MOX–Report No. 08/2012

Spatial Spline Regression Models

Sangalli, L.M.; Ramsay, J.O.; Ramsay, T.O.

MOX, Dipartimento di Matematica "F. Brioschi"
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

mox@mate.polimi.it http://mox.polimi.it

Spatial Spline Regression Models

Laura M. Sangalli
MOX - Dipartimento di Matematica, Politecnico di Milano, Milano, Italy.

James O. Ramsay
Department of Psychology, McGill University, Montréal, Canada.

Timothy O. Ramsay
Ottawa Health Research Institute, Ottawa, Canada.

Abstract. We describe a model for the analysis of data distributed over irregularly shaped spatial domains with complex boundaries, strong concavities and interior holes. Adopting an approach typical of functional data analysis, we propose a Spatial Spline Regression model that is computationally efficient, allows for spatially distributed covariate information and can impose various conditions over the boundaries of the domain. Accurate surface estimation is achieved by the use of finite elements, which provide a basis for piecewise polynomial surfaces.

Keywords: functional data analysis, spatial data analysis, semi-parametric model, penalized smoothing, finite elements

1. Introduction

We propose a semi-parametric model for the analysis of data distributed over spatial domains, including those with complex domain boundaries, strong concavities and interior holes. Figure 1, showing census tract locations over the Island of Montréal, Québec, Canada, but excluding the airport and rail yards in the south and an industrial park with an oil refinery tank farm in the north-east, illustrates the kind of problem that we consider. Population density, average per-capita income and other measures are available at each census tract, and a binary variable indicating whether a tract is predominantly residential or industrial/commercial is available as covariate for estimating the distributions of census quantities. Here in particular we are interested in population density; the airport and industrial park are thus not part of the domain of interest since people cannot live in those two areas. Census quantities can vary sharply across these uninhabited parts of the city; for instance, in the south of the industrial park there is a densely populated area with medium-low income, but north-east and west of it are wealthy neighborhoods, with low population density to the north-east, and high population density to the west. Hence, whilst it seems reasonable to assume that population density features a smooth spatial variation over the inhabited parts of the island, there is no reason to assume smooth spatial variation across interior boundaries. The figure also shows the island coasts as boundaries of the domain of interest; those parts of the boundary that are highlighted in red correspond respectively to the harbor, in the east shore, and to two public parks, in the south-west and north east shore; and no people live by the river banks in these boundary intervals. We thus want to

E-mail: laura.sangalli@polimi.it

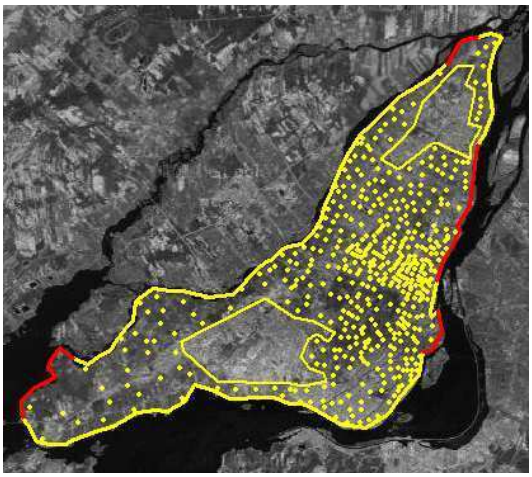

Figure 1. Island of Montréal census data. Dots indicates the centroids of census enumeration areas, for which population density and other census information are available. The two parts of the island where there are no data, encircled by yellow lines, are areas where people cannot live (the airport and rail yards in the south and an industrial park with an oil refinery tank farm in the north-east). The island boundary is also outlined in yellow and red, with red sections indicating the harbor and two public parks.

study population density, taking into account covariate information, being careful not to artificially link data across areas where people cannot live, and also efficiently including prior information concerning those stretches of coast where the population density should drop to zero. Well-known methods for spatial data analysis, such as kriging, kernel smoothing, wavelet-based smoothing, tensor product splines and thin-plate splines, are not appropriate for these data, since they do not take into account the shape of the domain and also smooth across concave boundary regions; moreover, these methods do not allow the specification of fixed values for the surface estimate at the domain boundary, or along parts of it.

We adopt functional data analysis approach in proposing a *Spatial Spline Regression* (SSR) model that overcomes these limitations, being able to efficiently deal with data distributed over irregularly shaped regions. The model incorporates the penalized bivariate spline smoother introduced by Ramsay (2002); in this smoother, the roughness penalty consists of a Laplace operator that is integrated only over the region of interest thanks to a finite element formulation that defines a system of local basis functions for continuous piecewise-polynomial surfaces. We improve the Ramsay (2002) smoother in many respects, from both computational and modeling perspectives. The modeling generalizations include the capacity to account for covariate measures and also to comply with different conditions at the boundary of the domain; the surfaces describing the spatial variation may either exhibit free behavior at the exterior or interior boundaries of the domain, with a zero gradient in the normal direction, or be constrained to desired fixed values. Moreover, these boundary conditions can vary over different intervals on the boundaries, and linear combinations of them may be applied as well. SSR estimators turn out to be linear in the observed data values, so that classical inferential tools may be readily derived. The finite element formulation is also computationally highly efficient.

SSR is compared to kriging, thin-plate splines and soap film smoothing, the latter intro-

duced by Wood et al. (2008) and used for instance in Marra et al. (2011) in a tensor product smoother that also includes the time dimension. Our simulation studies show that SSR and soap film smoothing provide a large advantage over the other more classical techniques when dealing with data scattered over irregular domains. SSR is also similar to the spatial data analysis models introduced by Guillas and Lai (2010) which also penalize roughness with a partial differential operator. Finally, SSR models have also strong connections with the work of Lindgren et al. (2011), that links Gaussian fields and Gaussian Markov random fields via a stochastic partial differential equation that induces a Matérn covariance and is solved over irregular grids of points resorting to finite elements.

1.1. Data and model

Let $\{\mathbf{p}_i = (x_i, y_i);\ i = 1 \ldots, n\}$ be a set of n points on a bounded regular domain $\Omega \subset \mathbb{R}^2$. Let z_i be the value of a real-valued variable observed at point $\mathbf{p}_i$, and let $\mathbf{w}_i = (w_{i1}, \ldots, w_{iq})^t$ be a q-vector of covariates associated to observation z_i.

The semi-parametric model for these data is

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

where ϵ_i, $i = 1, \ldots, n$, are residuals or errors distributed independently of each other, with zero mean and variance σ^2. Vector $\boldsymbol{\beta} \in \mathbb{R}^q$ contains regression coefficients and function f is real-valued and twice differentiable.

We estimate the regression coefficient vector $\boldsymbol{\beta}$ and the surface f by minimizing the penalized sum-of-square-error functional

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

The roughness penalty is the integral over Ω of the square of the Laplacian of f,

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

a measure of local curvature of f that is invariant with respect to Euclidean transformations of spatial coordinates, and therefore ensures that the concept of smoothness does not depend on the orientation of the coordinate system. This roughness penalty can be rewritten as $\int_\Omega \{(f_{xx})^2 + 2f_{xx}f_{yy} + (f_{yy})^2\}\, dx\, dy$ to facilitate comparison with thin-plate spline operator $\int_{\mathbb{R}^2} \{(f_{xx})^2 + 2(f_{xy})^2 + (f_{yy})^2\}\, dx\, dy$, the latter being integrated over the entire plane $\mathbb{R}^2$. As discussed in Section 8, the penalty in (2) can be generalized to deal with more complicated partial differential operators.

The estimation problem (2) is tackled by means of finite element analysis, a methodology mainly developed and used in engineering applications, to solve partial differential equations. The strategy of finite element analysis is very similar in spirit to that of univariate splines, and consists of partitioning the problem domain into small disjoint sub-domains and defining polynomial functions on each of these sub-domains in such a way that the union of these pieces closely approximates the solution. This simplified problem is made computationally tractable by the choice of the basis functions for the space of piecewise polynomials on the domain partition. Each piece of the partition, equipped with the basis functions defined over it, is named a *finite element*.

In our context, for instance, a convenient partition of the domain Ω is given by a triangulation $\mathcal{T}$ where each data point $\mathbf{p}_i$ is a triangle vertex. This triangulation can be obtained by a constrained Delaunay algorithm (see, e.g., Hjelle and Dæhlen 2006). Figure 14, for example, shows the triangulation of the Isle of Montréal census data.

Introductions to finite element analysis are offered, e.g., by Gockenbach (2006), Quarteroni (2009) and Braess (2007). Although not focussed on finite elements, Lai and Schumaker (2007) is a valuable reference on piecewise polynomial functions on triangulations.

The rest of the paper is organized as follows. Section 2 presents the estimation problem in functional or variational form as a preliminary to the development of the finite element approximation to the functional solution. Section 3 describes the finite element function spaces used for the approximation, and Section 4 derives the finite element solution to the estimation problem and the properties of the associated estimators. Section 5 focusses how to deal with various boundary conditions. Section 6 illustrates the performances of the proposed method via simulation studies. Section 7 is devoted to applications to real data, including the Island of Montréal census data and hemodynamics data on a carotid artery. Section 8 draws some conclusive considerations and discuss extensions of the model considered and directions of future research. Detailed proofs are deferred to the Appendix.

2. The estimation problem in variational terms

Let $H^m(\Omega)$ be the set of functions in $L^2(\Omega)$ having all derivatives up to order m in $L^2(\Omega)$. Let $H^m_{n0}(\Omega) \subset H^m(\Omega)$ consists of those functions whose normal derivatives are 0 on the boundary of Ω. Denote by $\mathbf{W}$ the $n \times q$ matrix whose ith row is given by $\mathbf{w}_i^t$, the vector of q covariates associated to observation z_i at $\mathbf{p}_i$, and assume that $\mathbf{W}$ has full rank. Let $\mathbf{P}$ be the orthogonal projection matrix that projects orthogonally on the subspace of $\mathbb{R}^n$ generated by the columns of $\mathbf{W}$, i.e., $\mathbf{P} := \mathbf{W}(\mathbf{W}^t\mathbf{W})^{-1}\mathbf{W}^t$, and let $\mathbf{Q} = \mathbf{I} - \mathbf{P}$, where $\mathbf{I}$ is the identity matrix. Furthermore, set $\mathbf{z} := (z_1, \ldots, z_n)^t$ and, for a given function f on Ω denote by $\mathbf{f}_n$ the vector of evaluations of f at the n data locations, i.e., $\mathbf{f}_n := (f(\mathbf{p}_1), \ldots, f(\mathbf{p}_n))^t$.

The penalized sum-of-square-error functional (2) is well defined for $\boldsymbol{\beta} \in \mathbb{R}^q$ and $f \in H^2(\Omega)$. Furthermore, as is shown in Appendix A.1, imposing a boundary condition on f, such as $f \in H^2_{n0}(\Omega)$, ensures that the estimation problem has a unique solution. For this reason, we shall start by considering the estimation problem over $\boldsymbol{\beta} \in \mathbb{R}^q$ and $f \in H^2_{n0}(\Omega)$. These assumptions on functional spaces, as well as other assumptions on functional spaces that we will make in the following, can be relaxed; however, for simplicity of exposition, we do not pursue this here. Proposition 1 characterizes the solution of the estimation problem.

PROPOSITION 1. *The estimators $\hat{\boldsymbol{\beta}}$ and $\hat{f}$ that jointly minimize (1) over $\boldsymbol{\beta} \in \mathbb{R}^q$ and $f \in H^2_{n0}(\Omega)$ are*

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

$$\mathbf{u}_n^t\, \mathbf{Q}\, \hat{\mathbf{f}}_n + \lambda \int_\Omega (\Delta u)(\Delta \hat{f}) = \mathbf{u}_n^t\, \mathbf{Q}\, \mathbf{z} \tag{3}$$

for every $u \in H^2_{n0}(\Omega)$.

The estimators $\hat{\boldsymbol{\beta}}$ and $\hat{f}$ are uniquely determined.

Proof. See Appendix A.1.

The following sections shows how an approximate solution to the estimation problem can be obtained using the finite element method. In particular, (3) is reformulated as a problem in $H^1(\Omega)$ and this problem is thus discretized using a finite element space included in $H^1(\Omega)$. The first part of the section briefly review the construction of this finite element space.

3. Lagrange triangular finite elements

We consider a regular triangulation $\mathcal{T}$, where adjacent triangles share either a vertex or a complete edge. Domain Ω is approximated by domain $\Omega_{\mathcal{T}}$ consisting of the union of all triangles, so that the boundary $\partial\Omega$ of Ω is approximated by a polygon (or more polygons, in the case for instance of domains with interior holes). It is assumed, therefore, that the number and density of triangles in $\mathcal{T}$ is sufficient to capture sharp features in $\partial\Omega$ as well providing a basis for adequately describing the data. We will assume in this paper, for purposes of simplicity only, that the triangulation points in $\mathcal{T}$ are numbered so that the first n are in fact data location points $\mathbf{p}_i$; i.e., letting $\mathbf{v}_j, j = 1, \ldots, J$ indicate vertices of triangles, we have that $\mathbf{v}_i = \mathbf{p}_i, i = 1, \ldots, n$, where $n \leq J$. Our methodology can easily be extended to allow the triangulation points and the data location points to be separate. It must be emphasized that the triangular mesh generation problem is far from trivial, is the subject of a large literature and constitutes a great deal of the technical challenge in applications of finite element analysis. Triangulation software is readily available in many free and commercial finite element packages.

The surface to be constructed over $\Omega_{\mathcal{T}}$ is assumed to be a polynomial in x and y over any triangle, and is continuous across edges and vertices. In this paper we consider the two cases where the polynomial is either linear or quadratic (the latter case being the one considered in Ramsay 2002). The linear case defines the linear polynomial over each triangle as linear combination of three basis functions, each having value one at a single vertex and zero at the other two; see Figure 2. The vertices used in this way are also called *control points* or *nodes*, and are indicated by $\boldsymbol{\xi}_k, k = 1, \ldots, K$. In the quadratic case, six nodes are required for each triangle, each node being associated with the unique quadratic polynomial that has value one at a single node and zero at the remaining five. In this quadratic case, we extend the set of vertices to include midpoints of edges in defining the nodes $\boldsymbol{\xi}_k$. To summarize, nodes and vertices coincide for piecewise linear surfaces, but quadratic surfaces involve roughly twice as many nodes as vertices.

Since adjacent triangles share an edge or a vertex, it is automatic that that a piece-wise linear or quadratic basis function ψ_k is associated with each node $\boldsymbol{\xi}_k$ that has value one at the node and value zero at all neighboring nodes. Each such piece-wise polynomial basis function is called a *Lagrange finite element*. Figure 3 show for the linear case one such basis function; its shape depends on the number of triangles sharing the vertex $\boldsymbol{\xi}_k$ and on the lengths of the triangle edges. The set of K basis functions defined in this way define a function subspace that we denote as $H^1_{1,\mathcal{T}}(\Omega) \subset H^1(\Omega)$ in the linear case and $H^1_{2,\mathcal{T}}(\Omega) \subset H^1(\Omega)$ in the quadratic case; $H^1_{\mathcal{T}}(\Omega)$ will be used to denote in general a finite element space, either $H^1_{1,\mathcal{T}}(\Omega)$ or $H^1_{2,\mathcal{T}}(\Omega)$.

Set $\boldsymbol{\psi} := (\psi_1, \ldots, \psi_K)^t$; moreover, for a given function f on Ω, denote by $\mathbf{f}$ the K-vector having as entries the evaluations of f at the K nodes, i.e., $\mathbf{f} := (f(\boldsymbol{\xi}_1), \ldots, f(\boldsymbol{\xi}_K))^t$. Being

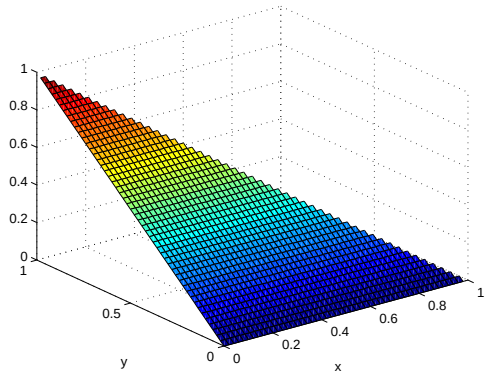

Figure 2. Nodal basis associated with node $\boldsymbol{\xi} = (0,1)$ over the triangle $\mathbf{t} = \{(0,0);(1,0);(0,1)\}$.

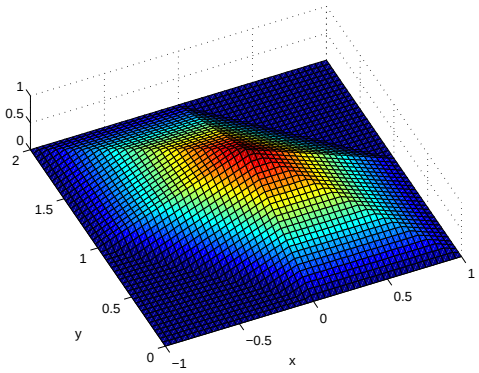

Figure 3. Nodal basis associated with node $\boldsymbol{\xi} = (0,1)$ over a uniform triangular mesh.

$\psi_k(\boldsymbol{\xi}_l) = \delta_{kl}$, the finite element space has been constructed precisely so that any function f in $H^1_{\mathcal{T}}(\Omega)$ is completely defined by its value at the K nodes:

$$f(x,y) = \sum_{k=1}^{K} c_k\, \psi_k(x,y) = \sum_{k=1}^{K} f(\boldsymbol{\xi}_k)\, \psi_k(x,y) = \mathbf{f}^t \boldsymbol{\psi}(x,y). \tag{4}$$

This property is very convenient from a computational point of view and will be exploited extensively in the following.

Using quadratic elements gives the possibility of estimating curvature over the interior of any triangle. But quadratic elements do not in general have continuous first derivatives at triangle edges; finite element systems with this property are rather more technical and are not considered in this paper.

4. Finite element solution to the estimation problem

In Appendix A.2 we show, by introducing an auxiliary function g and using integration by parts, that the problem of finding $\hat{f} \in H^2_{n0}(\Omega)$ that satisfies (3) for every $u \in H^2_{n0}(\Omega)$, may be reformulated as follows: find $(\hat{f}, g) \in \left(H^1_{n0}(\Omega) \cap C^0(\Omega)\right) \times H^1(\Omega)$ that satisfies

$$\begin{aligned} \mathbf{u}_n^t\, \mathbf{Q}\, \hat{\mathbf{f}}_n - \lambda \int_\Omega (\nabla u \cdot \nabla g) &= \mathbf{u}_n^t\, \mathbf{Q}\, \mathbf{z} \\ \int_\Omega v\, g + \int_\Omega (\nabla v \cdot \nabla \hat{f}) &= 0 \end{aligned} \tag{5}$$

for all $(u,v) \in \left(H^1_{n0}(\Omega) \cap C^0(\Omega)\right) \times H^1(\Omega)$; moreover, such $\hat{f}$ belongs to $H^2_{n0}(\Omega)$. This is a so-called weak formulation of the problem, and the finite element method requires its solution within a finite element space $H^1_{\mathcal{T}}(\Omega)$. Corollary 1 shows that, thanks to the choice for domain partition and function basis of the finite element space, finding the solution to this discrete counterpart of the estimation problem reduces to solving a linear system.

Let $\boldsymbol{\psi}_x := (\partial\psi_1/\partial x, \ldots, \partial\psi_k/\partial x)^t$ and $\boldsymbol{\psi}_y := (\partial\psi_1/\partial y, \ldots, \partial\psi_k/\partial y)^t$, and define the order K matrices

$$\mathbf{R}_0 := \int_{\Omega_\mathcal{T}} (\boldsymbol{\psi}\boldsymbol{\psi}^t) \qquad \text{and} \qquad \mathbf{R}_1 := \int_{\Omega_\mathcal{T}} (\boldsymbol{\psi}_x\boldsymbol{\psi}_x^t + \boldsymbol{\psi}_y\boldsymbol{\psi}_y^t).$$

Moreover, let us also introduce the order K block matrix $\mathbf{L}$, and the $K \times n$ block matrix $\mathbf{D}$, defined by

$$\mathbf{L} := \left[\begin{array}{c|c} \mathbf{Q} & \mathrm{O}_{n\times(K-n)} \\ \hline \mathrm{O}_{(K-n)\times n} & \mathrm{O}_{(K-n)\times(K-n)} \end{array}\right] \qquad \text{and} \qquad \mathbf{D} := \left[\begin{array}{c} \mathbf{I}_n \\ \hline \mathrm{O}_{(K-n)\times n} \end{array}\right]$$

where $\mathrm{O}_{m_1\times m_2}$ is a $m_1 \times m_2$ matrix with all entries equal to zero. Finally, denote by $\mathbf{0}$ the null vector.

COROLLARY 1. *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, are given by*

- $\hat{\boldsymbol{\beta}} = (\mathbf{W}^t\mathbf{W})^{-1}\mathbf{W}^t(\mathbf{z} - \hat{\mathbf{f}}_n)$
- *$\hat{f}$ is identified by the coefficient vector $\hat{\mathbf{f}}$ satisfying*

$$\begin{bmatrix} -\mathbf{L} & \lambda\mathbf{R}_1 \\ \lambda\mathbf{R}_1 & \lambda\mathbf{R}_0 \end{bmatrix} \begin{bmatrix} \mathbf{f} \\ \mathbf{g} \end{bmatrix} = \begin{bmatrix} -\mathbf{L}\,\mathbf{D}\,\mathbf{z} \\ \mathbf{0} \end{bmatrix} \tag{6}$$

The estimators $\hat{\boldsymbol{\beta}}$ and $\hat{f}$ are uniquely determined.

Proof. See Appendix A.3.

As mentioned in the introduction, the boundary condition that the normal derivatives of f are 0 on the boundary of Ω, i.e., $f \in H^2_{n0}(\Omega)$, guarantees the uniqueness of the solution to the estimation problem (2). Although the estimate $\hat{f} \in H^1_\mathcal{T}(\Omega)$ is an approximation of $f \in H^2_{n0}(\Omega)$, the estimate $\hat{f}$ itself will not generally have normal derivatives equal to 0 on the domain boundary, since this has not been enforced in the construction of the finite element space. The boundary condition on normal derivatives is rather used as a so-called *natural boundary condition*, being exploited when using integration by parts to obtain reformulation (5) of the estimation problem (see Appendix A.2).

Solving the linear system (6) is fast because, although the system is typically large, being of order $2K$, it is highly sparse because the matrices $\mathbf{R}_0$ and $\mathbf{R}_1$ are highly sparse, since the cross-products of nodal basis functions and of their partial derivatives are mostly zero. As an example, for the Isle of Montréal census data, we used 626 nodes and only about 1% of the entries of $\mathbf{R}_0$ and 0.2% of the entries of $\mathbf{R}_1$ were non-zero.

We compute the integrals of the cross-products of nodal basis functions and of their partial derivatives in $\mathbf{R}_0$ and $\mathbf{R}_1$ exactly, instead of using quadrature approximations as in Ramsay (2002).

4.1. Properties of the estimators

Denote by $\mathbf{B}$ the order $2K$ matrix in system (6), and set $\mathbf{A} := -\mathbf{B}^{-1}$. Moreover, denote by $\mathbf{A}_n$ the order n matrix corresponding to the first n rows and n columns of $\mathbf{A}$, and by

$\mathbf{A}_{Kn}$ the $K \times n$ matrix corresponding to the first K rows and n columns of $\mathbf{A}$. With a few simplifications we can thus write

$$\hat{\mathbf{f}}_n = \mathbf{A}_n \, \mathbf{Q} \, \mathbf{z} \qquad \text{and} \qquad \hat{\mathbf{f}} = \mathbf{A}_{Kn} \, \mathbf{Q} \, \mathbf{z} = \mathbf{A}_{Kn} \, \mathbf{A}_n^{-1} \, \hat{\mathbf{f}}_n. \tag{7}$$

This expression for $\hat{\mathbf{f}}$ highlights the fact that the finite element solution $\hat{f}$ at the $K - n$ nodes that are not data points is completely determined by the solution $\hat{f}$ at the n data points. Moreover, using the expression of $\hat{\mathbf{f}}_n$ in (7), we have

$$\hat{\boldsymbol{\beta}} = (\mathbf{W}^t\mathbf{W})^{-1}\mathbf{W}^t\big\{\mathbf{I} - \mathbf{A}_n \, \mathbf{Q}\big\} \, \mathbf{z} \, .$$

Consequently the estimators $\hat{\mathbf{f}}$, $\hat{\mathbf{f}}_n$ and $\boldsymbol{\beta}$ are linear functions of the data values; their properties are therefore straightforward to derive and classic inferential tools can be obtained. Recalling that $\mathrm{E}[\mathbf{z}] = \mathbf{W}\boldsymbol{\beta} + \mathbf{f}_n$ and $\mathtt{Var}(\mathbf{z}) = \sigma^2 \, I$, and exploiting the properties of the matrices involved (e.g., $\mathbf{Q}$ is symmetric and idempotent, $\mathbf{A}_n$ is symmetric, $\mathbf{Q}\,\mathbf{W} = \mathrm{O}_{n\times n}$, $\mathbf{Q}\,\mathbf{W}(\mathbf{W}^t\mathbf{W})^{-1} = (\mathbf{W}^t\mathbf{W})^{-1}\mathbf{W}^t\,\mathbf{Q} = \mathrm{O}_{n\times n}$), with a few simplifications we obtain the means and variances of the estimators $\hat{\mathbf{f}}_n$ and $\hat{\boldsymbol{\beta}}$:

$$\begin{aligned} \mathrm{E}[\hat{\mathbf{f}}_n] &= \mathbf{A}_n \, \mathbf{Q} \, \mathbf{f}_n \\ Var(\hat{\mathbf{f}}_n) &= \sigma^2 \mathbf{A}_n \, \mathbf{Q} \, \mathbf{A}_n \end{aligned} \tag{8}$$

and

$$\begin{aligned} E[\hat{\boldsymbol{\beta}}] &= \boldsymbol{\beta} + (\mathbf{W}^t\mathbf{W})^{-1}\mathbf{W}^t\big(\mathbf{I} - \mathbf{A}_n \, \mathbf{Q}\big)\mathbf{f}_n \\ Var(\hat{\boldsymbol{\beta}}) &= \sigma^2(\mathbf{W}^t\mathbf{W})^{-1} + \sigma^2(\mathbf{W}^t\mathbf{W})^{-1}\mathbf{W}^t\big\{\mathbf{A}_n \, \mathbf{Q} \, \mathbf{A}_n\big\}\mathbf{W}(\mathbf{W}^t\mathbf{W})^{-1}. \end{aligned} \tag{9}$$

Consider the vector $\hat{\mathbf{z}}$ of fitted values at the n data points

$$\hat{\mathbf{z}} = \mathbf{W}\hat{\boldsymbol{\beta}} + \hat{\mathbf{f}}_n = \big(\mathbf{P} + \mathbf{Q} \, \mathbf{A}_n \, \mathbf{Q}\big) \, \mathbf{z} = \mathbf{S} \, \mathbf{z}$$

where $\mathbf{S}$ denotes the smoothing matrix $\mathbf{S} := \mathbf{P} + \mathbf{Q} \, \mathbf{A}_n \, \mathbf{Q}$. The spatial spline regression estimator is thus a linear estimator, with the fitted values $\hat{\mathbf{z}}$ obtained from observations $\mathbf{z}$ via application of the linear operator $\mathbf{S}$, independent of $\mathbf{z}$. A commonly used measure of the equivalent degrees of freedom for linear estimators is given by $tr(\mathbf{S})$ (see, e.g., Buja et al. 1989, who first introduced this notion). The equivalent degrees of freedom of the SSR estimator,

$$tr(\mathbf{S}) = q + tr(\mathbf{A}_n \, \mathbf{Q}),$$

are given by the sum of the q degrees of freedom of the parametric part of the model (q being the number of covariates considered) and the equivalent degrees of freedom $tr(\mathbf{A}_n \, \mathbf{Q})$ corresponding to the non-parametric part of the model; recall in fact that $\hat{\mathbf{f}}_n = \big(\mathbf{A}_n \, \mathbf{Q}\big) \, \mathbf{z}$ as in (7).

We can now estimate σ^2 by

$$\hat{\sigma}^2 = \frac{1}{n - tr(\mathbf{S})}\big(\mathbf{z} - \hat{\mathbf{z}}\big)^t\big(\mathbf{z} - \hat{\mathbf{z}}\big).$$

This estimate, together with expressions (9) and (8), may be used to obtain approximate confidence intervals for $\boldsymbol{\beta}$ and approximate pointwise confidence bands for f. Furthermore, the value of the smoothing parameter λ may be selected by Generalized-Cross-Validation:

$$GCV(\lambda) = \frac{1}{n\big(1 - tr(\mathbf{S})/n\big)^2} \left(\mathbf{z} - \hat{\mathbf{z}}\right)^t \left(\mathbf{z} - \hat{\mathbf{z}}\right).$$

Finally, the value predicted for a new observation, at point $\mathbf{p}_{n+1}$ and with covariates $\mathbf{w}_{n+1}$, is given by

$$\hat{z}_{n+1} = \mathbf{w}_{n+1}^t \hat{\boldsymbol{\beta}} + \hat{f}(\mathbf{p}_{n+1}) = \mathbf{w}_{n+1}^t \hat{\boldsymbol{\beta}} + \hat{\mathbf{f}}^t \boldsymbol{\psi}(\mathbf{p}_{n+1}),$$

whose mean and variance can be obtained from expressions above; correspondingly, approximate prediction intervals may be also derived.

5. Boundary conditions

It is often the case that one would like the smoothing surface f to have specific values at the boundary of the domain, or in some part of the boundary domain. For instance, Azzimonti et al. (2011) study the blood-flow velocity field in a section of a carotid artery, using data provided by eco-color dopplers. In this applied problem it is a priori known that blood-flow velocity should be zero at the arterial wall because of the friction between blood cells and wall.

Thus, suppose we know that $f = f_{\partial\Omega}$ on $\partial\Omega$, where $f_{\partial\Omega}$ is sufficiently regular. We therefore want to estimate $\boldsymbol{\beta}$ and f by minimizing the penalized sum-of-square-error functional (2) over $\boldsymbol{\beta} \in \mathbb{R}^q$ and $f \in H^2(\Omega)$, conditioned on $f|_{\partial\Omega} = f_{\partial\Omega}$. Denote by $H_0^m(\Omega)$ the subset of $H^m(\Omega)$ consisting of those functions vanishing on the boundary of Ω. Proposition 2 characterizes the solution of this estimation problem with fixed value boundary conditions.

PROPOSITION 2. *The estimators $\hat{\boldsymbol{\beta}}$ and $\hat{f}$ that jointly minimize (1), over $\boldsymbol{\beta} \in \mathbb{R}^q$ and $f \in H^2(\Omega)$ with $f|_{\partial\Omega} = f_{\partial\Omega}$, are given by*

- $\hat{\boldsymbol{\beta}} = (\mathbf{W}^t\mathbf{W})^{-1}\mathbf{W}^t(\mathbf{z} - \hat{\mathbf{f}}_n)$
- $\hat{f} = \hat{s} + \bar{f}$, *where $\bar{f}$ is any fixed function in $H^2(\Omega)$ such that $\bar{f}|_{\partial\Omega} = f_{\partial\Omega}$ and with known gradient $\nabla\bar{f}$ and Laplacian $\Delta\bar{f}$, and $\hat{s} \in H_0^2(\Omega)$ satisfies*

$$\mathbf{u}_n^t\, \mathbf{Q}\, \hat{\mathbf{s}}_n + \lambda \int_\Omega (\Delta u)(\Delta \hat{s}) = \mathbf{u}_n^t\, \mathbf{Q}\, (\mathbf{z} - \bar{\mathbf{f}}_n^t) - \lambda \int_\Omega (\Delta u)(\Delta \bar{f}) \tag{10}$$

for every $u \in H_0^2(\Omega)$.

The estimators $\hat{\boldsymbol{\beta}}$ and $\hat{f}$ are uniquely determined.

Proof. See Appendix A.4.

Likewise in Section 4 we derive a reformulation of the estimation problem that will constitute the base for the application of the finite element method. Denote by ∂_ν the normal derivative on the boundary of Ω, $\partial_\nu u := \nu \cdot \nabla u$, where ν is the outward normal to $\partial\Omega$. Appendix A.5 shows that, introducing an auxiliary function g and using integration by

parts, the problem of finding $\hat{s}$ that satisfies (10) for every $u \in H_0^2(\Omega)$ may be formulated as follows: find $(\hat{s}, g) \in \big(H_0^1(\Omega) \cap C^0(\Omega)\big) \times H^1(\Omega)$ that satisfies

$$\begin{aligned} &\mathbf{u}_n^t\, \mathbf{Q}\, \hat{\mathbf{s}}_n - \lambda \int_\Omega (\nabla u \cdot \nabla g) + \lambda \int_{\partial\Omega} (\partial_\nu u)\, g = \mathbf{u}_n^t\, \mathbf{Q}\, (\mathbf{z} - \bar{\mathbf{f}}_n) \\ &\int_\Omega v\, g + \int_\Omega (\nabla v \cdot \nabla \hat{s}) - \int_{\partial\Omega} v\, (\partial_\nu \hat{s}) = -\int_\Omega (\nabla v \cdot \nabla \bar{f}) + \int_{\partial\Omega} v\, (\partial_\nu \bar{f}) \end{aligned} \tag{11}$$

for all $(u, v) \in \big(H_0^1(\Omega) \cap C^0(\Omega)\big) \times H^1(\Omega)$; moreover, such $\hat{s}$ belongs to $H_0^2(\Omega)$, and therefore $\hat{f} = \hat{s} + \bar{f} \in H^2(\Omega)$ with $\hat{f}|_{\partial\Omega} = f_{\partial\Omega}$. Corollary 2 gives the finite element solution to this problem.

Let m be the number of interior nodes. Denote by $\widetilde{\mathbf{L}}$ the order m matrix obtained from $\mathbf{L}$ by removing the $(K - m)$ rows and columns corresponding to boundary nodes, and by $\widetilde{\mathbf{R}}_1$ the $m \times K$ matrix obtained from $\mathbf{R}_1$ by removing the $(K - m)$ rows corresponding to boundary nodes. Moreover, define the order K matrix $\mathbf{R}_\nu$,

$$\mathbf{R}_\nu := \int_{\partial\Omega_{\mathcal{T}}} (\nu_x\, \boldsymbol{\psi}_x + \nu_y\, \boldsymbol{\psi}_y)\, \boldsymbol{\psi}^t,$$

and denote by $\widetilde{\mathbf{R}}_\nu$ the $m \times K$ matrix obtained from $\mathbf{R}_\nu$ by removing the $(K - m)$ rows corresponding to boundary nodes. The matrix $\mathbf{R}_\nu$ is highly sparse, having non-zero entries only for couples of nodal basis associated to boundary elements.

Corollary 2. *The estimators $\hat{\boldsymbol{\beta}} \in \mathbb{R}^q$ and $\hat{f} \in H^1_{\mathcal{T}}(\Omega)$ with $\hat{f} = f_{\partial\Omega}$ on all boundary nodes, that solve the discrete counterpart of the estimation problem with fixed value boundary conditions, are given by*

- $\hat{\boldsymbol{\beta}} = (\mathbf{W}^t\mathbf{W})^{-1}\mathbf{W}^t(\mathbf{z} - \hat{\mathbf{f}}_n)$
- *$\hat{f} = \hat{s} + \bar{f}$, where $\bar{f}$ is the finite element function that coincides with $f_{\partial\Omega}$ on the $K-m$ boundary nodes and equals 0 on the m interior nodes, and $\hat{s}$ is the finite element function that equals 0 on the $K-m$ boundary nodes and takes values on the m interior nodes identified by the vector $\hat{\tilde{\mathbf{s}}} \in \mathbb{R}^m$ satisfying*

$$\begin{bmatrix} -\widetilde{\mathbf{L}} & \lambda(\widetilde{\mathbf{R}}_1 - \widetilde{\mathbf{R}}_\nu) \\ \lambda(\widetilde{\mathbf{R}}_1 - \widetilde{\mathbf{R}}_\nu)^t & \lambda\mathbf{R}_0 \end{bmatrix} \begin{bmatrix} \mathbf{s} \\ \mathbf{g} \end{bmatrix} = \begin{bmatrix} -\widetilde{\mathbf{L}}\,\mathbf{z} \\ -\lambda(\mathbf{R}_1 - \mathbf{R}_\nu)^t\, \bar{\mathbf{f}} \end{bmatrix} \tag{12}$$

The estimators $\hat{\boldsymbol{\beta}}$ and $\hat{f}$ are uniquely determined.

Proof. See Appendix A.6.

We have considered the case where the values of the smoothing surface function are fixed along the whole boundary. This is been done only for clarity of explanation; with minor changes we can deal with mixed boundary conditions. For instance, we can specify fixed values only over a part of the boundary, say Γ_D, leaving the solution free on the other part, say Γ_N, where we ask that the normal derivatives are zero, as done in Section 2 (here, $\Gamma_D \cup \Gamma_N = \partial\Omega$). From a computational point of view, this is a minor modification of what has been explained above. In the simulation and applications presented in the

following section, we will look for estimates that satisfy mixed boundary conditions. In the numerical analysis terminology, fixing the values of the solution at the domain boundary is a *Dirichlet boundary condition*, fixing instead the values of the normal derivatives is a *Neumann boundary condition*. Within our framework, besides the *mixed boundary conditions* mentioned above (with different conditions on different part of the boundary), it is also straightforward to consider *Robin boundary conditions*, consisting in linear combinations of Dirichlet and Neumann conditions.

6. Simulation studies

In this Section we illustrate the performances of the proposed SSR models via simulation studies, following a similar simulation scheme as in Wood et al. (2008).

6.1. Simulations on C-domain

We shall here compare the performances of filtered kriging (KRIG), thin-plate splines (TPS), soap-film smoothing (SOAP) and SSR models, in a simulation study on a C-shaped domain with the surface test function used by Wood et al. (2008), which is in turn a modification of the one presented by Ramsay (2002). This function f, on its C-shaped domain, is shown in the top left panel of Figure 4. With respect to Wood et al. (2008) we also include covariates in the simulation study. In particular, for $N = 50$ replicates, we simulate data as follows. We sample $n = 200$ locations, $\mathbf{p}_1, \ldots, \mathbf{p}_n$, uniformly on the C-shaped domain. Independently for each $\mathbf{p}_i$, we sample two independent covariates w_{i1} and w_{i2}, from a $N(\mu_1, \sigma_1^2)$ and a $N(\mu_2, \sigma_2^2)$ distribution, respectively. We thus obtain $z_1, \ldots, z_n$ from

$$z_i = \beta_1 w_{i1} + \beta_2 w_{i2} + f(\mathbf{p}_i) + \epsilon_i \qquad i = 1, \ldots, n \tag{13}$$

where ϵ_i, $i = 1, \ldots, n$, are independent errors with $N(0, \sigma_\epsilon^2)$ distribution. The parameter values used in the simulation are: $\beta_1 = -0.5$, $\beta_2 = 0.2$, $\sigma_\epsilon = 0.5$, $\mu_1 = 3$, $\sigma_1 = 1.5$, $\mu_2 = 7$, $\sigma_2 = 5$. The top right panel of Figure 4 shows the data sampled in the first replicate, with the size of the point marker proportional to the size of the sampled value. The bottom center and right panels of the same figure display the corresponding scatter plots of the response variable versus the two covariates. From these plots, for instance, it is not apparent that the first covariate is significant for the explanation of the variability of the response.

KRIG and TPS estimates are obtained under standard implementation settings, using the R package **fields** (see Furrer et al. 2010), with the default exponential covariance for KRIG. SOAP is implemented using R package **soap** (see Wood 2010), and uses 32 interior knots and a rank 39 (40-knot) cyclic penalized cubic regression spline as the boundary curve. The bottom left panel of Figure 4 illustrate the triangulation used for SSR computations for the first replicate. The values of the smoothing parameters for the four methods are selected, for each replicate and method, by GCV.

Table 7 shows the Root Mean Square Errors (RMSE), over the 50 simulation replicates, of the estimate of the parameters β_1, β_2 and σ_ϵ provided by KRIG, TPS, SOAP and linear and quadratic SSR models (SSR1 and SSR2, respectively). SOAP and SSR models give the best estimates of the parameters β_1 and β_2; SSR models give better estimates of σ_ϵ than any of the other considered methods. Figure 8 shows the boxplots of the RMSE, over the 50 simulation replicates, of the estimate of f, evaluated on a fine grid of step 0.02 over the C-domain; the boxplots highlight that the best estimates of f are provided by SSR models,

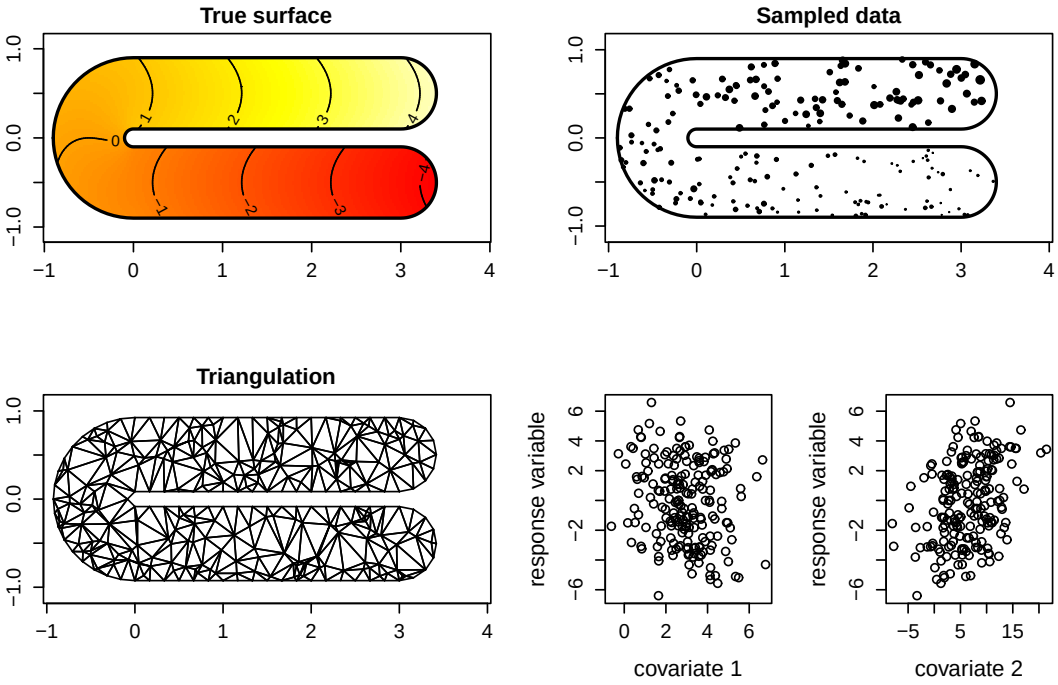

Figure 4. Top left: color map of true function (3D image shown in Figure 5). Top right: sampled data (size of the point marker proportional to sampled data value), replicate 1. Bottom left: domain triangulation, replicate 1. Bottom center and right: scatter plots of response variable vs covariates, replicate 1.

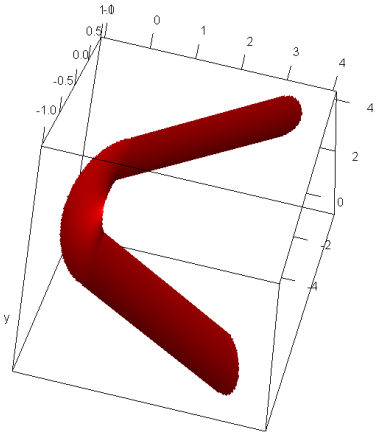

Figure 5. 3D image of the surface function on C-shaped domain (color map shown in Figure 4).

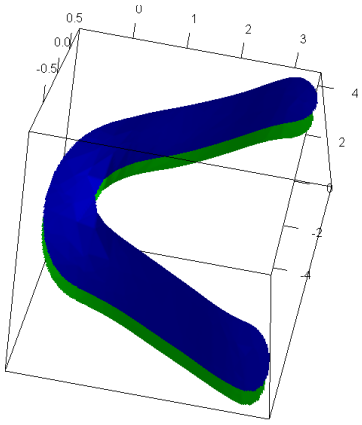

Figure 6. Pointwise 95% confidence regions for the linear SSR estimate, replicate 1.

	RMSE β_1	RMSE β_2	RMSE σ_ϵ
KRIG	0.0267	0.0099	0.1284
TPS	0.0247	0.0087	0.0612
SOAP	0.0232	0.0079	0.0314
SSR1	0.0232	0.0079	0.0289
SSR2	0.0232	0.0079	0.0289

Figure 7. RMSE of the estimates of the trend parameters β_1, β_2 and of the error standard deviation σ_ϵ, obtained via filtered kriging, thin-plate splines, soap-film smoothing and linear and quadratic SSR models.

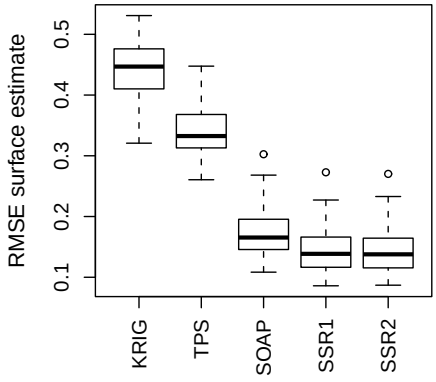

Figure 8. Boxplots of RMSE of the estimates of f, evaluated on a fine grid of step 0.02 over the C-domain, obtained by the different methods.

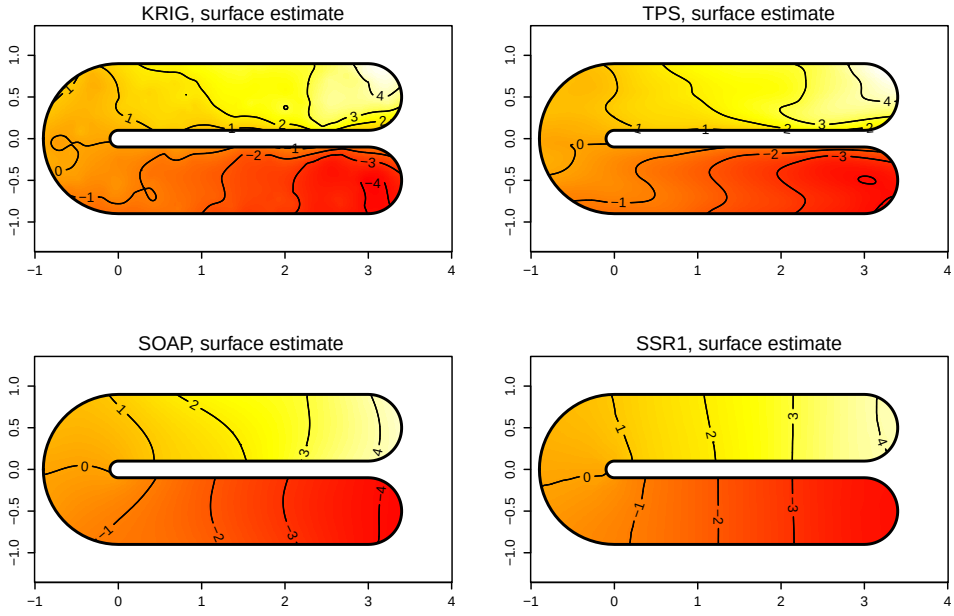

Figure 9. Color maps of f estimates obtained in replicate 1; quadratic SSR is not displayed here as almost indistinguishable from linear SSR.

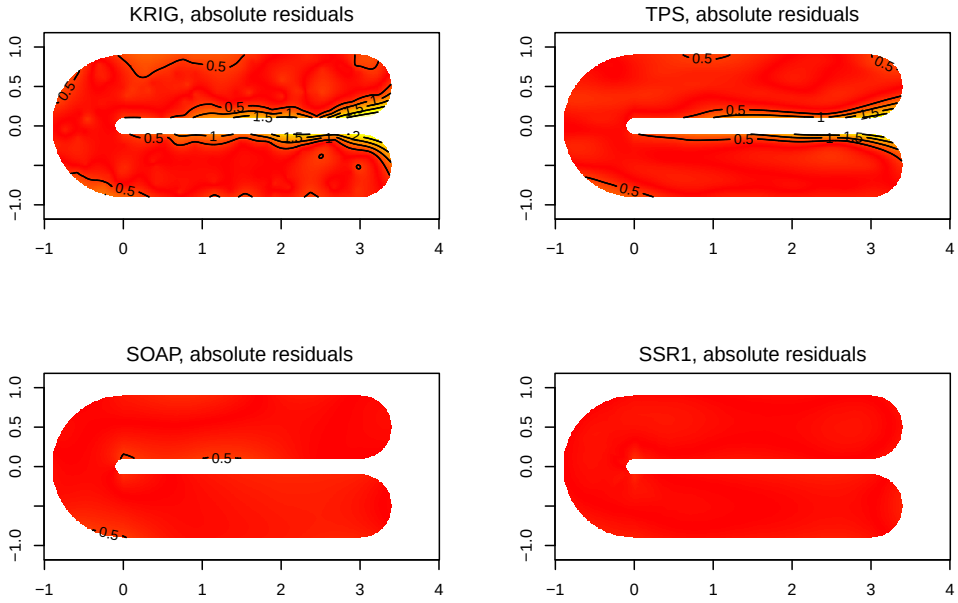

Figure 10. Absolute residuals of f estimates shown in Figure 9, with respect to true surface function f.

followed by those given by SOAP; these two methods provide significantly better estimates than TPS and KRIG. Figures 9 and 10 illustrate the reason of this comparative advantage; these figures show respectively the estimated functions provided by the different methods in the first simulation replicate (Figure 9), and the corresponding absolute residuals with respect to true f, always in the first replicate (Figure 10). Figure 10, in particular, shows that KRIG and TPS estimates have high absolute errors near the inner borders of the C arms; this "leakage effect" in KRIG and TPS estimates is due to the fact that these two methods do not take into any account domain boundaries, but instead smooth across them.

Individual 95% confidence intervals for β_1 and β_2 for the linear SSR estimate obtained in the first replicate, are respectively given by $[-0.521\,, -0.410]$ and $[0.171\,, 0.201]$ (the residuals are normally distributed, with Shapiro-Wilk normaly test p-value 0.86). Figure 6 displays pointwise 95% confidence regions for the corresponding f estimate.

6.2. Simulations with a test function having strong gradients at the boundary

We now show the results of a simulation study with the test function shown in Figures 11 and 12. Wood et al. (2008) used this test function, characterized by extremely strong gradients at the domain boundary, to show that in this case soap-film smoothing give a comparative advantage over the original bivariate spline smoother by Ramsay (2002), thanks to its additional specific smoothing of the solution along the boundary. We here show that the proposed SSR models are efficient even in the case of extremely strong gradients at domain boundary, yielding in this case comparable results to soap-film smoothing.

For $N = 50$ replicates, we sample data from (13) on a regular lattice of $n = 404$ points, shown in Figure 12, left panel, and with about 17% of the data points having distance from the domain boundary less than 0.05. The parameter values used in the simulation

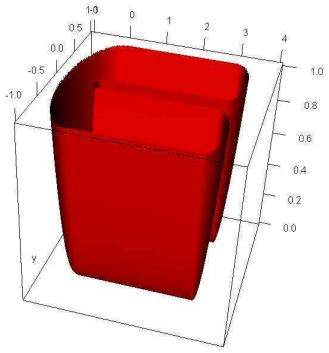

Figure 11. 3D image of the test function with extremely strong gradients at the domain boundary.

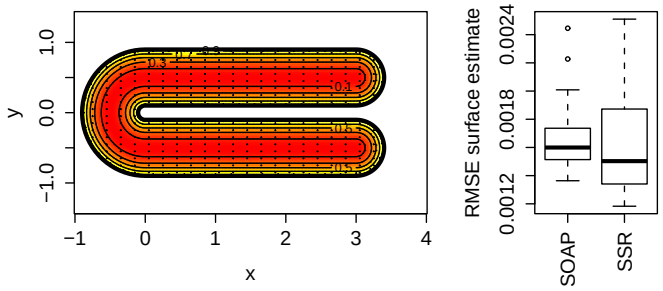

Figure 12. Left: Color map of the test function in Figure 11, with lattice of points used in the simulations. Right: boxplots of the RMSE of the estimates of f, evaluated on a fine grid of step 0.04, provided by soap-film smoothing and SSR models.

are: $\beta_1 = -0.2$, $\beta_2 = 0.1$, $\sigma_\epsilon = 0.05$, $\mu_1 = 0.3$, $\sigma_1 = 0.15$, $\mu_2 = 0.7$, $\sigma_2 = 0.5$. The RMSE for the estimates of the parameters $\beta_1, \beta_2, \sigma_\epsilon$ obtained by soap-film smoothing and SSR models are completely comparable (equal in fact when approximated to the third decimal digit). The boxplots of the RMSE of the estimate of f, evaluated on a fine grid of step 0.04, are displayed in the right panel of Figure 12. These results show that the improved spatial spline smoother here considered gives comparable estimates to soap-film smoothing in this example with a test function having extremely strong gradients at the boundary. Nevertheless, the extra-regularization of the solution along the boundary performed by soap-film smoothing would certainly help in reducing the error at the boundary if no data were observed close to the domain boundary.

6.3. Simulations on a rectangular domain

We finally consider a simulation case where there are no problems with irregularly shaped domains and complex boundaries, so that classical methods such as TPS would not encounter the difficulties highlighted in Section 6.1. This simulation study show that SSR models are as good as TPS in the case of simple domains. Moreover, in presence of information on the values of the surface at the domain boundary, this information can be included in the SSR model as described in Section 5, leading of course to significantly better estimates.

We consider the function f, defined on a square domain, shown in the top left panel of Figure 13. For $N = 50$ replicates, we sample data from (13) on a regular lattice of $n = 400$ points, using the following parameter values: $\beta_1 = -0.5$, $\beta_2 = 0.25$, $\sigma_\epsilon = 0.15$, $\mu_1 = 0.3$, $\sigma_1 = 0.5$, $\mu_2 = 0.7$, $\sigma_2 = 1.3$. The RMSE for the estimates of the parameters $\beta_1, \beta_2, \sigma_\epsilon$ obtained by TPS, SSR models and SSR models with known boundary value conditions (SSRbc) are completely comparable, equal in fact when approximated to the third decimal digit. The boxplots of the RMSE of the estimate of f, evaluated on a fine grid of step 0.03, are displayed in the bottom right panel of Figure 13: the estimates of f provided by TPS and SSR are comparable; SSRbc efficiently use the extra information about known boundary values, resulting in significantly better estimates.

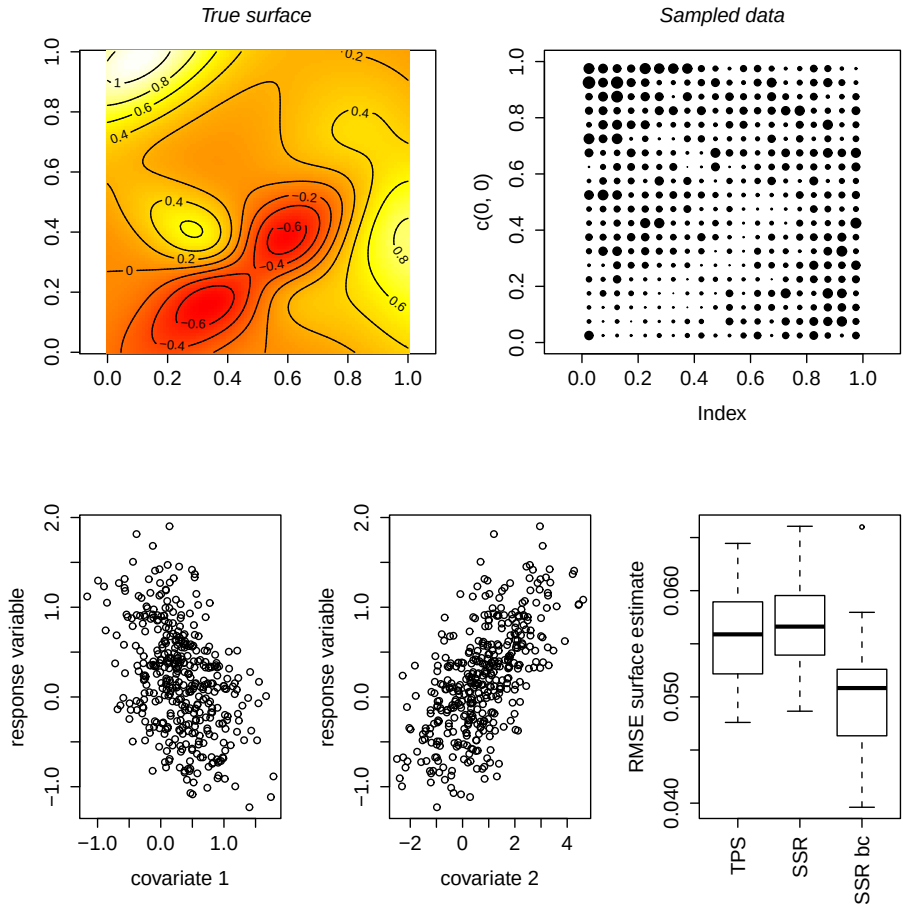

Figure 13. Top left: test function on a simple domain. Top right: sampled data on a regular lattice (size of the point marker proportional to sampled data value), replicate 1. Bottom left and center: scatter plots of response variable vs covariates, replicate 1. Right: boxplots of the RMSE of the estimate of the function f, evaluated on a fine grid of step 0.03, for TPS, SSR models and SSR with known boundary values (SSR bc).

7. Applications

7.1. Island of Montréal census data

We consider the problem of estimating population density over the Island of Montréal. The data are derived from the 1996 Canadian census. Figure 1 show the municipality of Montréal, along with 493 data points defined by the centroids of census enumeration areas. Population density is available at each census tract, measured as 1000 inhabitants per km^2. As covariate, we use the binary variable indicating whether a tract is predominantly residential (1) or commercial/industrial (0). Figures 1 and 14 highlight two areas that are not part the domain of interest for the study of population density: the Montréal airport, with some surrounding services and industries, in the south-wester part of the island, and the industrial park in the north-eastern tip, that includes an oil refinery tank farm and a water purification plant. As mentioned in the introduction, it is known that no people live by the

river banks in correspondence of the harbour (east shore) and the public parks (south-west and north-west shore). We hence impose appropriate (i.e., in this case, homogenous zero) Dirichlet boundary conditions along these stretches of coast, highlighted in red in Figures 1 and 14. Figure 15 shows the SSR estimate of the spatial variation structure, the non-parametric part of the model. Notice that the estimate complies with the imposed boundary conditions, dropping to zero along those stretches of coast, thus efficiently including this a priori information. Also, the estimate has not artificially linked data points on either side of the uninhabited parts; see for instance the densely populated area just in the south of the oil refinery and purification plant with respect to the low population density neighborhood just in the north-east of this industrial park. The β coefficient that corresponds to the binary covariate is estimated to be 1.300; this means that census tracts that are predominantly residential are in average expected to have 1300 more inhabitants per km^2, with respect to those classified as mostly commercial/industrial; the approximate 95% confidence interval is given by $[0.755; 1.845]$.

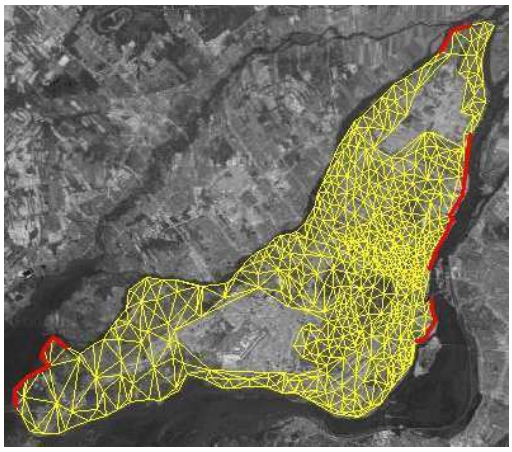

Figure 14. Constrained Delaunay triangulation of the Island of Montréal.

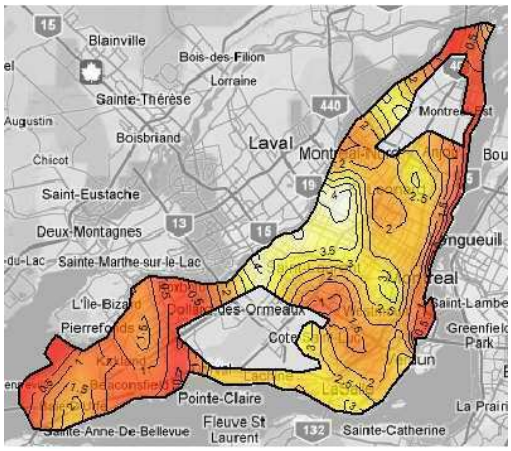

Figure 15. SSR estimate of spatial structure for population density over the Island of Montréal.

7.2. Hemodynamics data

In this section we present a brief application of the proposed technique to the analysis of data of the AneuRisk project†, a scientific endeavour that aimed at investigating the role of vessel morphology, blood fluid dynamics and biomechanical properties of the vascular wall, on the pathogenesis of cerebral aneurysms. The project has gathered together researchers of different scientific fields, ranging from neurosurgery and neuroradiology to statistics, numerical analysis and bio-engineering. See, e.g., Passerini et al. (2012) and Sangalli et al. (2009) for detailed descriptions of the project.

†The project involved MOX Laboratory for Modeling and Scientific Computing (Dip. di Matematica, Politecnico di Milano), Laboratory of Biological Structure Mechanics (Dip. di Ingegneria Strutturale, Politecnico di Milano), Istituto Mario Negri (Ranica), Ospedale Niguarda Ca' Granda (Milano) and Ospedale Maggiore Policlinico (Milano), and has been supported by Fondazione Politecnico di Milano and Siemens Medical Solutions Italia.

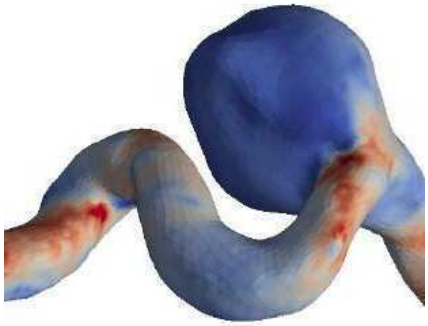

Figure 16. Wall shear stress modulus at the systolic peak, obtained via computational fluid dynamics, on real inner carotid artery geometry affected by aneurysm.

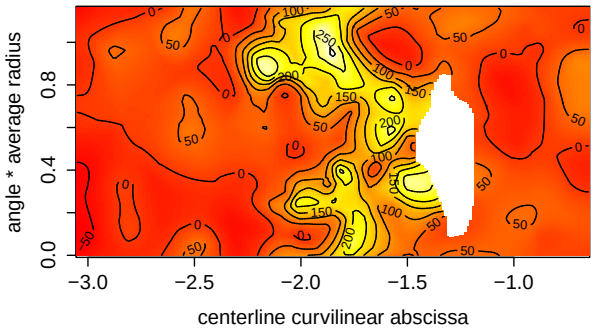

Figure 17. Estimated spatial structure of wall shear stress modulus, using pressure as covariate; the hole in the domain corresponds to the aneurysm neck, since the aneurismal sac is excluded from computations.

Cerebral aneurysms are deformations of cerebral vessels characterized by a bulge of the vessel wall. Figure 16, for instance, show an inner carotid artery, one of the main arteries bringing blood to the brain, affected by an aneurysm. The origin of the aneurysmal pathology is still unclear. Possible explanations that have been discussed in the medical literature focus on interactions between the biomechanical properties of artery walls and effects of hemodynamical forces exerted on the vessel walls, such as wall shear stress and pressure; the hemodynamics are in turn strictly dependent on vascular geometry. Whilst the first studies on the aneurismal pathology restricted their attention to the aneurismal sac, the AneuRisk project analyses have investigated the morphological and hemodynamical features of the parent vasculature, i.e., the vessel hosting the aneurysm and the upstream vasculature, with the goal of highlighting possible causes of aneurysm onset, development and possible rupture. See Passerini et al. (2012) and references therein.

Here in particular we analyze hemodynamical data on a real anatomy of inner carotid artery. These hemodynamical quantities (wall shear stress and pressure) have been simulated in Passerini (2009) via computational fluid dynamics; the Inner Carotid Artery geometry is reconstructed from three-dimensional angiographic images, belonging to the AneuRisk datawarehouse. Figure 16 in particular shows the simulated wall shear stress modulus at the systolic peak on the 3D geometry. The hemodynamical data are referred to points (x, y, z) on the artery wall, the latter being a bidimensional, but not-planar, domain; for this reason, this domain is first flattened via a bijective map. In particular a new coordinate system is considered, defined by (s, r, θ), where s is the curvilinear abscissa along the artery centerline, r the artery radius, and θ the angle of the surface point with respect to the artery centerline; the domain is thus reduced to the plane $(s, \theta * \bar{r})$, where $\bar{r}$ is the average carotid radius on the carotid tract considered. This planar rectangular domain, that is essentially obtained by cutting the artery wall along the axial direction given by s, and then by opening and flattening the artery wall, is equivalent to a simplified 3D artery geometry, where the radius has been fixed to a constant value. We shall exclude from the domain the aneurismal sac, since, as mentioned above, the focus of the analysis is on the parent vessel. Figure 17 shows the estimated spatial structure of wall shear stress modulus, using pressure as covariate; the figure only displays the carotid tract closer to the aneurysm neck, which here corresponds to an hole in the domain, since the aneurysmal sac has been excluded from

computations. Likewise in Boneschi (2010) and Boneschi and Vantini (2012), to obtain the necessary $(2\pi * \bar{r})-$periodicity of the estimate along the coordinate $\theta * \bar{r}$, the data have been augmented by two repetitions of the data values, at the same abscissas but with ordinates respectively given by $(\theta * \bar{r} + 2\pi * \bar{r})$ and $(\theta * \bar{r} - 2\pi * \bar{r})$. Figure 17 shows the very high values of the estimated spatial variation on the left of the aneurysm neck, i.e., upstream of the aneurysm, with respect to the rather low values on the right end side, i.e., downstream of the aneurysm. This very important feature of the hemodynamic regime is in large part lost if the domain is not correctly taken into account, resulting in distorted estimates of WSS in the area of highest interest, i.e., around the aneurysm neck. After appropriate registration of the patient-specific inner carotid artery geometries, the estimated spatial structures of wall shear stress could for instance undergo the functional principal component analysis of surfaces described in Boneschi (2010) and Boneschi and Vantini (2012), in order to detect recurrent hemodynamic patterns, common across patients, and relate them to presence and location of the pathology, and to rupture risk. For a detailed account of the analyses conducted within the AneuRisk project, see http://ecm2.mathcs.emory.edu/aneurisk.

8. Discussion

The described technique can be generalized in several directions, for instance to the case of repeated observations at each data location, to areal data (see Azzimonti et al. 2011), to general link functions such as the logit, to loss-functions other than the classical sum-of-squared-errors, thus allowing for very large potential application. The covariates themselves, when having a spatial structure, could be modeled as surfaces, in a functional regression model setting.

An important line of investigation within SSR framework concerns extensions of the roughness term to more complex partial differential operators; this model extension is particularly interesting for applications where some a priori knowledge of the problem (physical, chemical, mechanical or morphological) suggests the choice of a partial differential operator modeling to some extent the phenomenon under study. Azzimonti et al. (2011) investigates this research direction, generalizing SSR to penalties involving general second order elliptic operators. The applied problem there considered concerns the estimation of the blood-flow velocity field in a section of a carotid artery, using data provided by eco-doppler images. In this application, physiological knowledge of the phenomenon suggests the theoretical shape of the velocity field; this knowledge is thus translated into a partial differential operator and used in the roughness penalty in order to regularize the estimate. This extension also constitutes a very promising line of research toward the modeling of spatio-temporal phenomena, including dynamical functional data, such as curves and surfaces deforming over time. Moreover, an approach analogous to the one described here, could also be applied to the problem of parameter estimation for partial differential equations, and we will be exploring possible extensions in this direction of the work of Ramsay et al. (2007).

Finally, this model can also be extended to three dimensions, dealing with three-dimensional surfaces and volumes. Such model extension would have a strong impact on forefront applications, as for instance those concerning the analysis of three-dimensional images of the internal structures of a body provided by diagnostic medical scanners (angiographies, tomographies, magnetic resonance imaging devices, etc.). It is apparent that when analyzing these data it would be in many contexts desirable to use techniques that takes into account the boundary of the problem. For instance, organs such as brain have complex boundaries,

both external and internal, and when studying brain imaging data it would certainly be of great interest to be able to comply efficiently with those boundaries.

The R code for the implementation of the proposed models (including domain triangulation), as well as reproducing the examples shown in the paper, is fully integrated with the fda package; the same functions have also been coded in Matlab. Both versions shall be shortly released.

Acknowledgments. L. Sangalli acknowledges funding by MIUR Ministero dell'Istruzione dell'Università e della Ricerca, *FIRB Futuro in Ricerca* research project "Advanced statistical and numerical methods for the analysis of high dimensional functional data in life sciences and engineering", and by the program Dote Ricercatore Politecnico di Milano - Regione Lombardia, research project "Functional data analysis for life sciences". J. Ramsay acknowledges funding by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada. Part of this work has been carried out when L. Sangalli and J. Ramsay where visiting SAMSI Statistical and Applied Mathematical Sciences Institute, Research Triangle Park, North Carolina, USA. We are grateful to Laura Azzimonti, Simona Perotto and Giancarlo Sangalli for help on the numerical analysis aspects of this work, and to Fabio Nobile and Piercesare Secchi for thoughtful and constructive comments.

A. Appendix

A.1. Proof of Proposition 1

The functional $J_\lambda(\boldsymbol{\beta}, f)$ in (2) can be rewritten as

$$J_\lambda(\boldsymbol{\beta}, f) = \left(\mathbf{z} - \mathbf{W}\boldsymbol{\beta} - \mathbf{f}_n\right)^t \left(\mathbf{z} - \mathbf{W}\boldsymbol{\beta} - \mathbf{f}_n\right) + \lambda \int_\Omega (\Delta f)^2.$$

Given f, the unique minimizer $\hat{\boldsymbol{\beta}}(f)$ of $J_\lambda(\boldsymbol{\beta}, f)$ with respect to $\boldsymbol{\beta}$ is given by

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

Plugging $\hat{\boldsymbol{\beta}}(f)$ into $J_\lambda(\boldsymbol{\beta}, f)$ we obtain

$$J_\lambda(\hat{\boldsymbol{\beta}}(f), f) = \mathbf{z}^t\,\mathbf{Q}\,\mathbf{z} - 2\,\mathbf{f}_n^t\,\mathbf{Q}\,\mathbf{z} + \mathbf{f}_n^t\,\mathbf{Q}\,\mathbf{f}_n + \lambda \int_\Omega (\Delta f)^2\,.$$

Our estimation problem is thus reduced to an optimization problem over f only: find $f \in H^2_{n0}(\Omega)$ that minimizes

$$\widetilde{J}_\lambda(f) = \left[\mathbf{f}_n^t\,\mathbf{Q}\,\mathbf{f}_n + \lambda \int_\Omega (\Delta f)^2\right] - 2\,\mathbf{f}_n^t\,\mathbf{Q}\,\mathbf{z} \tag{14}$$

where we have dropped the term not depending on f.

The proof is completed by showing that $\hat{f}$, the minimizer of (14) over $H^2_{n0}(\Omega)$, is unique and satisfies (3) for every $u \in H^2_{n0}(\Omega)$. To prove this result, we exploit a characterization theorem (see e.g., Braess 2007, Chapter 2) which states that if G is a symmetric, positive definite, bilinear form on a vector space V, and F is a linear functional on V, then v is the unique minimizer of

$$G(v, v) - 2F(v)$$

in V if and only if

$$G(v,u) = F(u) \qquad \text{for all } u \in V. \tag{15}$$

Moreover, there is at most one solution of (15).

The desired result follows immediately from application of the above theorem considering: $V = H^2_{n0}(\Omega)$; the symmetric, positive definite, bilinear form

$$G(f,u) := \mathbf{u}_n^t \, \mathbf{Q} \, \mathbf{f}_n + \lambda \int_\Omega (\Delta u)(\Delta f) \tag{16}$$

and the linear functional $F(f) = \mathbf{f}_n^t \, \mathbf{Q} \, \mathbf{z}$.

Positive definitiveness of the form G in (16), on $H^2_{n0}(\Omega)$, is shown by the following argument. Suppose $G(f,f) = 0$ for some $f \in H^2_{n0}(\Omega)$; then $\int_\Omega (\Delta f)^2 = 0$ and $\mathbf{f}_n^t \, \mathbf{Q} \, \mathbf{f}_n = 0$. The fact that $\int_\Omega (\Delta f)^2 = 0$, for $f \in H^2_{n0}(\Omega)$, implies that f is a constant function on Ω. Thus, let $f(\cdot) \equiv c$ on Ω, for some constant c. Then $\mathbf{f}_n^t \, \mathbf{Q} \, \mathbf{f}_n = c^2 \, \mathbf{1}_n^t \, \mathbf{Q} \, \mathbf{1}_n$, where $\mathbf{1}_n$ is the n-vector having all entries equal to 1. Note that $\mathbf{1}_n^t \, \mathbf{Q} \, \mathbf{1}_n \neq 0$ because $\mathbf{1}_n$ does not belong to the subspace of $\mathbb{R}^n$ generated by the columns of $\mathbf{W}$, and hence the two vectors $\mathbf{1}_n$ and $\mathbf{Q} \, \mathbf{1}_n$ are not orthogonal. Therefore, $\mathbf{f}_n^t \, \mathbf{Q} \, \mathbf{f}_n = 0$ implies that $c = 0$, so that $f(\cdot) \equiv 0$ on Ω, i.e., G is positive definite on $H^2_{n0}(\Omega)$. This concludes the proof.

A.2. Derivation of problem (5)

In this Section we derive the reformulation of the estimation problem given in (5), that constitutes our base for the implementation of the finite element method.

The problem of finding $\hat{f} \in H^2_{n0}(\Omega)$ that satisfies (3) for every $u \in H^2_{n0}(\Omega)$, can be rewritten as the problem of finding $(\hat{f}, g) \in H^2_{n0}(\Omega) \times L^2(\Omega)$ that satisfy

$$\begin{aligned} \mathbf{u}_n^t \, \mathbf{Q} \, \hat{\mathbf{f}}_n + \lambda \int_\Omega g \, (\Delta u) &= \mathbf{u}_n^t \, \mathbf{Q} \, \mathbf{z} \\ \int_\Omega g \, v - \int_\Omega (\Delta \hat{f}) \, v &= 0 \end{aligned} \tag{17}$$

for all $(u,v) \in H^2_{n0}(\Omega) \times L^2(\Omega)$. In fact, if the pair of functions $(\hat{f}, g) \in H^2_{n0}(\Omega) \times L^2(\Omega)$ satisfies (17) for all $(u,v) \in H^2_{n0}(\Omega) \times L^2(\Omega)$, then $\hat{f}$ also satisfies (3). On the other hand, if $\hat{f} \in H^2_{n0}(\Omega)$ satisfies (3), then the pair $(\hat{f}, \Delta\hat{f})$ automatically satisfies the two equations in (17). Now, asking a slightly higher regularity of the auxiliary function g and of the test functions v, namely $g, v \in H^1(\Omega)$ instead of $g, v \in L^2(\Omega)$, the problem above may in turn be reformulated as the problem of finding $(\hat{f}, g) \in \left(H^1_{n0}(\Omega) \cap C^0(\Omega)\right) \times H^1(\Omega)$ that satisfies (17) for all $(u,v) \in \left(H^1_{n0}(\Omega) \cap C^0(\Omega)\right) \times H^1(\Omega)$. Moreover, the theory of problems of elliptic regularity ensure that such $\hat{f}$ still belongs to $H^2_{n0}(\Omega)$ (see, e.g., Lions and Magenes 1973, Chapter 8). Reformulating the problem as a problem in $H^1(\Omega)$ is done here in view of its discretization via the finite element space $H^1_{\mathcal{T}}(\Omega)$. Finally, Green's theorem or integration by parts, together with the boundary condition on the normal derivatives of f, implies that the latter reformulation is equivalent to the one given in (5).

A.3. Proof of Corollary 1

Problem (5) has the following discrete counterpart: find $(f,g) \in H^1_{\mathcal{T}}(\Omega) \times H^1_{\mathcal{T}}(\Omega)$ that satisfy (5) for all $(u,v) \in H^1_{\mathcal{T}}(\Omega) \times H^1_{\mathcal{T}}(\Omega)$, with the integrals now computed over $\Omega_{\mathcal{T}}$. Exploiting

property (4), we have that for any functions f, g, v in $H^1_{\mathcal{T}}(\Omega)$

$$\int_{\Omega_{\mathcal{T}}} v\, g = \mathbf{v}^t\, \mathbf{R}_0\, \mathbf{g} \qquad \text{and} \qquad \int_{\Omega_{\mathcal{T}}} (\nabla v \cdot \nabla f) = \mathbf{v}^t\, \mathbf{R}_1\, \mathbf{f}.$$

Moreover,

$$\mathbf{u}_n^t\, \mathbf{Q}\, \mathbf{f}_n = \mathbf{u}^t\, \mathbf{L}\, \mathbf{f} \qquad \text{and} \qquad \mathbf{u}_n^t\, \mathbf{Q}\, \mathbf{z} = \mathbf{u}^t \mathbf{L}\, \mathbf{D}\, \mathbf{z}.$$

Using these identities, solving the discrete counterpart of the estimation problem reduces to finding a couple of vectors $(\mathbf{f}, \mathbf{g}) \in \mathbb{R}^K \times \mathbb{R}^K$ that satisfy

$$\begin{aligned} &\mathbf{u}^t\, \mathbf{L}\, \mathbf{f} - \lambda\, \mathbf{u}^t\, \mathbf{R}_1\, \mathbf{g} = \mathbf{u}^t\, \mathbf{L}\, \mathbf{D}\, \mathbf{z} \\ &\mathbf{v}^t\, \mathbf{R}_0\, \mathbf{g} + \mathbf{v}^t\, \mathbf{R}_1\, \mathbf{f} = 0 \end{aligned} \tag{18}$$

for all $(\mathbf{u}, \mathbf{v}) \in \mathbb{R}^K \times \mathbb{R}^K$. Since (18) must hold for all $(\mathbf{u}, \mathbf{v}) \in \mathbb{R}^K \times \mathbb{R}^K$, this is equivalent to finding $(\mathbf{f}, \mathbf{g}) \in \mathbb{R}^K \times \mathbb{R}^K$ that satisfies the linear system (6), that has been symmetrized for computational convenience. Solving this system (6) provides $\hat{\mathbf{f}}$ and, thanks to (4), $\hat{\mathbf{f}}$ identifies the estimate $\hat{f} \in H^1_{\mathcal{T}}(\Omega)$. Since the system has a unique solution, the finite element solution to the estimation problem is unique. This can be shown working sequentially on the two equations composing the system. From the second equation, exploiting the positive-definiteness and thus invertibility of the matrix $\mathbf{R}_0$ (proved below), we get $\mathbf{g} = -\mathbf{R}_0^{-1}\, \mathbf{R}_1\, \mathbf{f}$. Then plugging this expression for $\mathbf{g}$ into the first equation composing the system, thanks to the fact that $(\mathbf{L} + \lambda \mathbf{R}_1\, \mathbf{R}_0^{-1}\, \mathbf{R}_1)$ is also invertible (as proved below), we obtain the unique solution $\hat{\mathbf{f}} = (\mathbf{L} + \lambda\, \mathbf{R}_1\, \mathbf{R}_0^{-1}\, \mathbf{R}_1)^{-1}\, \mathbf{L}\, \mathbf{D}\, \mathbf{z}$.

Positive-definiteness of $\mathbf{R}_0$ follows immediately from the fact that $\mathbf{R}_0$ is the Gramm matrix associated to set of K linearly independent vectors $\big(\psi_k(\boldsymbol{\xi}_1), \ldots, \psi_k(\boldsymbol{\xi}_K)\big)^t$, for $k = 1, \ldots, K$, in the inner product space $(H^1_{\mathcal{T}}(\Omega), <v_1, v_2>)$, where where $<v_1, v_2> = \int_\Omega v_1\, v_2\, d\Omega$.

Positive-definiteness of $(\mathbf{L} + \lambda\, \mathbf{R}_1\, \mathbf{R}_0^{-1}\, \mathbf{R}_1)$ is proved by the following argument. Since $\mathbf{L}$ is positive semi-definite by construction, and $\mathbf{R}_0$, and thus $\mathbf{R}_0^{-1}$, are positive definite, then $(\mathbf{L} + \lambda \mathbf{R}_1\, \mathbf{R}_0^{-1}\, \mathbf{R}_1)$ is at least positive semi-definite. Suppose now that $\mathbf{c}^t\, (\mathbf{L} + \lambda \mathbf{R}_1\, \mathbf{R}_0^{-1}\, \mathbf{R}_1)\, \mathbf{c} = 0$ for some K-vector $\mathbf{c}$. Since

$$0 = \mathbf{c}^t\, (\mathbf{L} + \lambda\, \mathbf{R}_1\, \mathbf{R}_0^{-1}\, \mathbf{R}_1)\, \mathbf{c} = \mathbf{c}^t\, \mathbf{L}\, \mathbf{c} + \lambda\, \mathbf{c}^t \mathbf{R}_1\, \mathbf{R}_0^{-1}\, \mathbf{R}_1\, \mathbf{c}$$

and both terms on the right-hand side are non-negative, it follows that $\mathbf{c}^t\, \mathbf{L}\, \mathbf{c} = 0$ and $\mathbf{c}^t \mathbf{R}_1\, \mathbf{R}_0^{-1}\, \mathbf{R}_1\, \mathbf{c} = 0$. Being $\mathbf{R}_0^{-1}$ positive definite, the latter implies that $\mathbf{R}_1\, \mathbf{c} = \mathbf{0}$, and this in turn implies that $\mathbf{c}^t\, \mathbf{R}_1\, \mathbf{c} = 0$. Hence

$$\begin{aligned} 0 = \mathbf{c}^t\, \mathbf{R}_1\, \mathbf{c} &= \mathbf{c}^t \bigg\{ \int_\Omega (\boldsymbol{\psi}_x \boldsymbol{\psi}_x^t + \boldsymbol{\psi}_y \boldsymbol{\psi}_y^t) \bigg\} \mathbf{c} = \int_\Omega \big\{ \mathbf{c}^t (\boldsymbol{\psi}_x \boldsymbol{\psi}_x^t + \boldsymbol{\psi}_y \boldsymbol{\psi}_y^t) \mathbf{c} \big\} \\ &= ||\mathbf{c}^t \boldsymbol{\psi}_x||_2^2 + ||\mathbf{c}^t \boldsymbol{\psi}_y||_2^2 \end{aligned}$$

where $||\cdot||_2$ denotes the L^2-norm, and this implies that $||\mathbf{c}^t \boldsymbol{\psi}_x||_2^2 = 0$ and $||\mathbf{c}^t \boldsymbol{\psi}_y||_2^2 = 0$. Thus, both partial derivatives of the piecewise quadratic finite element function $\mathbf{c}^t\, \boldsymbol{\psi}$ vanish, and this means that $\mathbf{c}^t\, \boldsymbol{\psi}$ is a constant function. Since the entries of the vector $\mathbf{c}$ are the values of the function $\mathbf{c}^t\, \boldsymbol{\psi}$ evaluated at the K nodes, $\mathbf{c}$ must have the form $(c, \ldots, c)^t$ for some real constant c, i.e., $\mathbf{c} = c\, \mathbf{1}_k$. Then,

$$\mathbf{c}^t\, \mathbf{L}\, \mathbf{c} = c^2\, \mathbf{1}_n\, \mathbf{Q}\, \mathbf{1}_n$$

Exploiting this identity, the fact that $\mathbf{c}^t\,\mathbf{L}\,\mathbf{c} = 0$ implies that $c = 0$, since $\mathbf{1}_n\,\mathbf{Q}\,\mathbf{1}_n \neq 0$, as already noticed in Appendix A.1. It follows that $\mathbf{c} = \mathbf{0}$, and this proves that the matrix $(\mathbf{L} + \lambda \mathbf{R}_1\,\mathbf{R}_0^{-1}\,\mathbf{R}_1)$ is positive definite.

A.4. Proof of Proposition 2

Likewise in the proof of Proposition 1, we reduce the estimation problem to an optimization problem over f only: find f that minimizes the functional $\widetilde{J}_\lambda(f)$ in (14) over $f \in H^2(\Omega)$ with $f|_{\partial\Omega} = f_{\partial\Omega}$. Now fix $\bar{f}$ to be any function in $H^2(\Omega)$ such that $\bar{f}|_{\partial\Omega} = f_{\partial\Omega}$ and $\bar{f}$ has known gradient $\nabla\bar{f}$ and Laplacian $\Delta\bar{f}$. If $f_{\partial\Omega}$ is sufficiently regular, e.g., $f_{\partial\Omega} \in H^{3/2}(\partial\Omega)$, this function exists. Moreover, let s be the function in $H_0^2(\Omega)$ such that $f = s + \bar{f}$. Holding $\bar{f}$ fixed, our optimization problem becomes finding $s \in H_0^2(\Omega)$ that minimizes

$$G_\lambda(s) = \left[(\mathbf{s}_n^t + \overline{\mathbf{f}}_n^t)\,\mathbf{Q}\,(\mathbf{s}_n + \bar{\mathbf{f}}_n) + \lambda \int_\Omega (\Delta s + \Delta\bar{f})^2 d\Omega\right] - 2\,(\mathbf{s}_n^t + \overline{\mathbf{f}}_n^t)\,\mathbf{Q}\,\mathbf{z}\,.$$

This is in turn equivalent to finding $s \in H_0^2(\Omega)$ that minimizes

$$\tilde{G}_\lambda(s) = \left[\mathbf{s}_n^t\,\mathbf{Q}\,\mathbf{s}_n + \lambda \int_\Omega (\Delta s)^2 d\Omega\right] - 2\left\{\mathbf{s}_n^t\,\mathbf{Q}\,(\mathbf{z} - \overline{\mathbf{f}}_n^t) - \lambda \int_\Omega (\Delta s)(\Delta\bar{f})\right\} \tag{19}$$

where we have dropped the terms constant with respect to s. To complete the proof, we need to show that the function $\hat{s}$ is the unique minimizer over $H_0^2(\Omega)$ of $\widetilde{G}_\lambda(s)$ in (19) if and only if it satisfies (10) for every $u \in H_0^2(\Omega)$. This result is proved resorting to the characterization theorem already evoked in Appendix A.1 and considering: the vector space $V = H_0^2(\Omega)$; the symmetric positive definite bilinear form $G(s,u) := \mathbf{u}_n^t\,\mathbf{Q}\,\mathbf{s}_n + \lambda \int_\Omega (\Delta u)(\Delta s)$; the linear functional $F(s) = \mathbf{s}_n^t\,\mathbf{Q}\,(\mathbf{z} - \overline{\mathbf{f}}_n^t) - \lambda \int_\Omega (\Delta s)(\Delta\bar{f})$.

A.5. Derivation of problem (11)

Likewise in Appendix A.2, the problem of finding $s \in H_0^2(\Omega)$ that satisfies (10) for every $u \in H_0^2(\Omega)$ can be rewritten as follows: find $(\hat{s}, g) \in H_0^2(\Omega) \times L^2(\Omega)$ that satisfies

$$\begin{aligned} &\mathbf{u}_n^t\,\mathbf{Q}\,\hat{\mathbf{s}}_n + \lambda \int_\Omega g\,(\Delta u) = \mathbf{u}_n^t\,\mathbf{Q}\,(\mathbf{z} - \overline{\mathbf{f}}_n^t) \\ &\int_\Omega g\,v - \int_\Omega (\Delta\hat{s} + \Delta\bar{f})\,v = 0 \end{aligned} \tag{20}$$

for all $(u, v) \in H_0^2(\Omega) \times L^2(\Omega)$. Reasoning as in Appendix A.2, we hence notice that, asking a slightly higher regularity of the auxiliary function g and of the test functions v, namely $g, v \in H^1(\Omega)$ instead of $g, v \in L^2(\Omega)$, the problem above may in turn be reformulated as the problem of finding $(\hat{s}, g) \in \left(H_0^1(\Omega) \cap C^0(\Omega)\right) \times H^1(\Omega)$ that satisfies (20) for all $(u, v) \in \left(H_0^1(\Omega) \cap C^0(\Omega)\right) \times H^1(\Omega)$; moreover, the theory of problems of elliptic regularity ensures that such $\hat{s}$ still belongs to $H_0^2(\Omega)$ (see, e.g., Lions and Magenes 1973, Chapter 8). Also in this case, this is done in view of the discretization of the problem via the finite element space $H_{\mathcal{T}}^1(\Omega)$. Finally, Green's theorem now implies that the latter reformulation is equivalent to the one given in (11).

A.6. Proof of Corollary 2

It is convenient to choose $\bar{f}$ to be the finite element function that coincides with $f_{\partial\Omega}$ on the boundary nodes and is 0 on the m interior nodes; recall that this is a datum in our system. Problem (11) has the following discrete counterpart: find $(\hat{s}, g) \in H^1_{0,\mathcal{T}}(\Omega) \times H^1_{\mathcal{T}}(\Omega)$ that satisfies (11) for all $(u, v) \in H^1_{0,\mathcal{T}}(\Omega) \times H^1_{\mathcal{T}}(\Omega)$, with the integrals now computed over $\Omega_{\mathcal{T}}$. Here $H^1_{0,\mathcal{T}}(\Omega)$ is the subspace of $H^1_{\mathcal{T}}(\Omega)$ consisting of the finite element functions that equal zero on boundary nodes; this space is spanned by the m nodal basis functions corresponding to m interior nodes. The finite element functions $s, u \in H^1_{0,\mathcal{T}}(\Omega)$ are thus identified by vectors of $\mathbb{R}^m$ having as entries the values of these functions at the m interior nodes. Hence, problem above reduces to finding a couple of vectors $(\mathbf{s}, \mathbf{g}) \in \mathbb{R}^m \times \mathbb{R}^K$ that satisfies

$$\begin{aligned} &\mathbf{u}^t \widetilde{\mathbf{L}}\, \mathbf{s} - \lambda\, \mathbf{u}^t \widetilde{K}_1\, \mathbf{g} + \lambda\, \mathbf{u}^t \widetilde{K}_\nu\, \mathbf{g} = \mathbf{u}^t \widetilde{\mathbf{L}}\, \mathbf{z} \\ &\mathbf{v}^t \mathbf{R}_0\, \mathbf{g} + \mathbf{v}^t \widetilde{K}_1^t\, \mathbf{s} - \mathbf{v}^t \widetilde{K}_\nu^t\, \mathbf{s} = -\mathbf{v}^t \mathbf{R}_1\, \bar{\mathbf{f}} + \mathbf{v}^t \mathbf{R}_\nu^t\, \bar{\mathbf{f}} \end{aligned} \tag{21}$$

for all $(\mathbf{u}, \mathbf{v}) \in \mathbb{R}^m \times \mathbb{R}^K$. Since (21) must hold for all $(\mathbf{u}, \mathbf{v}) \in \mathbb{R}^m \times \mathbb{R}^K$, this is equivalent to finding $(\mathbf{s}, \mathbf{g}) \in \mathbb{R}^m \times \mathbb{R}^K$ satisfying the linear system (21). Solving the system provides the solution $\hat{\mathbf{s}} \in \mathbb{R}^m$ that identifies a finite element function $\hat{s} \in H^1_{0,\mathcal{T}}(\Omega)$. The estimate $\hat{f}$, having the required fixed values at the domain boundary, is thus given by $\hat{f} = \hat{s} + \bar{f}$.

References

Azzimonti, L., Domanin, M., Sangalli, L. M., and Secchi, P. (2011), "Surface estimation via spatial spline models with PDE penalization," in *Proceedings of S.Co.2011 Conference*.

Boneschi, A. (2010), "Functional Data Analysis of CFD Simulations: the Systolic Wall Shear Stress Map of the Internal Carotid Artery," Master Thesis. Dipartimento di Matematica, Politecnico di Milano, Italy. Available at http://mox.polimi.it/it/progetti/pubblicazioni/tesi/boneschi.pdf

Boneschi, A. and Vantini, S. (2012), "Functional Data Analysis of CFD Simulations: the Systolic Wall Shear Stress Map of the Internal Carotid Artery," Tech. rep., MOX, Dipartimento di Matematica, Politecnico di Milano.

Braess, D. (2007), *Finite elements*, Cambridge: Cambridge University Press, 3rd ed., theory, fast solvers, and applications in elasticity theory.

Buja, A., Hastie, T., and Tibshirani, R. (1989), "Linear smoothers and additive models," *Ann. Statist.*, 17, 453–555.

Furrer, R., Nychka, D., and Sain, S. (2010), *fields: Tools for spatial data*, r package version 6.3.

Gockenbach, M. S. (2006), *Understanding and implementing the finite element method*, Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM).

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

Hjelle, Ø. and Dæhlen, M. (2006), *Triangulations and applications*, Mathematics and Visualization, Berlin: Springer-Verlag.

Lai, M.-J. and Schumaker, L. L. (2007), *Spline functions on triangulations*, vol. 110 of *Encyclopedia of Mathematics and its Applications*, Cambridge: Cambridge University Press.

Lindgren, F., Håvard, R., and Lindström, 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, with discussions and a reply by the authors.

Lions, J.-L. and Magenes, E. (1973), *Non-homogeneous boundary value problems and applications. Vol. III*, New York: Springer-Verlag.

Marra, G., Miller, D., and Zanin, L. (2011), "Modelling the spatiotemporal distribution of the incidence of resident foreign population," *Statistica Neerlandica*, in press.

Passerini, T. (2009), "Computational hemodynamics of the cerebral circulation: multiscale modeling from the circle of Willis to cerebral aneurysms," PhD Thesis. Dipartimento di Matematica, Politecnico di Milano, Italy. Available at http://mathcs.emory.edu/.

Passerini, T., Sangalli, L., Vantini, S., Piccinelli, M., Bacigaluppi, S., Antiga, L., Boccardi, E., Secchi, P., and Veneziani, V. (2012), "An Integrated Statistical Investigation of Internal Carotid Arteries of Patients affected by Cerebral Aneurysms," *Cardiovascular Engineering and Technology*, in press.

Quarteroni, A. (2009), *Numerical models for differential problems*, vol. 2 of *MS&A. Modeling, Simulation and Applications*, Springer-Verlag Italia, Milan.

Ramsay, J. O., Hooker, G., Campbell, D., and Cao, J. (2007), "Parameter estimation for differential equations: a generalized smoothing approach," *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 69, 741–796, with discussions and a reply by the authors.

Ramsay, T. (2002), "Spline Smoothing over Difficult Regions," *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 64, 307–319.

Sangalli, L. M., Secchi, P., Vantini, S., and Veneziani, A. (2009), "A Case Study in Exploratory Functional Data Analysis: Geometrical Features of the Internal Carotid Artery," *J. Amer. Statist. Assoc.*, 104, 37–48.

Wood, S. (2010), *soap: Soap film smoothing*, r package version 0.1-4.

Wood, S. N., Bravington, M. V., and Hedley, S. L. (2008), "Soap film smoothing," *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 70, 931–955.

MOX Technical Reports, last issues

Dipartimento di Matematica "F. Brioschi",
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

08/2012 Sangalli, L.M.; Ramsay, J.O.; Ramsay, T.O.
Spatial Spline Regression Models

07/2012 Perotto, S; Zilio, A.
Hierarchical model reduction: three different approaches

06/2012 Micheletti, S.; Perotto, S.
Anisotropic recovery-based a posteriori error estimators for advection-diffusion-reaction problems

05/2012 Ambrosi, D; Arioli, G; Koch, H.
A homoclinic solution for excitation waves on a contractile substratum

04/2012 Tumolo, G.; Bonaventura, L.; Restelli, M.
A semi-implicit, semi-Lagrangian, p-adaptive Discontinuous Galerkin method for the shallow water equations

03/2012 Fumagalli, A.; Scotti, A.
A reduced model for flow and transport in fractured porous media with non-matching grids

02/2012 Arioli, G.
Optimization of the forcing term for the solution of two point boundary value problems

01/2012 Lassila, T.; Manzoni, A.; Quarteroni, A.; Rozza, G.
A reduced computational and geometrical framework for inverse problems in haemodynamics

47/2011 Antonietti, P.F.; Borz, A.; Verani, M.
Multigrid shape optimization governed by elliptic PDEs

46/2011 Migliorati, G.; Nobile, F.; von Schwerin, E.; Tempone, R.
Analysis of the discrete L2 projection on polynomial spaces with random evaluations