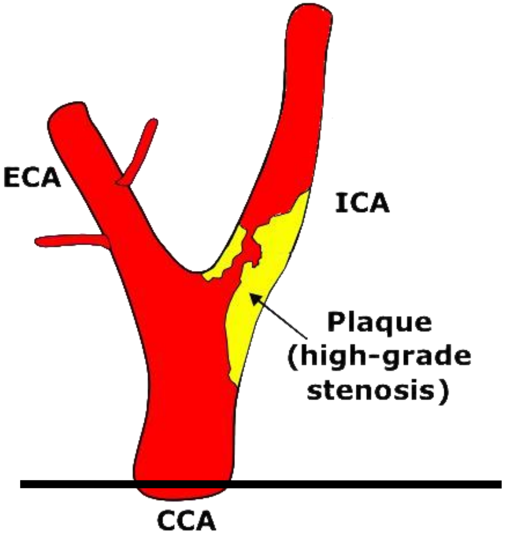
Buoy data

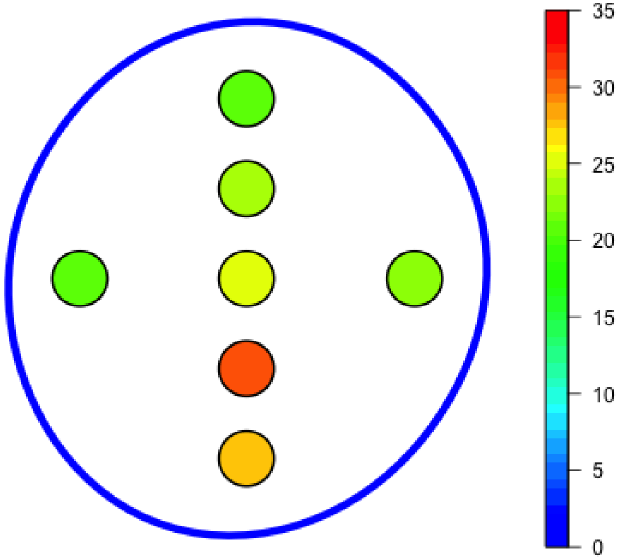
Prior (Gulf Stream)

Motivating applied problem: *MACAREN@MOX*

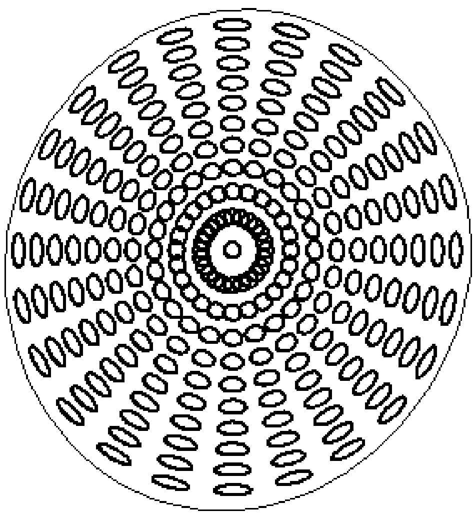
ECD measurements over 7 beams

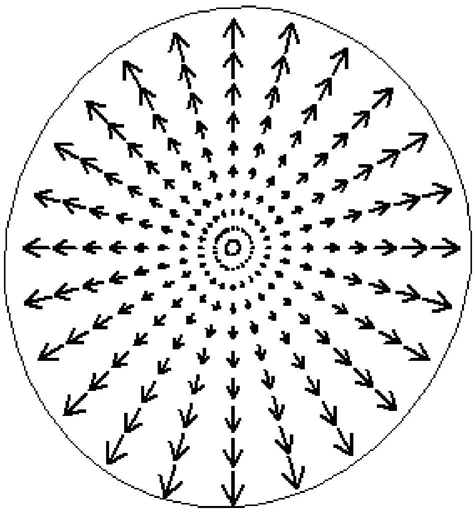






A priori knowledge described by a PDE

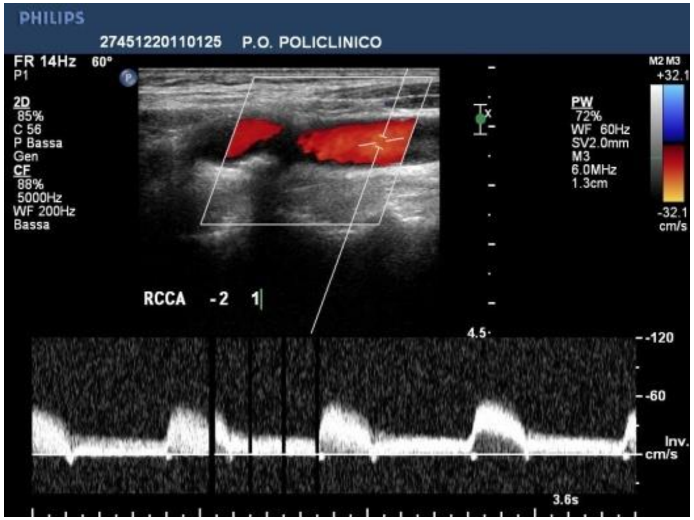


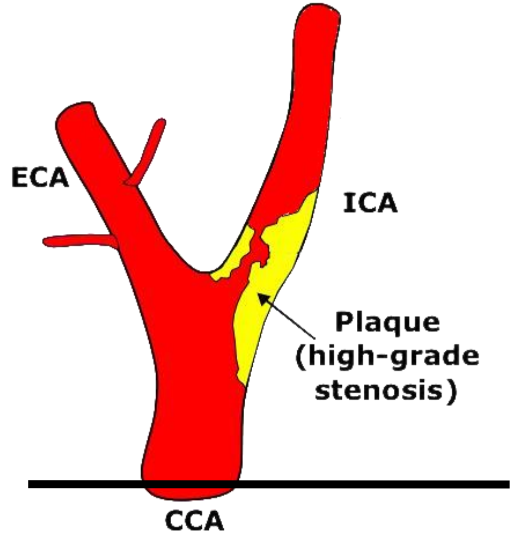


Mathematichs for CARotid ENdarterectomy @ MOX
MACAREN@MOX

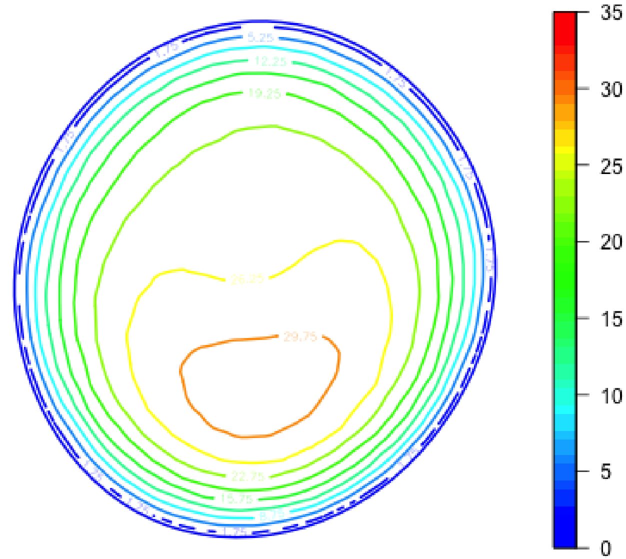
Azzimonti et al., 2013a, TechRep

Motivating applied problem: *MACAREN@MOX*

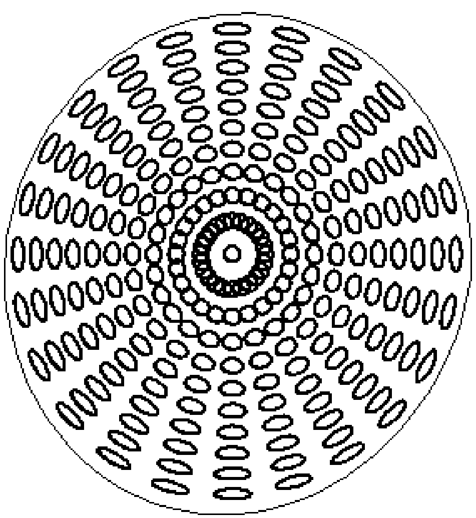


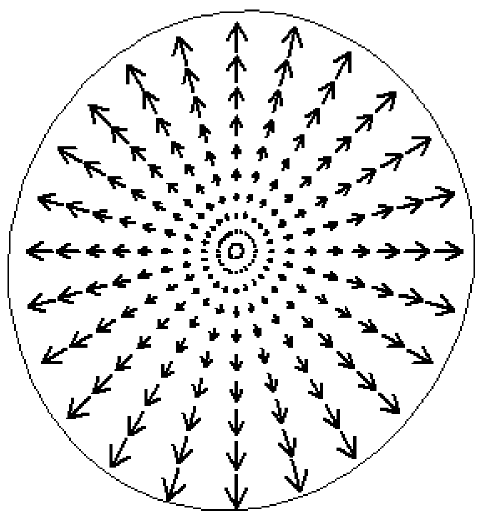


Blood flow velocity field estimate

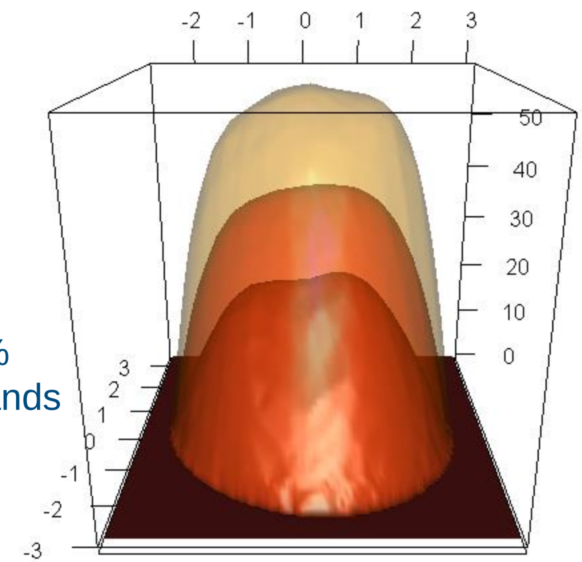


A priori knowledge described by a PDE





pointwise 95% confidence bands



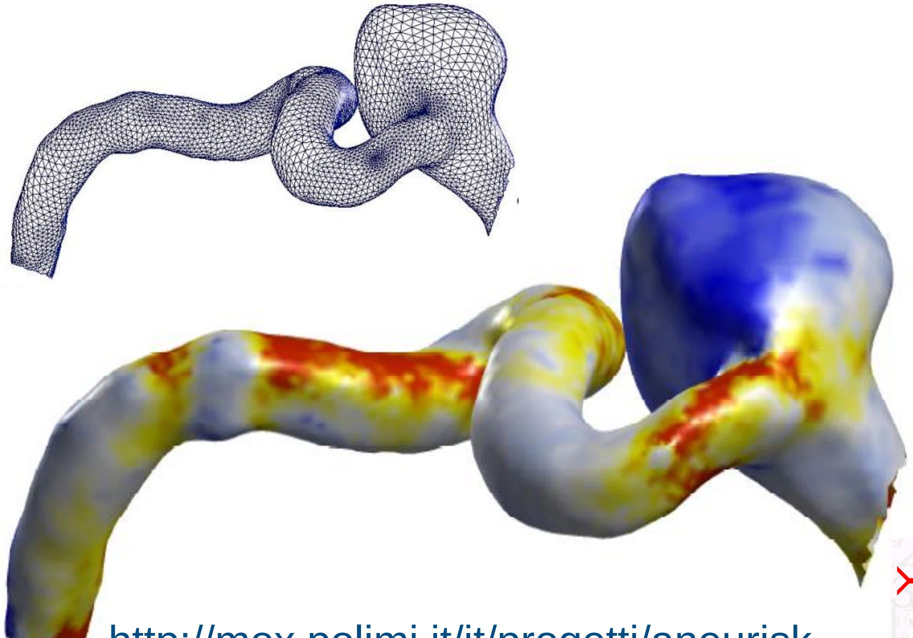
Manifold domains

Ettinger et al. 2012, TechRep

Spatial regression over bi-dimensional Riemannian manifold domains

$$J_{\Gamma,\lambda}(\boldsymbol{\beta}, f) = \sum_{i=1}^{n} \left(z_i - \mathbf{w}_i^t \boldsymbol{\beta} - f(\mathbf{x}_i)\right)^2 + \lambda \int_{\Gamma} (\Delta_{\Gamma} f(\mathbf{x}))^2 d\mathbf{x}$$

Laplace-Beltrami operator associated to Γ



Γ

Γ: surface embedded in $\mathbb{R}^3$

$\mathbf{x}_i \in \Gamma$

$f : \Gamma \to \mathbb{R}$

http://mox.polimi.it/it/progetti/aneurisk

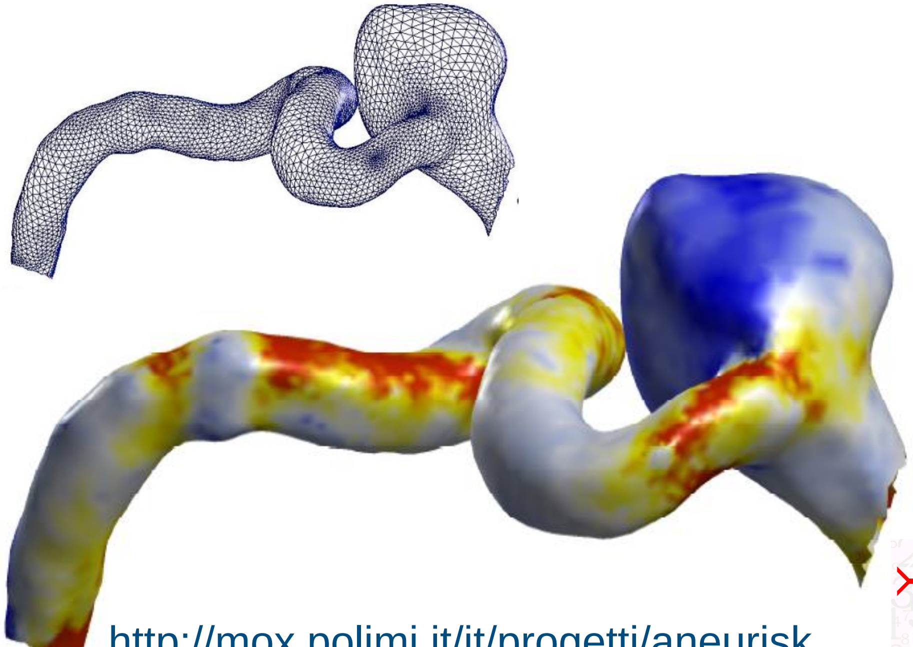
Manifold domains

Ettinger et al. 2012, TechRep

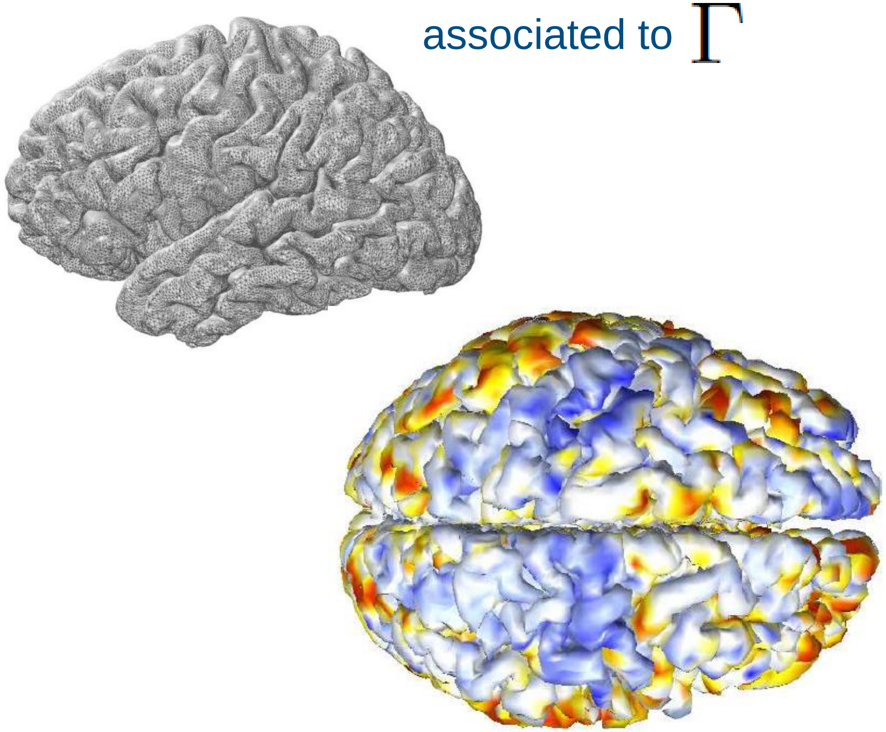
Spatial regression over bi-dimensional Riemannian manifold domains

$$J_{\Gamma,\lambda}(\boldsymbol{\beta}, f) = \sum_{i=1}^{n} \left(z_i - \mathbf{w}_i^t \boldsymbol{\beta} - f(\mathbf{x}_i)\right)^2 + \lambda \int_{\Gamma} (\Delta_\Gamma f(\mathbf{x}))^2 d\mathbf{x}$$

Laplace-Beltrami operator associated to Γ



Γ



http://mox.polimi.it/it/progetti/aneurisk

Motivating applied problem: AneuRisk project

€

ANEURISK

A CONJECTURE

The pathogenesis of cerebral aneurysms is conditioned by the geometry of the cerebral vessels through its effects on blood fluid dynamics

Statistics

Numerical Analysis

BioEngineering

Computer science

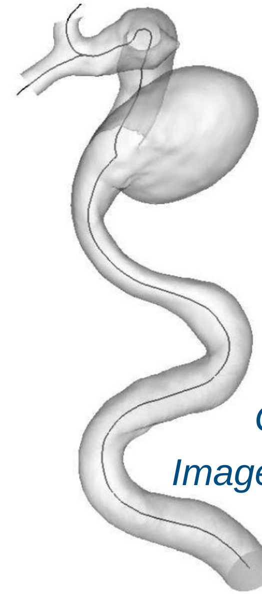
Neurosurgery

Neuroradiology

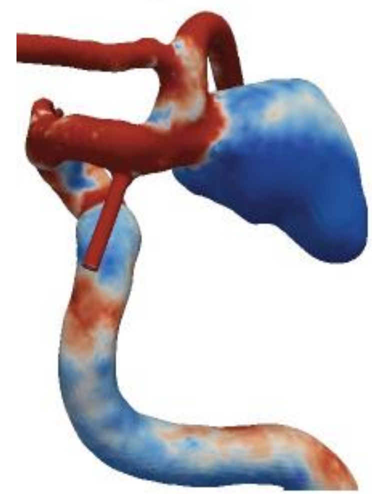
Data: 3D-angiographies

Sangalli Secchi Vantini Veneziani JASA 2009
Sangalli Secchi Vantini Veneziani JRSSC 2009
Sangalli Secchi Vantini Vitelli CSDA 2010

Observational Study conducted at Ospedale Ca' Granda Niguarda, Milano, relative to 65 patients hospitalized from 2002 to 2005.



Geometrical data
Image reconstructions

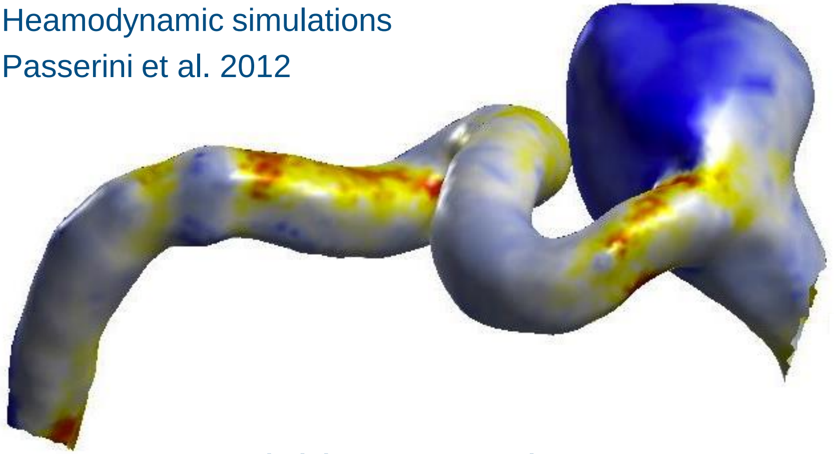


Hemodynamic data
Computational Fluid Dynamics

Ettinger et al., 2012, TechRep

Heamodynamic simulations

Passerini et al. 2012



Object Oriented Data Analysis

- Nearest Neighbor Averaging
  - Hagler, Saygin, Sereno, 2006, *NeuroImage*
- Heat Kernel Smoothing
  - Chung et al., 2005, *NeuroImage*
- Methods for data over spheres, hyperspheres and other manifolds
  - Baramidze, Lai, Shum, 2006, *SIAM J.S.C.*
  - Wahba,1981, *SIAM J.S.C.*
  - Lindgren, Rue, Lindstrom, 2011, *JRSSB*
  - Yun, 2011, *Scandinavian*
  - Gneiting, 2013, *Bernoulli*

- $\Gamma \subset \mathbb{R}^3$ - a non-planar surface domain

  Artery wall

- $\{\mathbf{x}_i = (x_{1i}, x_{2i}, x_{3i}) \in \Gamma\}$ - data locations

- $z_i \in \mathbb{R}$ - variable of interest observed at $\mathbf{x}_i$

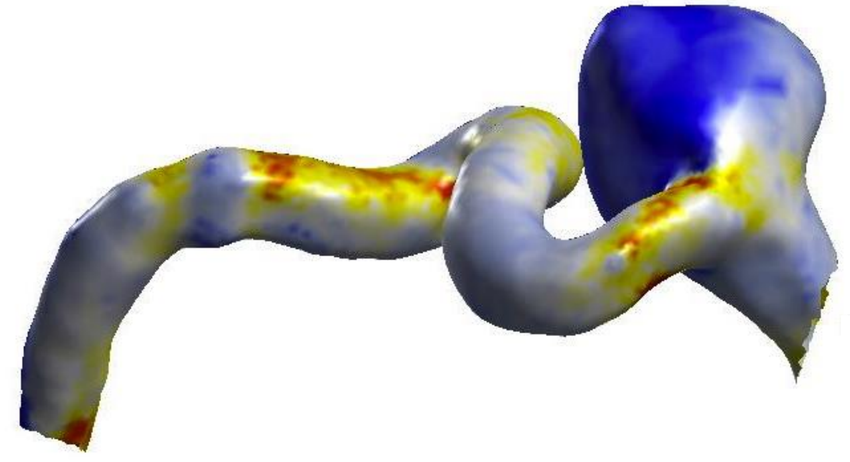
  Wall shear stress modulus at systolic peak

- $\mathbf{w}_i = (w_{i1}, \ldots, w_{iq}) \in \mathbb{R}^q$ - space varying covariates

  Local curvature of vessel wall

  Curvature of vessel

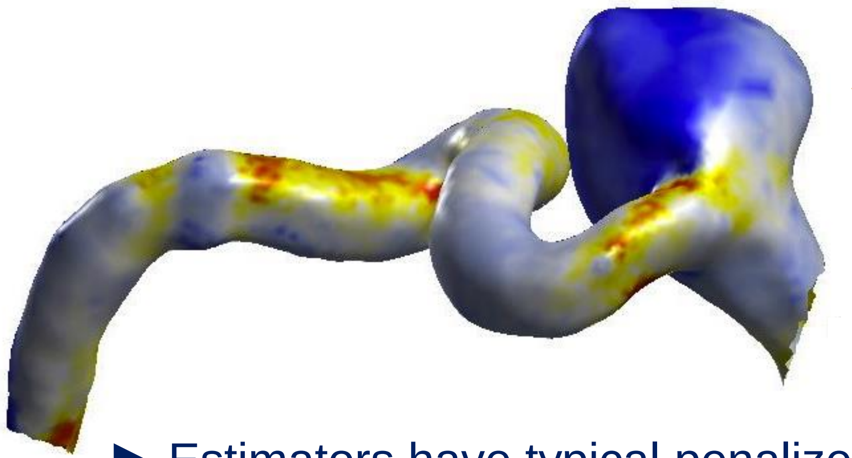
  Local radius of the vessel

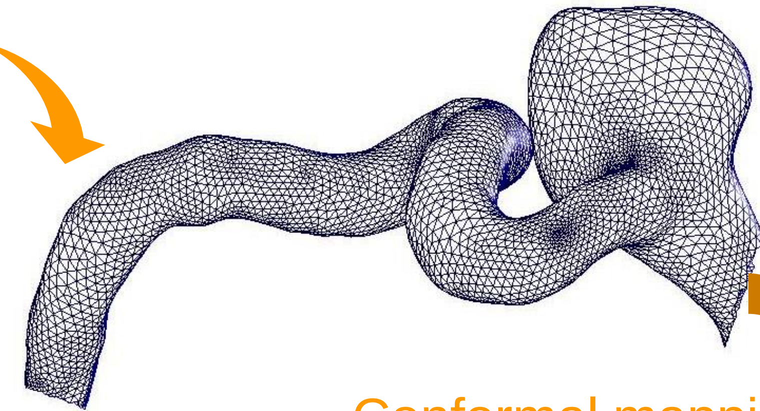


$$z_i = \mathbf{w}_i^t \boldsymbol{\beta} + f(\mathbf{x}_i) + \epsilon_i \qquad\qquad i = 1, \ldots, n$$

$$J_{\Gamma,\lambda}(\boldsymbol{\beta}, f) = \sum_{i=1}^{n} \left(z_i - \mathbf{w}_i^t \boldsymbol{\beta} - f(\mathbf{x}_i)\right)^2 + \lambda \int_\Gamma (\Delta_\Gamma f(\mathbf{x}))^2 d\mathbf{x}$$

Spatial regression over Riemannian manifolds

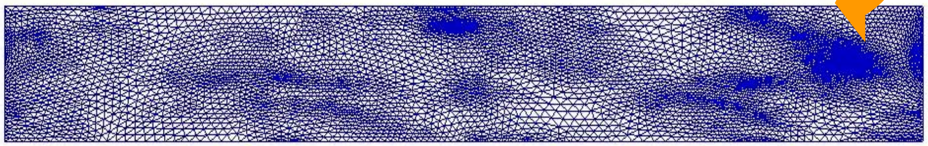


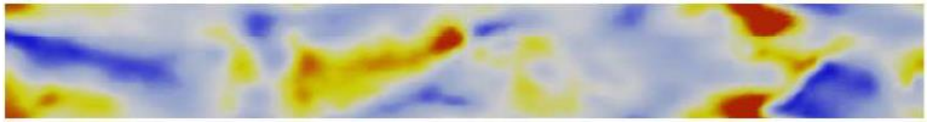


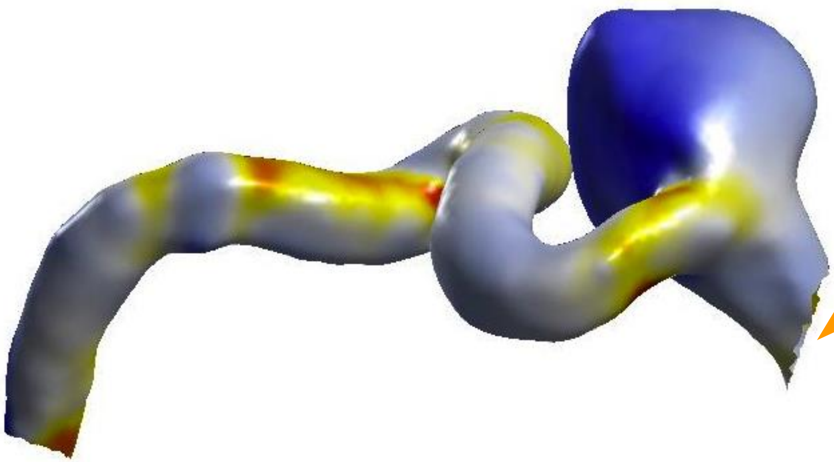
- ► Estimators have typical penalized regression form
- ► Linear in observed data values
- ► Classical inferential tools

Conformal mapping

(fully encode information about complex 3D geometry)





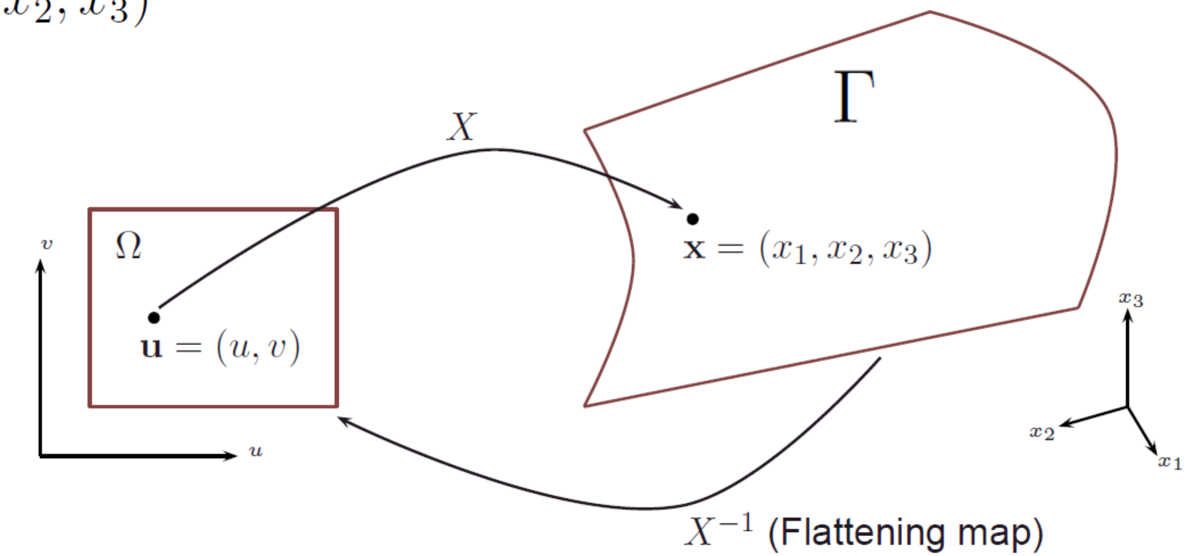


Equivalent estimation problem on planar domain

⟶ Extend method for planar domains

▷ $X : \Omega \to \Gamma$ ($\Omega :$ open, convex, bounded set in $\mathbb{R}^2$)

$$\mathbf{u} = (u, v) \mapsto \mathbf{x} = (x_1, x_2, x_3)$$



- $\frac{\partial X}{\partial u}(\mathbf{u})$, $\frac{\partial X}{\partial v}(\mathbf{u})$: column vectors

- space varying metric tensor:

$$G(\mathbf{u}) \;:=\; \nabla X(\mathbf{u})' \nabla X(\mathbf{u}) = \begin{pmatrix} \left\| \frac{\partial X}{\partial u}(\mathbf{u}) \right\|^2 & \langle \frac{\partial X}{\partial u}(\mathbf{u}), \frac{\partial X}{\partial v}(\mathbf{u}) \rangle \\ \langle \frac{\partial X}{\partial u}(\mathbf{u}), \frac{\partial X}{\partial v}(\mathbf{u}) \rangle & \left\| \frac{\partial X}{\partial v}(\mathbf{u}) \right\|^2 \end{pmatrix}$$

- $\mathcal{W}(\mathbf{u}) := \sqrt{\det(G(\mathbf{u}))}\,;$ $\qquad \boxed{\mathcal{W}(\mathbf{u}) d\mathbf{u} = d\mathbf{x}}$

- $\mathbf{K}(\mathbf{u}) = \mathcal{W}(\mathbf{u})\, G^{-1}(\mathbf{u})$

- For $f \circ X \in \mathcal{C}^2(\Omega)$, $\qquad \mathbf{u} = X^{-1}(\mathbf{x})$

$$\nabla_\Gamma f(\mathbf{x}) \;=\; \nabla X(\mathbf{u})\; G^{-1}(\mathbf{u})(\nabla f(X(\mathbf{u})))$$

$$\boxed{\Delta_\Gamma f(\mathbf{x})} \;=\; \mathrm{div}_\Gamma(\nabla_\Gamma f(X(\mathbf{u}))) = \boxed{\frac{1}{\mathcal{W}(\mathbf{u})} \mathrm{div}\big(\mathbf{K}(\mathbf{u}) \nabla f(X(\mathbf{u}))\big)}$$

▷ $H^m_{n0,\mathbf{K}}(\Omega) = \{h \in H^m(\Omega) \ : \ \mathbf{K}\nabla h \cdot n = 0 \text{ on } \partial\Omega\} \subset H^m(\Omega)$

▷ **Equivalent estimation problem over the planar domain Ω**

Find $\boldsymbol{\beta} \in \mathbb{R}^q$ and f with $(f \circ X) \in H^2_{n0,\mathbf{K}}(\Omega)$ that minimizes

$$J_{\Omega,\lambda}(\boldsymbol{\beta}, f \circ X) = \sum_{i=1}^{n} \left(z_i - \mathbf{w}_i'\boldsymbol{\beta} - f(X(\mathbf{u}_i))\right)^2 + \lambda \boxed{\int_\Omega \frac{1}{W}\Big(\operatorname{div}(\mathbf{K}\nabla(f \circ X))\Big)^2 d\Omega}$$

where $\mathbf{u}_i = X^{-1}(\mathbf{x}_i)$

- $z_i = \mathbf{w}_i^t \boldsymbol{\beta} + f(\mathbf{x}_i) + \epsilon_i\,,\ i = 1, \ldots, n,$ $\qquad \mathbf{z} = W\boldsymbol{\beta} + \mathbf{f}_n + \boldsymbol{\epsilon}$

▷ $\mathbf{z} := (z_1, \ldots, z_n)^t$ $\qquad$ ▷ $\mathbf{f}_n := \big(f(\mathbf{x}_1), \ldots, f(\mathbf{x}_n)\big)^t = \big(f(X(\mathbf{u}_1)), \ldots, f(X(\mathbf{u}_n))\big)^t$

▷ $W := \begin{bmatrix} \mathbf{w}_1^t \\ \vdots \\ \mathbf{w}_n^t \end{bmatrix}$ $\qquad H := W(W^tW)^{-1}W^t \qquad Q := I - H$

Proposition. The estimators $\hat{\boldsymbol{\beta}} \in \mathbb{R}^q$ and $\hat{f} \in H^2_{n0,\mathbf{K}}(\Omega)$ exist unique

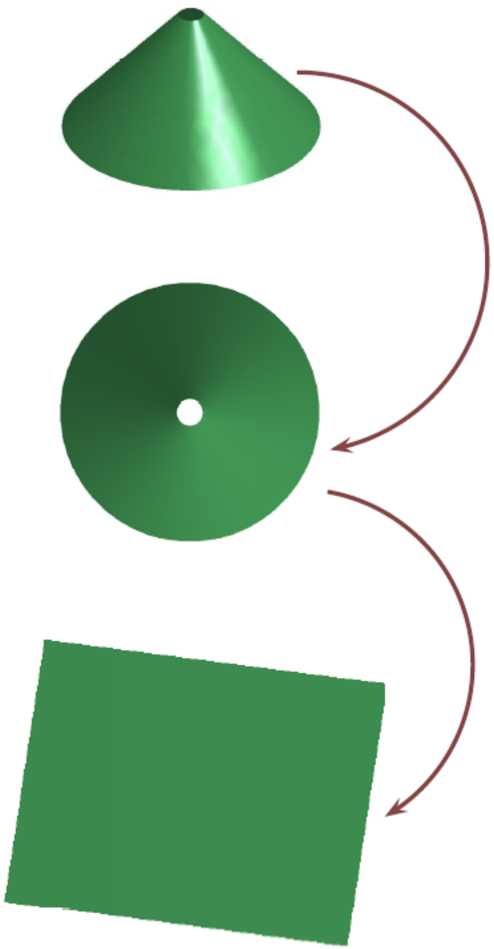
($\star$) $\hat{\boldsymbol{\beta}} = (W^tW)^{-1}W^t(\mathbf{z} - \hat{\mathbf{f}}_n)$

($\star\star$) $\hat{f}$ satisfies

$$\boldsymbol{\mu}_n^t Q \hat{\mathbf{f}}_n + \lambda \int_\Omega \frac{1}{\mathcal{W}} \Big(\mathrm{div}(\mathbf{K}\nabla(\mu \circ X))\Big)\Big(\mathrm{div}\Big(\mathbf{K}\nabla(\hat{f} \circ X)\Big)\Big) d\Omega = \boldsymbol{\mu}_n^t Q \mathbf{z}$$

for any μ defined on Γ such that $\mu \circ X \in H^2_{n0,\mathbf{K}}(\Omega)$.

Conformal parametrization



$$\begin{cases} -\Delta_\Gamma u = 0 \text{ on } \Gamma \\ u = 0 \text{ on } \sigma_0 \\ u = 1 \text{ on } \sigma_1 \end{cases}$$

$$\boxed{E_D(u) = \frac{1}{2}\int_\Gamma \|\nabla_\Gamma u\|^2 d\Gamma}$$

$$\begin{cases} -\Delta_\Gamma v = 0 \text{ on } \Gamma \\ v(\zeta) = \int_{\zeta_0}^{\zeta} \frac{\partial u}{\partial \nu}\, ds \ \text{ on } B \end{cases}$$

$$\boxed{E_D(v) = \frac{1}{2}\int_\Gamma \|\nabla_\Gamma v\|^2 d\Gamma}$$

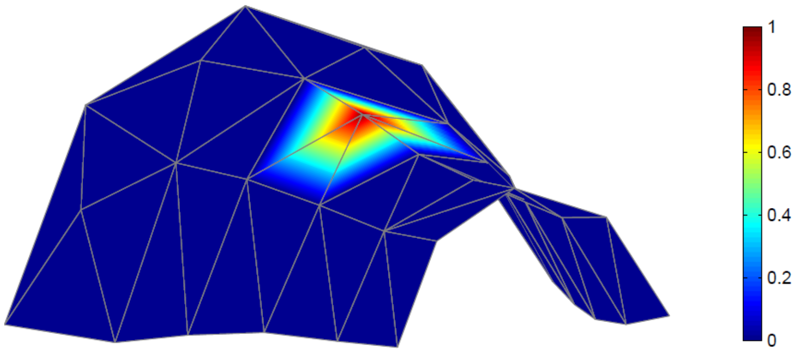
Haker *et al*, 2000, *IEEE Trans. Med. Imag*

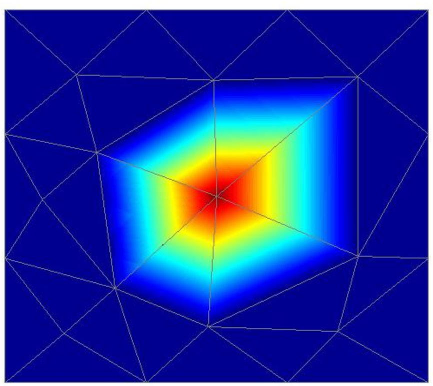
Infinite-dimensional problem

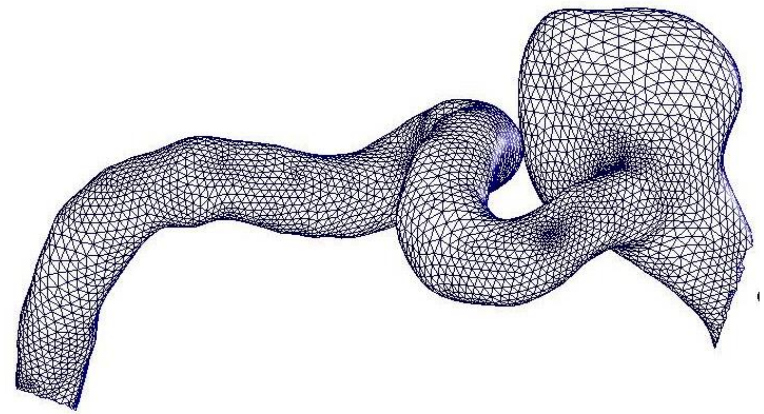
Basis expansion

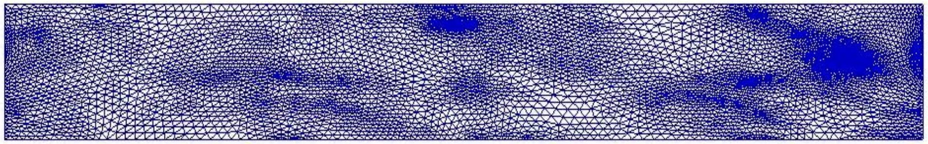
Finite Elements

Finite-dimensional problem





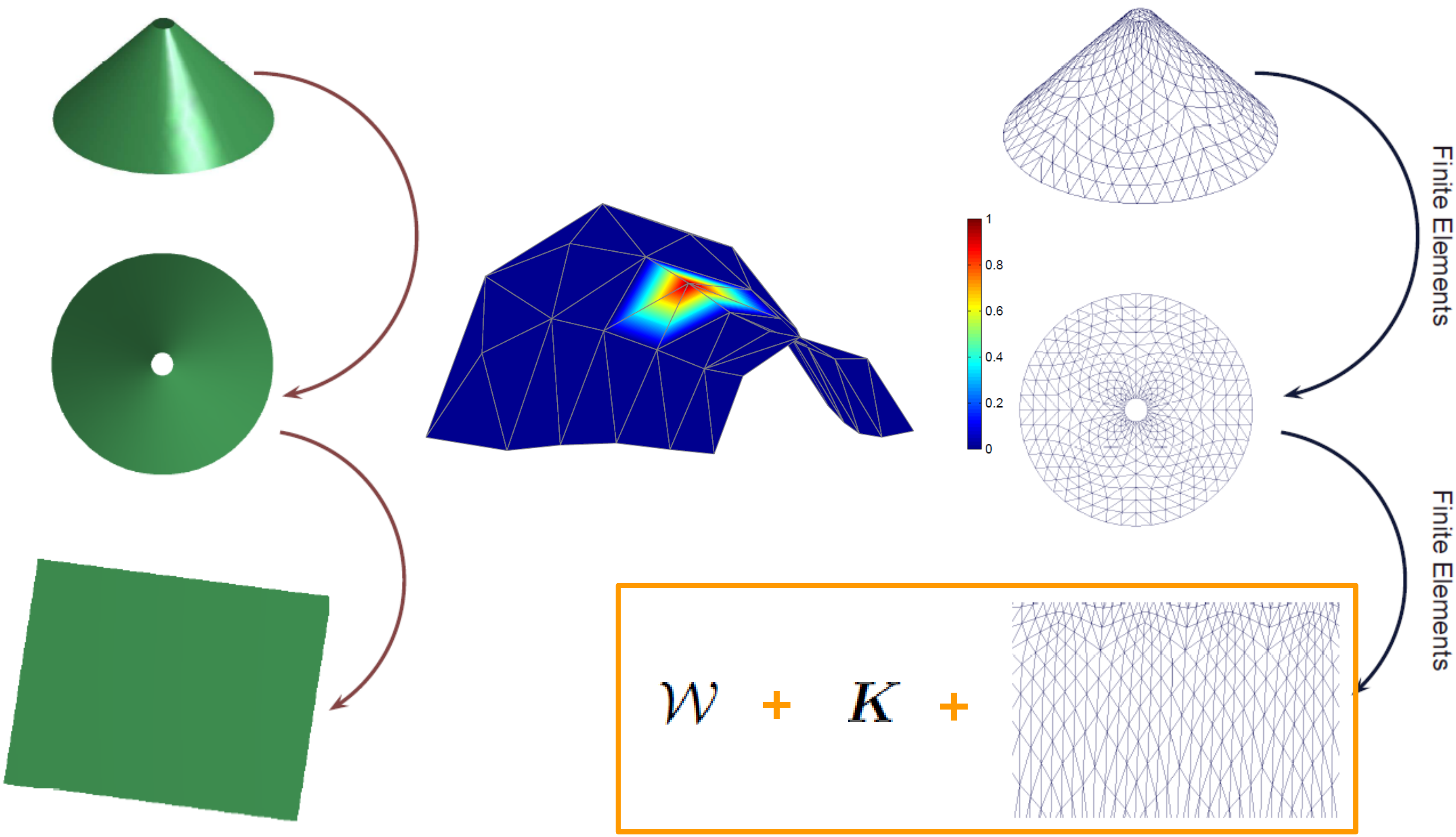




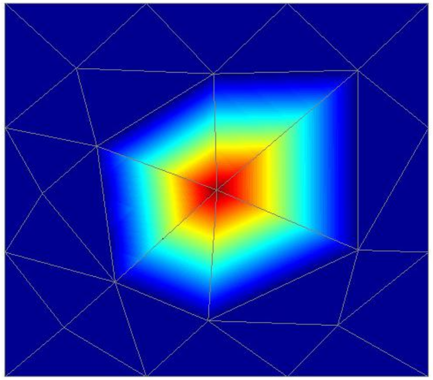
Finite element analysis has been mainly developed and used in engineering applications, to solve partial differential equations

Finite element space: space of continuous piecewise-polynomial surfaces over a triangulation $\mathcal{T}$ of the domain

Conformal parametrization



Haker *et al*, 2000, *IEEE Trans. Med. Imag*



- $\{\boldsymbol{\xi}_1, ..., \boldsymbol{\xi}_K\}$: nodes of planar triangulation $\mathcal{T}$

- $\Omega_\mathcal{T}$: planar triangulated domain; $H^1_\mathcal{T}(\Omega)$: finite element space

- $\boldsymbol{\psi} = (\psi_1, \ldots, \psi_K)^t$: finite element basis $\qquad \Psi = \{\Psi\}_{ij} := \psi_j(\mathbf{p}_i)$

- for any h in the finite element space, $h = \mathbf{h}^t\boldsymbol{\psi}$ where $\mathbf{h} := (h(\boldsymbol{\xi}_1), \ldots, h(\boldsymbol{\xi}_K))^t$

- $R_0 := \int_{\Omega_\mathcal{T}} (\boldsymbol{\psi}\boldsymbol{\psi}^t)\mathcal{W} \qquad R_1 := \int_{\Omega_\mathcal{T}} \nabla\boldsymbol{\psi}' \boldsymbol{K} \nabla\boldsymbol{\psi}$

Corollary. The estimators $\hat{\boldsymbol{\beta}} \in \mathbb{R}^q$ and $\hat{f} \in H^1_\mathcal{T}(\Omega)$, that solve the discrete counterpart of the estimation problem, exist unique

- $\hat{\boldsymbol{\beta}} = (W^tW)^{-1}W^t(\boldsymbol{z} - \hat{\mathbf{f}}_n)$

- $\hat{f} = \hat{\mathbf{f}}^t\boldsymbol{\psi}$, with $\hat{\mathbf{f}}$ satisfying

$$\begin{bmatrix} -\Psi^t Q \Psi & \lambda R_1 \\ \lambda R_1 & \lambda R_0 \end{bmatrix} \begin{bmatrix} \mathbf{f} \\ \boldsymbol{g} \end{bmatrix} = \begin{bmatrix} -\Psi^t Q\, \boldsymbol{z} \\ \mathbf{0} \end{bmatrix}$$

- $\hat{\boldsymbol{\beta}}$ and $\hat{f}$ are linear in $\mathbf{z}$ $\rightarrow$ *linear estimators*

  $\hat{f}$ has typical penalized regression form, being identified by

  $$\hat{\mathbf{f}}_n = (\Psi^t Q \Psi + \lambda P)^{-1} \Psi^t Q\, \mathbf{z} \qquad\qquad P = R_1\, R_0^{-1}\, R_1$$

- Classical inferential tools are readily derived
  - mean and variances of $\hat{\boldsymbol{\beta}}$ and $\hat{f}$
  - confidence intervals for $\boldsymbol{\beta}$
  - confidence bands for f
  - *prediction intervals for new observations*
  - estimate of error variance σ^2 $\qquad \boxed{\sum_{i=1}^{n} \left(z_i - \mathbf{w}_i^t \boldsymbol{\beta} - f(\mathbf{x}_i)\right)^2 + \lambda \int_\Gamma (\Delta_\Gamma f(\mathbf{x}))^2 d\mathbf{x}}$
  - selection of smoothing parameter λ via generalized cross-validation

Azzimonti et al., 2013b, TechRep

Two sources of bias:

- ▷ $\hat{f} \in H^1_{\mathcal{T}}(\Omega)$ is affected by bias due to discretization:

  This bias disappears as $n \to \infty$ with $h \to 0$ (infill asymptotic)

- ▷ $\hat{f} \in H^2_{n0}(\Omega)$ and $\hat{f} \in H^1_{\mathcal{T}}(\Omega)$ are affected by bias due to regularization

  This bias disappears as $n \to \infty$ with $\lambda \to 0$

- $\hat{f} \in H^2_{n0}(\Omega)$ and $\hat{f} \in H^1_{\mathcal{T}}(\Omega)$ are consistent
- (Hopefully also) asymptotically normal

Open problems

Simulation studies

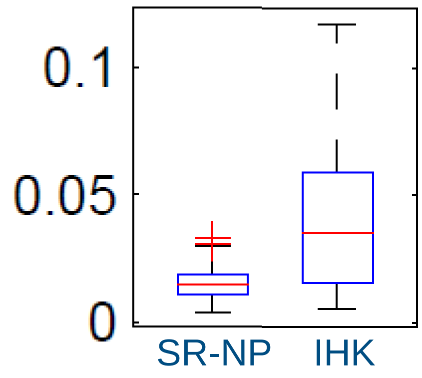
Simulation (without covariates)

TRUE

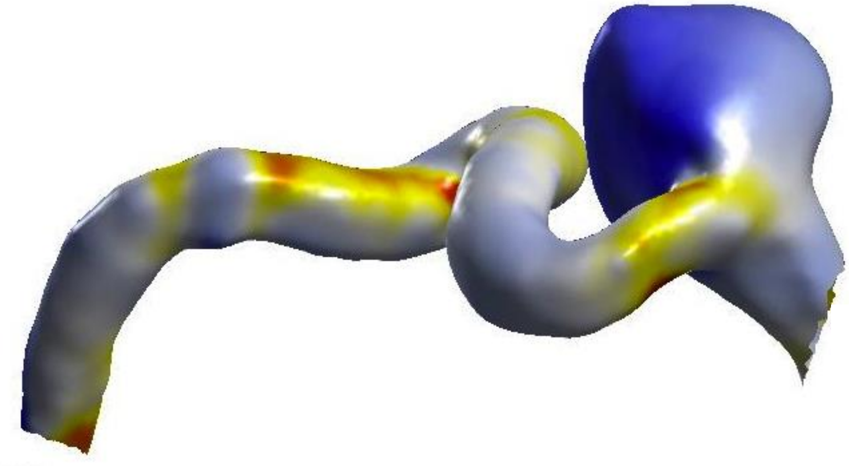
TRUE + NOISE

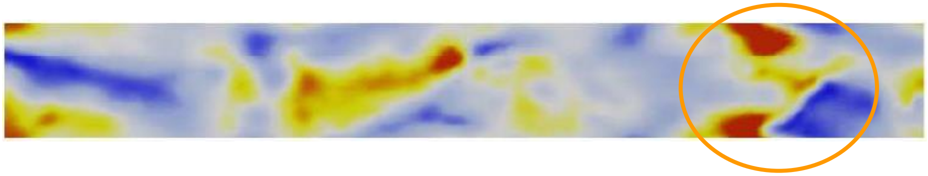
ESTIMATE

RMSE



50 simulation replicates

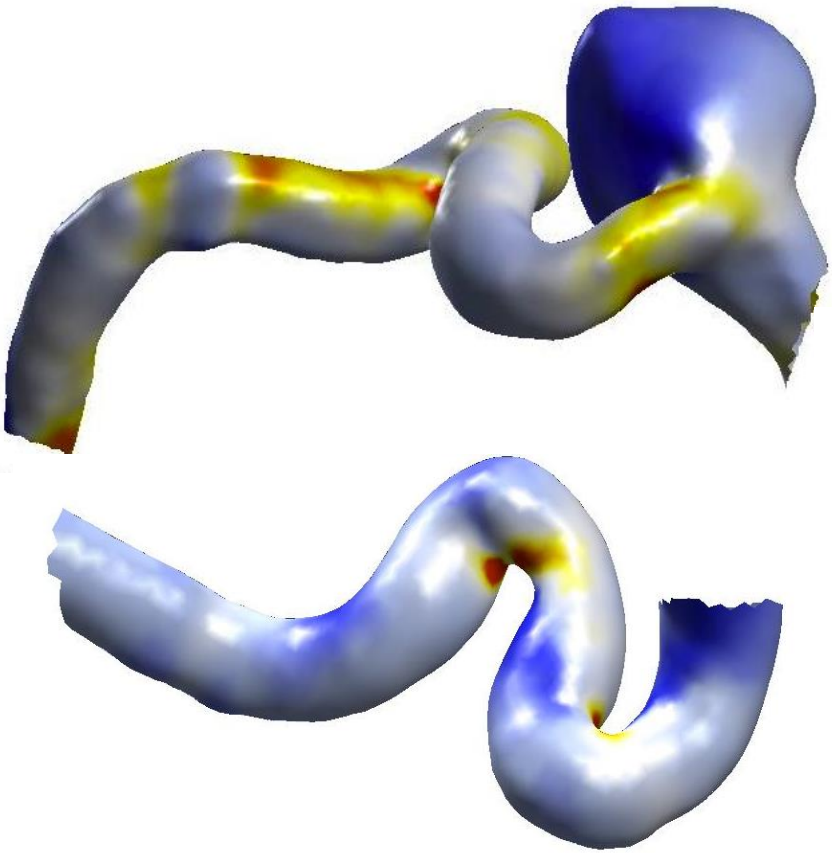


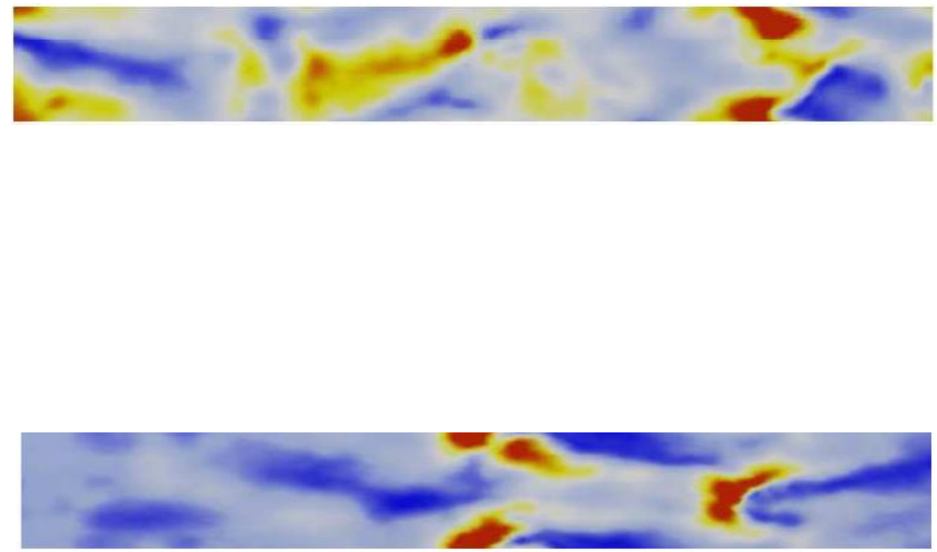


Covariates:

- Local curvature of vessel wall → Negative association
- Curvature of vessel → Positive association
- Local radius of vessel → Negative association

Future work

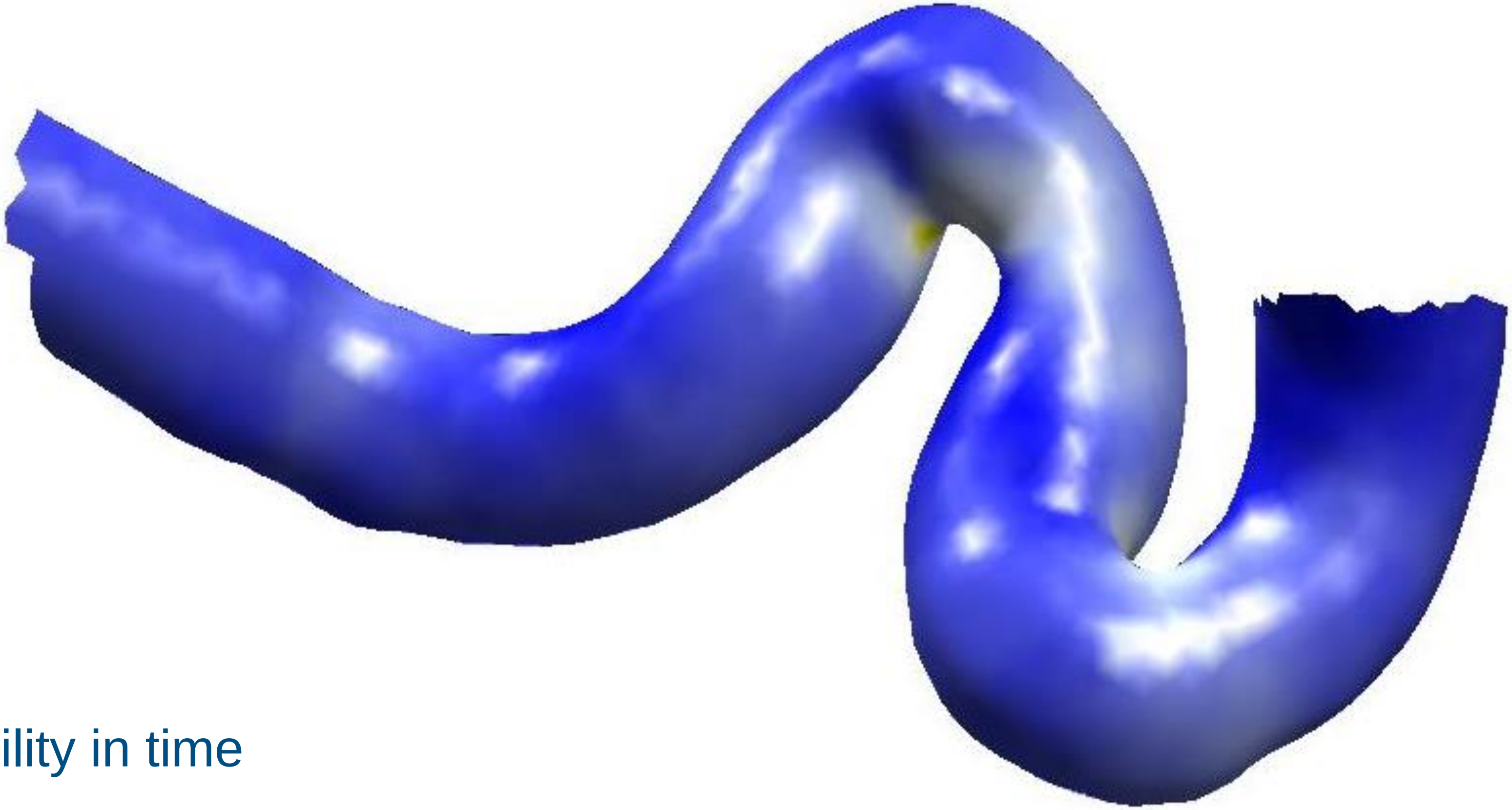




Variability across patients

. . .

(data registration)

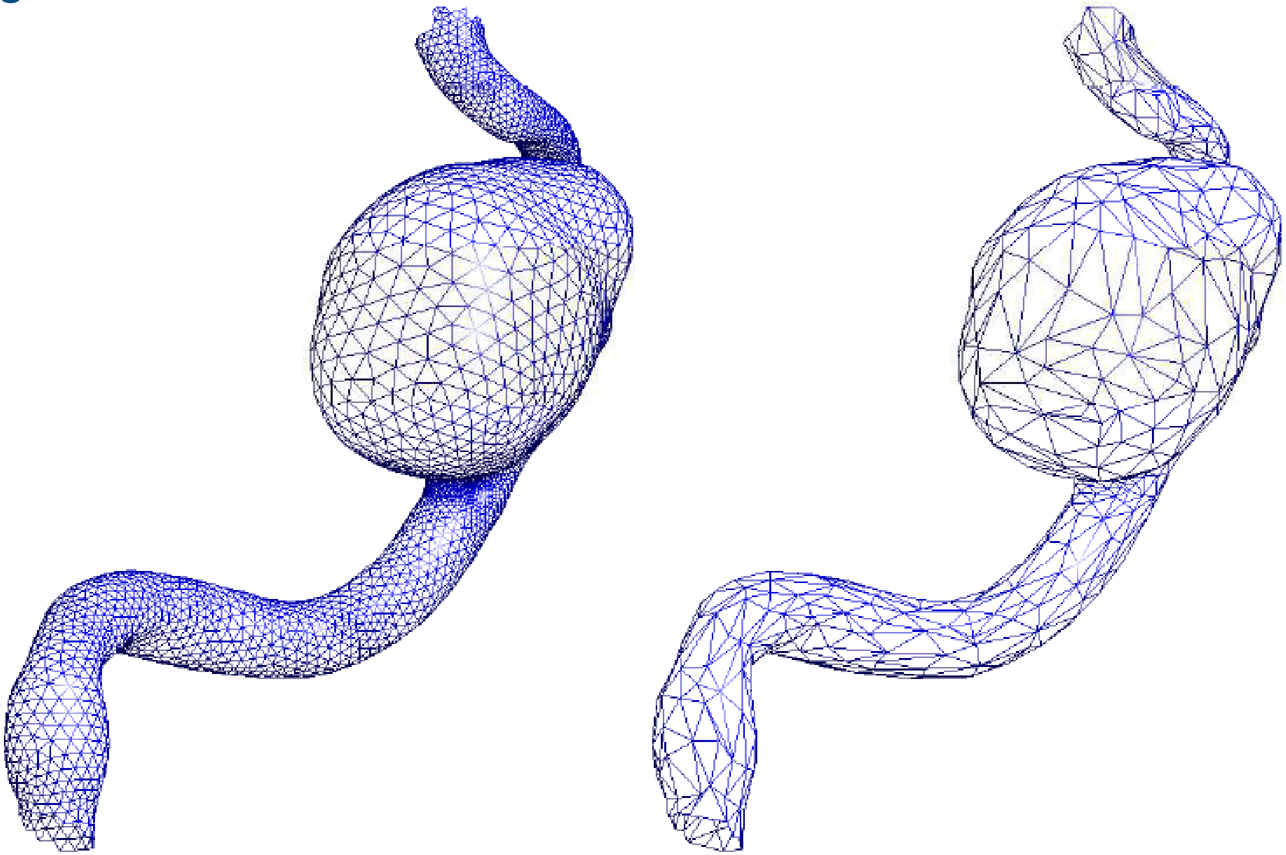


Variability in time

Facing big data challenges:

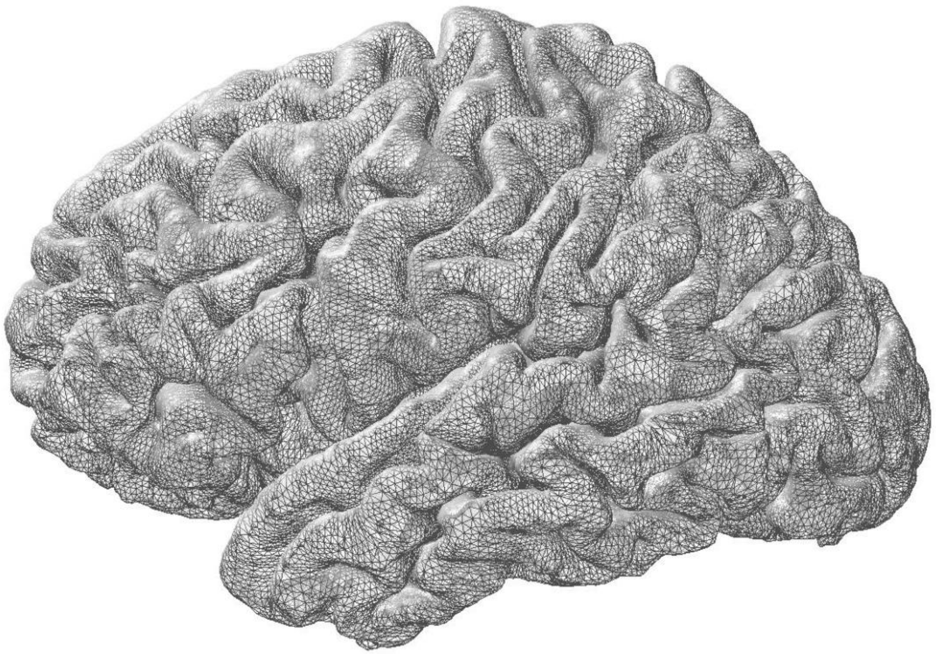
- iterative algorithms

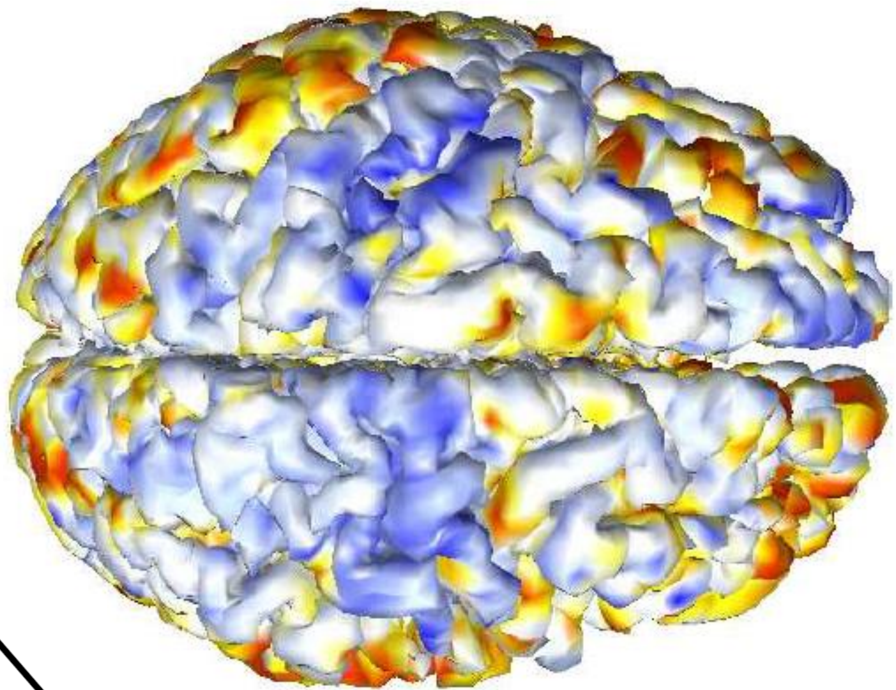
- mesh simplification algorithms

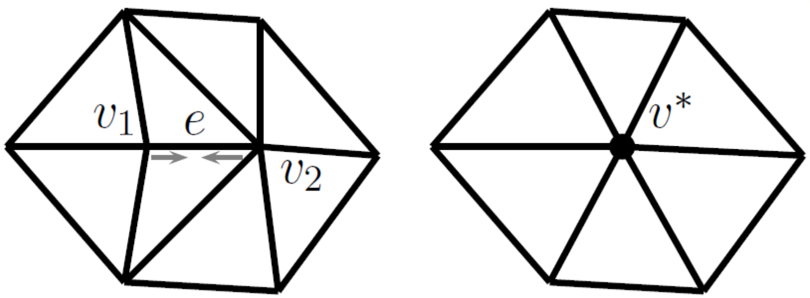


Spatial Regression models over Riemannian manifolds

Dassi et al., 2013, TechRep





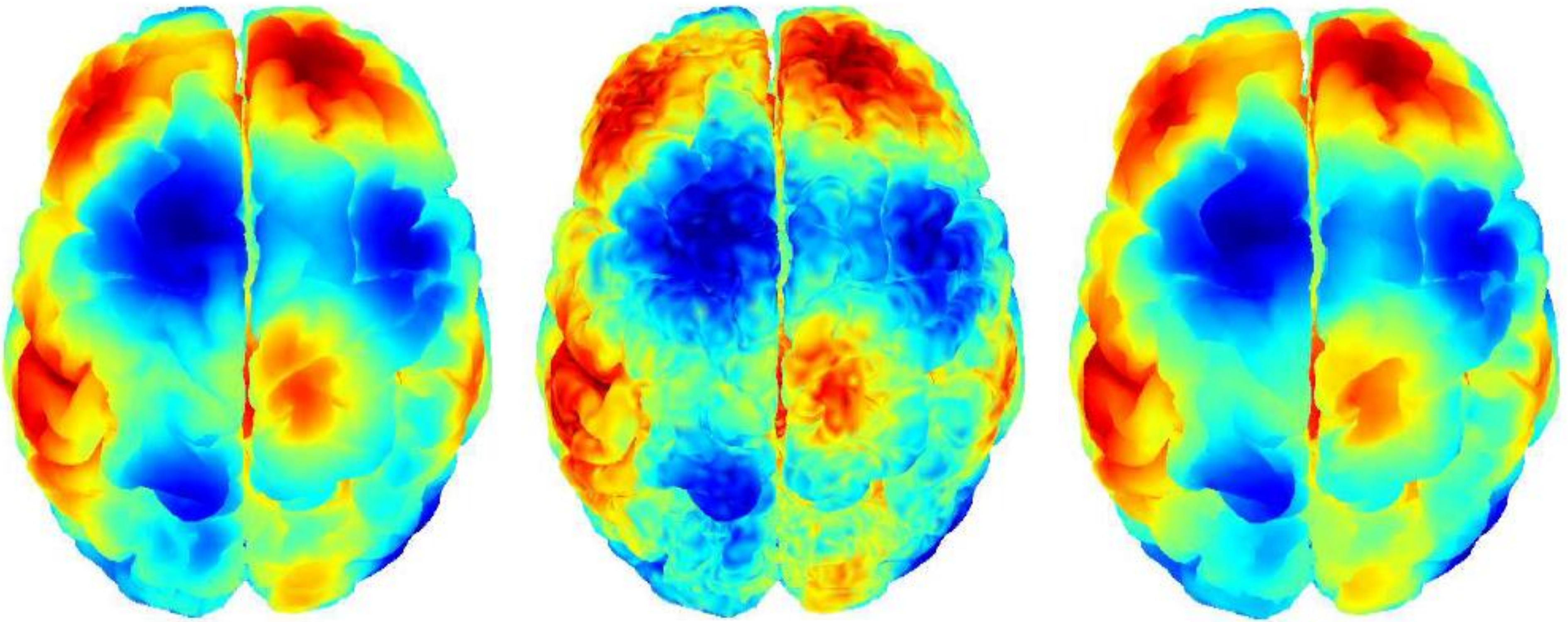


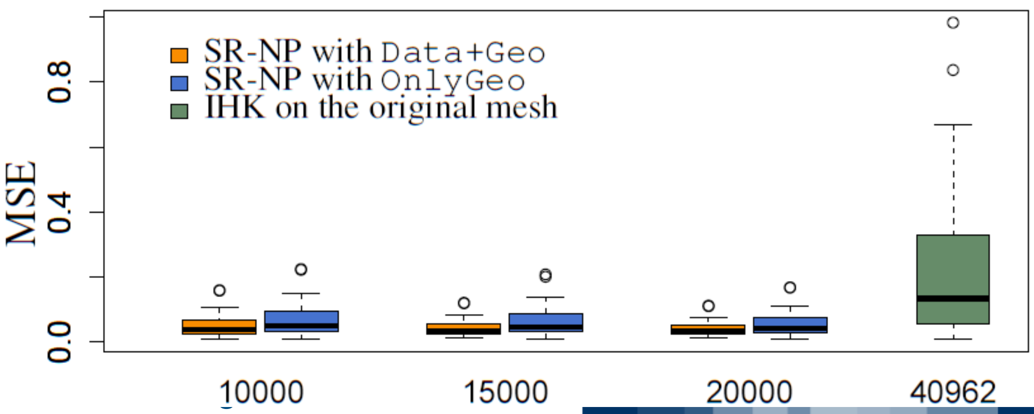
$$c(e,v^*) := \alpha c_{\mathrm{geo}}(e,v^*) + (1-\alpha)c_{\mathrm{data}}(e,v^*), \quad 0 \le \alpha \le 1$$

TRUE

TRUE + NOISE

ESTIMATE





Current features:

- Irregular shaped domains
- Boundary conditions
- Areal data
- Incorporate priori knowledge about spatial field
- Data over bi-dimensional manifolds
- Mesh simplification that preserves inferential properties of the estimators

Current research:

- *Binomial, Gamma e Poisson outcomes*
- *Space/time models*
- *Asymptotic properties*
- *Mixed effect models*

Future directions

- Combine these features to create *class of models with very broad applicability* that aims at handling data structures for which no statistical modeling currently exists
- Strong synergy of various approaches from different scientific disciplines, with an *intense interplay of statistics, mathematics and engineering*
- ***R* and *Matlab* code**

Aknowledgments

Acknowledgments:

FUTURO IN RICERCA SNAPLE

FIRB2008 Futuro in Ricerca 2010 - 2014 Starting grant from Ministero dell'Istruzione dell'Università e della Ricerca "SNAPLE: Statistical and Numerical methods for the Analysis of Problems in Life sciences and Engineering". http://mox.polimi.it/users/sangalli/firbSNAPLE.html

► *CLOSING WORKSHOP 15-16 May 2014, Politecnico di Milano*

Invited speakers: John Aston (Cambridge), Marc Genton (Kaust), Tilmann Gneiting (Heidelberg), Hans Georg Mueller (UC Davis), Fabio Nobile (EPFL), James Ramsay (Mc Gill)

RegioneLombardia

Dote Ricercatore Politecnico di Milano – Regione Lombardia: Research Project (2010 - 2013) "Functional data analysis for life sciences".

McGill University: Visiting research grant (Apr.- July 2010), Department of Mathematics and Statistics, McGill University, Montreal, Canada.

SAMSI: Visiting research grant (Jan.- Feb. 2011), SAMSI Statistical and Applied Mathematical Sciences Institute, Research Triangle Park, North Carolina, US

Main references on spatial regression with PDE regularization

Irregularly shaped domains and boundary conditions

- Sangalli L.M., Ramsay J.O., Ramsay T.O. (2013), “Spatial splines regression models”, *Journal of the Royal Statistical Society Ser. B, Statistical Methodology*, 75, 4, 681-703.
- Ramsay J.O., Ramsay T.O., Sangalli L.M. (2011), “Spatial Functional Data Analysis”, in Recent Advances in Functional Data Analysis and Related Topics, Springer Series Contribution to Statistics, pp. 269--276.

Incorporating a priori knowledge

- Azzimonti L., “Blood flow velocity field estimation via spatial regression with PDE penalization”, PhD Thesis, Politecnico di Milano, 2013.
- Azzimonti L., Sangalli L.M., Secchi P., Domanin M. and Nobile F. (2013), “Blood flow velocity field estimation via spatial regression with PDE penalization”, Tech.rep N. 19/2013, MOX, Dip. di Matematica, Politecnico di Milano.
- Azzimonti L., Nobile F., Sangalli L.M., Secchi P. (2013), “Mixed Finite Elements for spatial regression with PDE penalization”, Tech.rep N. 20/2013, MOX, Dip. di Matematica, Politecnico di Milano.

Manifold domains

- Ettinger B., Perotto S., Sangalli L.M. (2012), “Regression models over two-dimensional manifolds”, Tech.rep N. 54/2012, MOX, Dip. di Matematica, Politecnico di Milano.
- Ettinger B., Passerini T., Perotto S., Sangalli L.M. (2013), “Spatial smoothing for data distributed over non-planar domains”, in Complex Models and Computational Methods in Statistics, Springer Series Contribution to Statistics, pp. 123-136.
- Dassi F., Ettinger B., Perotto S., Sangalli L.M. (2013), “ Mesh simplification for spatial regression of cortical surface data”, Tech.rep N. 31/2013, MOX, Dip. di Matematica, Politecnico di Milano.

Other references

- Baramidze, V., Lai,M. J., and Shum, C. K. (2006), “Spherical splines for data interpolation and fitting,” *SIAM J. Sci. Comput.*, 28, 241–259.

- Chung, M. K., Robbins, S. M., Dalton, K. M., Davidson, R. J., Alexander, A. L., and Evans, A. C. (2005), “Cortical thickness analysis in autism with heat kernel smoothing,” *NeuroImage*, 25, 1256–1265.

- Gneiting, T. (2013), Strictly and non-strictly positive definite functions on spheres, Bernoulli, 19, 4, 1087-1500

- Guillas, S. and Lai, M. (2010), “Bivariate splines for spatial functional regression models,” *J. Nonparametric Stat*, 22, 477–497.

- Hagler, Jr., D. J., Saygin, A. P., and Sereno, M. I. (2006), “Smoothing and cluster thresholding for cortical surface-based group analysis of fMRI data,” *NeuroImage*, 33, 1093–1103.

- Haker, S., Angenent, S., Tannenbaum, A., and Kikinis, R. (2000), “Nondistorting flattening maps and the 3-D visualization of colon CT images,” *IEEE Trans. Med. Imag.*, 19, 665–670.

- Jun, M. (2011), “Non-stationary Cross-Covariance Models for Multivariate Processes on a Globe”, *Scandinavian Journal of Statistics*, 38, 726-747.

- Jun, M., Stein, M.L. (2007), “An approach to producing space-time covariance functions on spheres”, *Technometrics*, 49, 468-479.

- Lindgren, F., Rue, H., and Lindstr ¨om, J. (2011), “An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach,” *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 73, 423–498.

- Passerini, T., Sangalli, L. M., Vantini, S., Piccinelli, M., Bacigaluppi, S., Antiga, L., Boccardi, E., Secchi, P., and Veneziani, A. (2012), “An integrated CFD-statistical investigation of parent vasculature of cerebral aneurysms,” *Cardio. Eng. and Tech.*, 3, 26–40.

- Ramsay. Spline smoothing over difficult regions. J. R. Stat. Soc. Ser. B Stat. Methodol., 64, 307–319, 2002.

Other references

- Sangalli, L.M., Secchi, P., Vantini, S., Veneziani, A. (2009), “A Case Study in Exploratory Functional Data Analysis: Geometrical Features of the Internal Carotid Artery”,*Journal of the American Statistical Association*, 104, 37-48.

- Sangalli, L.M., Secchi, P., Vantini, S., Veneziani, A. (2009), “Efficient estimation of 3-dimensional curves and their derivatives by free-knot regression splines, applied to the analysis of inner carotid artery centerlines”, *Journal of the Royal Statistical Society Ser C*, 58, 285-306.

- Sangalli, L.M., Secchi, P., Vantini, S., Vitelli, V. (2010), “K-mean alignment for curve clustering”, *Computational Statistics and Data Analysis*, 54, pp. 1219-1233.

- Stone, G. (1988). Bivariate splines. PhD thesis, University of Bath, Bath.

- Stuart, A. (2010), “Inverse problems: a Bayesian perspective,” *Acta Numerica*, 19, 451–559.

- Wahba, G. (1990). Spline models for observational data. Society for Industrial and Applied Mathematics.

- Wahba, G. (1981), “Spline interpolation and smoothing on the sphere,” *SIAM J. Sci. Stat. Comput.*, 2, 5–16.

- Wood, S., Bravington, M., and Hedley, S. (2008), “Soap film smoothing,” *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 70, 931–955.