Metadata of the chapter that will be visualized online

Chapter Title	Estimating Surfaces and Spatial Fields via Regression Models with Differential Regularization	
Copyright Year	2014	
Copyright Holder	Springer-Verlag	
Author	Family Name	**Sangalli**
	Particle	
	Given Name	**Laura M.**
	Suffix	
	Email	laura.sangalli@polimi.it
Abstract	This paper gives an overview presentation of spatial regression with differential regularization, a novel class of models for the accurate estimation of surfaces and spatial fields, that merges that merges advanced statistical methodology with numerical analysis techniques. Thanks to the combination of potentialities from these two scientific areas, the proposed class of models has important advantages with respect to classical techniques used to analyze spatially distributed data. The models are able to efficiently deal with data distributed over irregularly shaped domains, including non-planar domains, only few methods existing in literature for this type of data structures. Moreover, they can incorporate problem-specific prior information about the spatial structure of the phenomenon under study, with a very flexible modeling of space variation, allowing naturally for anisotropy and non-stationarity. The models have a generalized additive framework with a regularizing term involving a differential quantity of the spatial field. The estimators have good inferential properties; moreover, thanks to the use of numerical analysis techniques, they are computationally highly efficient. The method is illustrated in various applied contexts, including demographic data and medical imaging data.	

Estimating Surfaces and Spatial Fields via Regression Models with Differential Regularization

Laura M. Sangalli

1 Introduction

This work briefly reviews progress to date on spatial regression with differential regularization. These are penalized regression models for the accurate estimation of surfaces and spatial fields, that have regularizing terms involving partial differential operators. Partial Differential Equations (PDEs) are commonly used to describe complex phenomena behavior in many fields of engineering and sciences, including Biosciences, Geosciences and Physical sciences. Here they are used to model the shape of the surface and the spatial variation of the problem under study.

In the simpler context of curve estimation and univariate smoothing problems, the idea of regularization with ordinary differential operators has already proved to be very effective and it is in general playing a central role in the functional data analysis literature. See, e.g., [23]. In the more complex case of surface estimation and spatial smoothing, a few methods have been introduced that use roughness penalties involving simple partial differential operators. A classical example is given by thin-plate-splines, while more recent proposals are offered for instance by Ramsay [22], Wood et al. [27], and Guillas and Lai [15]; see also the applications in [1, 11, 19]. Finally, although in a different framework, the use of simple form of (stochastic) PDEs is also at the core of the Bayesian spatial models introduced by Lindgren et al. [18] and more generally of the larger literature on Bayesian inverse problems [25] and data assimilation in inverse problems [9].

Regression models with partial differential regularizations [3, 4, 8, 12, 24] merge advanced statistical methodology with numerical analysis techniques. Thanks to the interactions of these two areas, the proposed class of models have important

---

L.M. Sangalli (✉)
MOX - Dipartimento di Matematica, Politecnico di Milano, Milano, Italy
e-mail: laura.sangalli@polimi.it

A.M. Paganoni, P. Secchi (eds.), *Advances in Complex Data Modeling and Computational Methods in Statistics*, Contributions to Statistics,
DOI 10.1007/978-3-319-11149-0_13

advantages with respect to classical techniques used in spatial data analysis. Spatial regression with differential regularization is able to efficiently deal with data distributed over irregularly shaped domains with complicated geometries [24]. Moreover, it can comply with specific conditions at the boundaries of the problem domain [3, 24], which is fundamental in many applications to obtain meaningful estimates. The proposed models can also deal with data scattered over general bidimensional Riemannian manifold domains [8, 12]. Moreover, spatial regression with differential regularization has the capacity to incorporate problem-specific prior information about the spatial structure of the phenomenon under study [2–4], formalized in terms of a governing PDE. This also allows for a very flexible modeling of space variation, that accounts naturally for anisotropy and non-stationarity. Space-varying covariate information is included in the models via a semiparametric framework. The estimators have a penalized regression form, they are linear in the observed data values, and have good inferential properties. The use of advanced numerical analysis techniques, and specifically of finite elements, makes the models computationally very efficient. The method is implemented in both R [21] and Matlab.

This work gives a unified summary review of [3, 4, 8, 12, 24] and is organized as follows. Section 2 introduces the model in the simplest setting, characterizes the estimation problem (Sect. 2.2), illustrates how to compute the estimators using finite elements (Sect. 2.3) and reports some distributional properties of the estimators (Sect. 2.4). Section 3 shows how to include in the model prior information about the space variation of the phenomenon under study. Section 4 reviews the works on spatial regression over manifolds. The method is illustrated in various applied contexts in Sects. 2–4, including demographic data and medical imaging data. Finally, Sect. 5 outlines current research directions.

2 Regression Models with Partial Differential Regularization

Let $\{\mathbf{p}_i = (p_{i1}, p_{i2}); i = 1 \ldots, n\}$ be a set of n points in a bounded domain $\Omega \subset \mathbb{R}^2$ with boundary $\partial\Omega \in C^2$. Let z_i be the value of a real-valued variable observed at location $\mathbf{p}_i$, and let $\mathbf{w}_i = (w_{i1}, \ldots, w_{iq})^t$ be a q-vector of covariates associated to observation z_i at $\mathbf{p}_i$.

Consider the semi-parametric generalized additive model

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

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 the function $f : \Omega \to \mathbb{R}$ captures the spatial structure of the phenomenon.

Generalizing the bivariate smoothing method introduced in [22, 24] the vector of regression coefficients $\boldsymbol{\beta}$ and the surface or spatial field f are estimated by minimizing the penalized sum-of-square-error functional

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

where λ is a positive smoothing parameter. Here Δf denotes the Laplacian of the function f. To define this partial differential operator let us consider a generic surface $f = f(\mathbf{p})$ defined on Ω. Denote the gradient of f by

$$\nabla f(\mathbf{p}) := \left(\frac{\partial f}{\partial p_1}(\mathbf{p}), \frac{\partial f}{\partial p_2}(\mathbf{p})\right)^t$$

where ^{t} is the transpose operator. Moreover, given a vector field $\mathbf{f} = \left(f_1(\mathbf{p}), f_2(\mathbf{p})\right)^t$ on Ω, where f_1 and f_2 are two surfaces on Ω, define the divergence of the vector field as

$$\operatorname{div} \mathbf{f}(\mathbf{p}) := \frac{\partial f_1}{\partial p_1}(\mathbf{p}) + \frac{\partial f_2}{\partial p_2}(\mathbf{p}).$$

Then, the Laplacian of the surface f is defined as

$$\Delta f(\mathbf{p}) := \operatorname{div} \nabla f(\mathbf{p}) = \frac{\partial^2 f}{\partial p_1^2}(\mathbf{p}) + \frac{\partial^2 f}{\partial p_2^2}(\mathbf{p}).$$

The Laplacian Δf provides a simple measure of the local curvature of the surface f defined on a planar domain Ω, and is invariant with respect to rigid transformations (rotations, translations and reflections) of the spatial coordinates of the domain. The use of the Laplace operator in the roughness penalty in (2) therefore ensures that the concept of smoothness does not depend on the orientation of the coordinate system. This roughness penalty can be seen as a generalization of the penalty considered for one-dimensional smoothing splines, normally consisting in the L^2-norm of the second order derivative of the curve to be estimated. Likewise for one-dimensional splines, the higher the smoothing parameter λ, the more we are controlling the roughness of the field f, the smaller the smoothing parameter, the more we are allowing for local curvature of f.

The proposed model is able to efficiently handle data distributed over domains Ω having shapes with complicated geometries. This constitutes an important advantage with respect to classical methods for surface estimation in all the applicative contexts where the shape of the problem domain is important for the behavior of the phenomenon under study. Classical methods for surface estimation, such as tensor product of unidimensional splines, thin-plate splines, bidimensional kernel smoothing, bidimensional wavelet-based smoothing and kriging, are in fact naturally defined on tensorized domains and cannot efficiently deal with more complex domains. Moreover, the method proposed can also comply with general conditions at the boundary $\partial\Omega$ of the domain [3, 24], which in many applied problems is a crucial feature to obtain meaningful estimates. These boundary conditions, homogeneous or not, may involve the evaluation of the function and/or its normal derivative at the boundary, allowing for different behaviors of the surface at the boundary of the domain of interest.

As will be summarized in Sect. 3, instead of the Laplacian and other simple partial differential operators, the roughness penalty may also involve more complex partial differential operators. This model extension, developed in [2–4], is particularly interesting whenever prior knowledge is available on the phenomenon under study, coming for instance from the Physics, Chemistry, Biology or Mechanics of the problem at hand, that can be formalized in terms of a partial differential equation modeling the phenomenon under study. Moreover, it allows for a very flexible modeling of the space variation.

Another important result is that the models can be extended to handle data distributed over general bidimensional Riemannian manifold domains. In such a case the differential operator considered in the roughness penalty is computed over the manifold domain. This generalization, developed in [8, 12], is reviewed in Sect. 4.

2.1 Modeling Data Distributed over Irregular Domains and Complying with General Conditions at the Domain Boundaries

To illustrate the issue of spatial smoothing over irregularly shaped domains and with boundary conditions, consider the problem of estimating population density over the Island of Montréal (QC, Canada), starting from census data (1996 Canadian census). Figure 1 displays census tract locations over the Island of Montréal; population density and other census information are available at each census tract, together with a binary covariate indicating whether a tract is predominantly residential or industrial/commercial. The figure highlights two parts of the island without data: the airport and rail yards in the south and an industrial park with an oil refinery tank farm in the north-east; these two areas are not part of the domain of interest when studying population density, since people cannot live there. Notice that 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 and other census quantities feature a smooth spatial variation over the inhabited parts of the island, there is no reason to assume smooth spatial variation across uninhabited areas. The figure also shows the island coasts as boundaries of the domain of interest, as it does not make sense to smooth population density into the rivers. Those parts of the boundary that are

Fig. 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. Figure from [24]

highlighted in red correspond respectively to the harbor, in the east shore, and to two public parks, in the south-west and north-west shore; no people live by the river banks in these boundary intervals. We thus want to 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.

2.2 Characterization of the Estimators

Denote by W the $n \times q$ matrix whose ith row is given by $\mathbf{w}_i^t$, the vector of q covariates associated with observation z_i at $\mathbf{p}_i$, and assume that W has full rank. Let P be the matrix that projects orthogonally on the subspace of $\mathbb{R}^n$ generated by the columns of W, i.e., $P := W(W^tW)^{-1}W^t$, and let $Q = I - P$, where 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., $\mathbf{f}_n := (f(\mathbf{p}_1), \ldots, f(\mathbf{p}_n))^t$.

Let $H^m(\Omega)$ be the Hilbert space of all functions which belong to $L^2(\Omega)$ along with all their distributional derivatives up to the order m. 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, imposing appropriate boundary conditions on f ensures that the estimation problem has a unique solution. Three classic boundary conditions are Dirichlet, Neumann and Robin conditions. The Dirichlet condition controls the value of f at the boundary, i.e., $f|_{\partial\Omega} = \gamma_D$, the Neumann condition controls instead the flow across the boundary, concerning the value of the normal derivative of f at the boundary, i.e., $\partial_{\boldsymbol{\nu}} f|_{\partial\Omega} = \gamma_N$, where $\boldsymbol{\nu}$ is the outward unit normal vector to $\partial\Omega$, while the Robin condition involves linear combinations of the above conditions, i.e., $\left(\partial_{\boldsymbol{\nu}} f + \alpha f\right)\big|_{\partial\Omega} = \gamma_R$. The functions γ_D, γ_N and γ_R have to satisfy some regularity conditions in order to obtain a well defined functional $J(f)$; when these functions coincide with null functions, the condition is said homogeneous. Moreover, it is possible to impose different boundary conditions on different portions of the boundary, forming a partition of $\partial\Omega$, as will be illustrated for instance in the application to Montréal census data. For simplicity of exposition, in the following we consider homogeneous Neumann (or homogeneous Dirichlet) boundary conditions, and $V(\Omega)$ will denote the subspace of $H^2(\Omega)$ characterized by the chosen boundary conditions. The interested reader is referred to [3, 4, 24] for the case of general boundary conditions.

The following Proposition characterizes the estimators (see [24]).

Proposition 1 *There exists a unique pair of estimators* $(\hat{\boldsymbol{\beta}} \in \mathbb{R}^q, \hat{f} \in V(\Omega))$ *which minimize (1). Moreover,*

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

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

for every $h \in V(\Omega)$.

Problem (3) is an infinite dimensional problem and cannot be solved analytically. We thus solve the problem numerically, reducing it to a finite dimensional one. Since (3) is a fourth order problem, a convenient way to tackle it is to first rewrite it as a coupled system of second order problems, by introducing an auxiliary variable. The latter problem is hence reformulated in order to involve only first order derivatives, i.e., in a suitable $H^1(\Omega)$ subspace. This weak or variational formulation can be obtained by integrating the differential equations against a test function and integrating by parts the second order terms. This problem reformulation is particularly well suited to be solved numerically, as H^1 spaces can be approximated by convenient finite dimensional spaces, and specifically by standard finite element spaces. Finite element analysis has been mainly developed and used in engineering applications, to solve partial differential equations. In the finite element space, solving the estimation problem reduces to solving a linear system. Next section describes the finite element spaces we shall be using.

2.3 Finite Element Solution to the Estimation Problem

To construct a finite element space, we start by partitioning the domain Ω of interest into small subdomains. Convenient domain partitions are given for instance by triangular meshes; Fig. 3, left panel, shows for example a triangulation of the domain of interest for the Island of Montréal data. In particular, we consider a regular triangulation $\mathscr{T}$ of Ω, where adjacent triangles share either a vertex or a complete edge. Domain Ω is hence approximated by domain $\Omega_{\mathscr{T}}$ 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 $\mathscr{T}$, with the associated finite element basis, is sufficient to adequately describe the data. The triangulation points in $\mathscr{T}$ may or may not coincide with the data locations $\mathbf{p}_i$. In any case, for the inferential properties of the estimators, it is convenient to consider triangulations that are finer where there are more data points, and coarser where there are fewer data points.

Starting from the triangulation $\mathscr{T}$, we can introduce a locally supported basis that spans the space of continuous surfaces on $\Omega_{\mathscr{T}}$, coinciding with polynomials of a given order over each triangle of $\mathscr{T}$. The resulting finite element space, denoted by $H^1_{\mathscr{T}}(\Omega)$, provides an approximation of the infinite dimensional space $H^1(\Omega)$. Linear finite elements are for instance obtained considering a basis system where each basis function ψ_j is associated with a triangle vertex ξ_j, $j = 1, \ldots, N$, in the triangulation $\mathscr{T}$. This basis function ψ_j is a piecewise linear polynomial which takes the value one at the vertex ξ_j and the value zero on all the other vertices of the mesh, i.e., $\psi_j(\xi_l) = \delta_{jl}$, where δ_{jl} denotes the Kronecker delta symbol. Figure 2 shows an example of such linear finite element basis function on a planar mesh, highlighting the locally supported nature of the basis.

Now, let $\boldsymbol{\psi} = (\psi_1, \ldots, \psi_N)^t$ be the column vector collecting the N basis functions associated with the N vertices ξ_j, $j = 1, \ldots, N$. Then, each function h in the finite element space $H^1_{\mathscr{T}}$ can be represented as an expansion in terms of the basis function $\psi_1, \ldots, \psi_N$. In particular,

$$h(\cdot) = \sum_{j=1}^{N} h(\xi_j)\psi_j(\cdot) = \mathbf{h}^t \boldsymbol{\psi}(\cdot) \qquad (4)$$

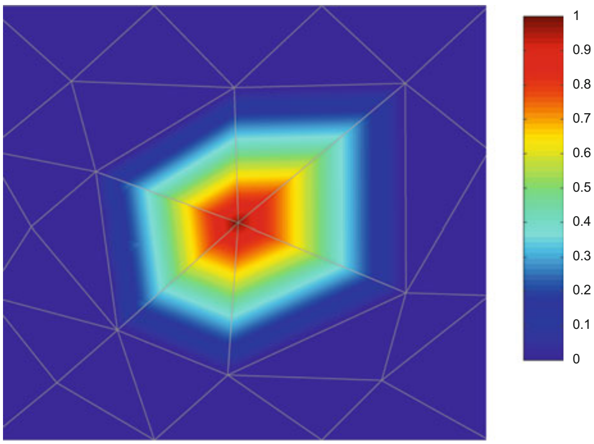

Fig. 2 Example of a linear finite element basis on a planar triangulation

where

$$\mathbf{h} = (h(\xi_1), \ldots, h(\xi_N))^t \tag{5}$$

is the column vector of the evaluations of h at the N mesh nodes. Each function $h \in H^1_{\mathscr{T}}$ is thus uniquely identified by its evaluations $\mathbf{h}$ on the mesh nodes.

Let Ψ be the $n \times N$ matrix of the evaluations of the N basis at the n data locations $\mathbf{p}_1, \cdots, \mathbf{p}_n$,

$$\Psi = \begin{bmatrix} \boldsymbol{\psi}^t(\mathbf{p}_1) \\ \vdots \\ \boldsymbol{\psi}^t(\mathbf{p}_n) \end{bmatrix}$$

and consider the $N \times N$ matrices

$$R_0 := \int_{\Omega_{\mathscr{T}}} (\boldsymbol{\psi}\boldsymbol{\psi}^t) \qquad R_1 := \int_{\Omega_{\mathscr{T}}} \nabla\boldsymbol{\psi}^t \nabla\boldsymbol{\psi}.$$

Corollary 1 shows that, once recast in the finite element space, solving the estimation problem reduces to solving a linear system.

Corollary 1 *There exists a unique pair of estimators* $(\hat{\boldsymbol{\beta}} \in \mathbb{R}^q, \hat{f} \in H^1_{\mathscr{T}}(\Omega))$ *which solve the discrete counterpart of the estimation problem. Moreover,*

- $\hat{\boldsymbol{\beta}} = (W^t W)^{-1} W^t (\mathbf{z} - \hat{\mathbf{f}}_n)$*;*
- $\hat{f} = \hat{\mathbf{f}}^t \boldsymbol{\psi}$*, with* $\hat{\mathbf{f}}$ *satisfying*

$$\begin{bmatrix} -\Psi^t Q \Psi & \lambda R_1 \\ \lambda R_1 & \lambda R_0 \end{bmatrix} \begin{bmatrix} \mathbf{f} \\ \mathbf{g} \end{bmatrix} = \begin{bmatrix} -\Psi^t Q\, \mathbf{z} \\ \mathbf{0} \end{bmatrix}. \tag{6}$$

Solving the linear system (6) is fast. In fact, although the system is typically large, being of order $2N$, it is highly sparse because the matrices R_0 and 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 R_0 and 0.2 % of the entries of R_1 were non-zero.

2.4 *Properties of the Estimators*

Corollary 1 highlights that the estimators $\hat{\boldsymbol{\beta}}$ and $\hat{f}$ are linear in the observed data values $\mathbf{z}$. Moreover $\hat{f}$ has a penalized regression form, being identified by the vector

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

where the positive definite matrix $R_1 R_0^- 1 R_1$ represents the discretization of the penalty term in (2). Notice that, thanks to the variational formulation of the estimation problem, this penalty matrix only involves the computation of first order derivatives.

Denote by S_f the $n \times n$ matrix

$$S_f = \Psi(\Psi^t Q \Psi + \lambda R_1 R_0^- 1 R_1)^{-1} \Psi^t Q.$$

Using this notation,

$$\hat{\mathbf{f}}_n = S_f\,\mathbf{z}$$

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

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

$$E[\hat{\mathbf{f}}_n] = S_f\,\mathbf{f}_n \tag{7}$$

$$Var(\hat{\mathbf{f}}_n) = \sigma^2\,S_f\,S_f^t \tag{8}$$

and

$$E[\hat{\boldsymbol{\beta}}] = \boldsymbol{\beta} + (W^t W)^{-1} W^t \big(I - S_f\big)\mathbf{f}_n \tag{9}$$

$$Var(\hat{\boldsymbol{\beta}}) = \sigma^2 (W^t W)^{-1} + \sigma^2 (W^t W)^{-1} W^t \big\{S_f\,S_f^t\big\} W(W^t W)^{-1}. \tag{10}$$

Denote now by S the smoothing matrix $S := P + Q\,S_f$ and consider the vector $\hat{\mathbf{z}}$ of fitted values at the n data locations

$$\hat{\mathbf{z}} = W\hat{\boldsymbol{\beta}} + \hat{\mathbf{f}}_n = S\,\mathbf{z}$$

The fitted values $\hat{\mathbf{z}}$ are thus obtained from observations $\mathbf{z}$ via application of the linear operator S, independent of $\mathbf{z}$. A commonly used measure of the equivalent degrees of freedom for linear estimators is given by the trace of the smoothing matrix; see, e.g., [6], who first introduced this notion. The equivalent degrees of freedom of the estimator $\hat{\mathbf{z}}$ are thus given by

$$tr(S) = q + tr(S_f),$$

and coincide with the sum of the q degrees of freedom of the parametric part of the model (q being the number of covariates considered) and of the $tr(S_f)$ equivalent degrees of freedom of the non-parametric part of the model. We can now estimate σ^2 by

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

This estimate, together with expressions (10) and (8), may be used to obtain approximate confidence intervals for $\boldsymbol{\beta}$ and approximate confidence bands for f. Furthermore, the value of the smoothing parameter λ may be selected by Generalized-Cross-Validation, see, e.g., [23] and references therein:

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

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.

The expressions (7) and (9) highlight that the estimators are biased. In particular, there are two sources of bias in the proposed model. The first source is the discretization and it is common to any model employing a basis expansion. This source of bias disappears as the number n of observations increases (in the sense of infill asymptotic) if meanwhile the mesh is correspondingly refined. The second source is the penalty term, and this is typical of regression models involving a roughness penalty: unless the true function f is such that it annihilates the penalty term, this term will of course induce a bias in the estimate. As shown in [3], this source of bias disappears as n increases, if the smoothing parameter λ decreases with n. This appears to be a natural request since having more observations decreases the need to impose a regularization. In all the simulations we carried out, the bias always appeared negligible.

2.5 Applied Illustrative Problem: Island of Montréal Census Data

Figure 3, left panel, displays a triangulation of the domain of interest for Montréal census data application. We shall estimate population density, measured as 1,000 inhabitants per km^2, using as covariate the binary variable that indicates whether a tract is predominantly residential (1) or commercial/industrial (0). We use here mixed boundary conditions: homogeneous Dirichlet along the stretches of coast corresponding to the harbor and the public parks, implying that the estimate of

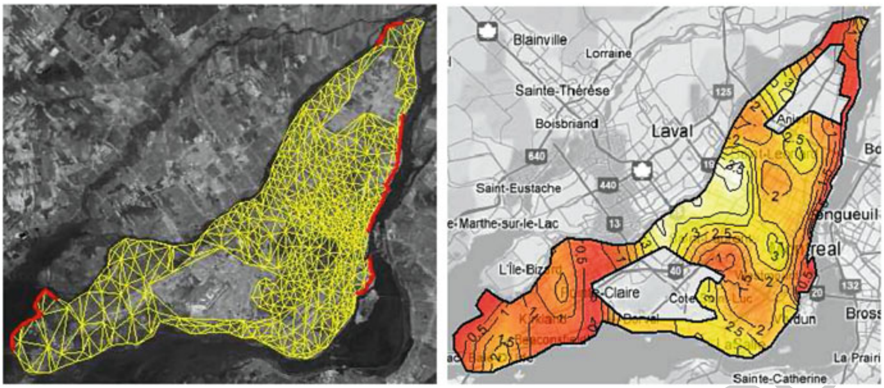

Fig. 3 *Left*: constrained Delaunay triangulation of the Island of Montréal. *Right*: estimate of spatial structure for population density over the Island of Montréal. Figures from [24]

population density should drop to zero along these stretches, and homogeneous Neumann along the remaining shores, meaning there is no immigration-emigration across shores. Figure 3, right panel, shows the estimated spatial structure of population density. The estimate complies with the specified boundary conditions and does not artificially link data points on either side of the uninhabited parts; see for instance the densely populated areas just in the south and in the west of the industrial park, with respect to the low population density neighborhood north-east of it. 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 1,300 more inhabitants per km^2, with respect to those predominantly commercial; the approximate 95 % confidence interval is given by [0.755; 1.845].

3 Incorporating Prior Knowledge About the Phenomenon Under Study

In this section we briefly summarize the generalization of the models developed in [2–4]. Suppose that prior knowledge is available on the phenomenon under study, that can be formalized through a partial differential equation $Lf = u$ modeling the phenomenon (where $u \in L^2(\Omega)$ is some forcing term). Partial differential models are indeed commonly used to describe complex phenomena behaviors in many fields of engineering and sciences. The prior knowledge, formalized in the PDE, is thus incorporated into the statistical model, looking for estimates of $\boldsymbol{\beta}$ and f that minimize the functional

$$J_\lambda(\boldsymbol{\beta}, f) = \sum_{i=1}^{n} \left(z_i - \mathbf{w}_i^t \boldsymbol{\beta} - f(\mathbf{p}_i)\right)^2 + \lambda \int_\Omega \left(Lf(\mathbf{p}) - u(\mathbf{p})\right)^2 d\,\mathbf{p} \qquad (11)$$

with respect to $f \in V$. The penalized error functional hence trades off a data fitting criterion, the sum-of-square-error, and a model fitting criterion, that penalizes departures from a PDE problem-specific description of the phenomenon. The proposed method can be seen as a regularized least square analogous to the Bayesian inverse problems presented, e.g., in [25]. In particular, the least square term in $J(f)$ corresponds to a log-likelihood for Gaussian errors, while the regularizing term effectively translates the prior knowledge on the surface. With respect to [25], besides the different model framework and estimation approaches, we also deal with a larger class of operators, including also non-stationary (i.e., spatially inhomogeneous) and anisotropic diffusion, transport and reaction coefficients. This also allows for a very flexible modeling of space variation, that accounts naturally for anisotropy and non-stationarity (space inhomogeneity).

In particular, in [3, 4] we consider phenomena that are well described in terms of linear second order elliptic operators L and forcing term $u \in L^2(\Omega)$ that can be either the null function $u = 0$, homogeneous case, or $u \neq 0$, non-homogeneous case. The operator L is a general differential operator that can, for instance, include second, first and zero order differential operators. Consider a symmetric and positive definite matrix $\mathbf{K} = \{\mathbf{K}_{ij}\} \in \mathbb{R}^{2\times 2}$, named diffusion tensor, a vector $\mathbf{b} = \{\mathbf{b}_j\} \in \mathbb{R}^2$, named transport vector, and a positive scalar $c \in \mathbb{R}^+$, named reaction term. Then, the operator can include: second order differential operators as the divergence of the gradient, i.e.,

$$\mathrm{div}(\mathbf{K}\nabla f) = \frac{\partial}{\partial x}\left(\mathbf{K}_{11}\frac{\partial f}{\partial x} + \mathbf{K}_{12}\frac{\partial f}{\partial y}\right) + \frac{\partial}{\partial y}\left(\mathbf{K}_{21}\frac{\partial f}{\partial x} + \mathbf{K}_{22}\frac{\partial f}{\partial y}\right),$$

first order differential operators as the gradient, i.e.,

$$\mathbf{b}\cdot\nabla f = \mathbf{b}_1\frac{\partial f}{\partial x} + \mathbf{b}_2\frac{\partial f}{\partial y} ,$$

and also zero order operators, i.e., cf. The general form that we consider is

$$Lf = -\mathrm{div}(\mathbf{K}\nabla f) + \mathbf{b}\cdot\nabla f + cf. \tag{12}$$

Moreover, the parameters of the differential operator L can be space-varying on Ω; i.e., $\mathbf{K} = \mathbf{K}(p_1, p_2)$, $\mathbf{b} = \mathbf{b}(p_1, p_2)$ and $c = c(p_1, p_2)$. The three terms that compose the general second order operator (12) induce an anisotropic and non-stationary smoothing, providing different regularizing effects. The diffusion term $-\mathrm{div}(\mathbf{K}\nabla f)$ induces anisotropic and non-stationary smoothing in all directions; the transport term $\mathbf{b}\cdot\nabla f$ induces a non-stationary smoothing only in the direction specified by the transport vector $\mathbf{b}$. Finally, the reaction term cf has instead a non-stationary shrinkage effect, since penalization of the L^2 norm of f induces a shrinkage of the surface to zero. Setting $\mathbf{K} = \mathbf{I}$, $\mathbf{b} = \mathbf{0}$, $c = 0$ and u equal to the null function we obtain the special case described in Sect. 2, where the penalization of the Laplacian Δf induces an isotropic and stationary smoothing.

In [2–4] the estimation problem is shown to be well defined. Its discretization with the finite element space follows the same lines described in Sect. 2.3, with the important difference that the matrix R_1 is now defined as

$$R_1 = \int_{\Omega_{\mathcal{T}}} \left(\nabla \boldsymbol{\psi}^t \mathbf{K} \nabla \boldsymbol{\psi} + \nabla \boldsymbol{\psi}^t \mathbf{b} \boldsymbol{\psi} + c\, \boldsymbol{\psi} \boldsymbol{\psi}^t \right).$$

This change is due to the different penalty in (11) with respect to (2). The matrix R_1 is in fact used in the discretization of the penalty term. In the case where the considered forcing term $u \in L^2(\Omega)$ is not homogeneous ($u \neq 0$), the vector $\mathbf{0}$ in the right hand side of (6) is replaced by the discretization $\mathbf{u} = (u(\xi_1), \ldots, u(\xi_N))^t$ of the forcing term. Moreover, when the forcing term is homogeneous the estimator properties are as in Sect. 2.4; otherwise, additional terms need instead to be considered, but the estimators remain linear in the observed data values, and their properties follows along the lines described in Sect. 2.4.

In [3] we study in detail the numerical convergence properties of the Finite Element approximation to the estimation problem. In addition, the case of areal data is considered in [3, 4].

3.1 Applied Illustrative Problem: Blood-Flow Velocity Field Estimation

The motivating applied problem driving the generalization to more complex penalty terms concerns the estimation of the blood-flow velocity field on a cross-section of the common carotid artery, using data provided by echo-color doppler acquisitions. This applied problem arises within the research project MAthematichs for CARotid ENdarterectomy@MOX (MACAREN@MOX), which aims at studying atherosclerosis pathogenesis. Carotid Echo-Color Doppler (ECD) is a medical imaging procedure that uses reflected ultrasound waves to create images of an artery and to measure the velocity of blood cells in some locations within the artery. Specifically, the ECD data measure the velocity of blood within beams located on the considered artery cross-section; see Fig. 4, top left panel. For this study, during the ECD scan 7 beams are considered, located in the cross-shaped pattern shown in Fig. 4, bottom left panel.

In this applied problem we have prior knowledge on the phenomenon under study that could be exploited to derive accurate physiological estimates. There is in fact a vast literature devoted to the study of fluid dynamics and hemodynamics, see for example [13] and references therein. This prior information concerns both the shape of the field, which can be conveniently described via a linear second order elliptic PDE, and the conditions at the boundary of the problem domain, i.e., specifically, at the wall of the carotid cross-section. The proposed method efficiently uses the prior information on the phenomenon under study and gives a realistic and physiological estimate of the dynamic blood flow, which is not affected by the pattern of the

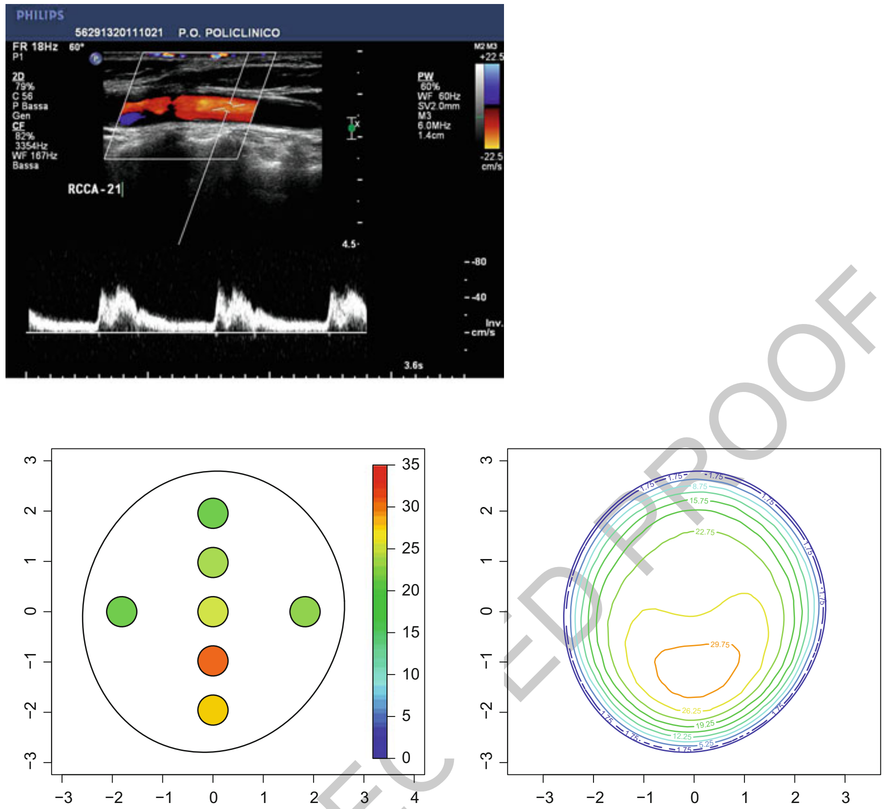

Fig. 4 *Top left*: ECD image corresponding to the central point of the carotid section located 2 cm before the carotid bifurcation. *Bottom left*: MRI reconstruction of the cross-section of the carotid artery located 2 cm before the bifurcation; cross-shaped pattern of observations with each beam colored according to the mean blood-velocity measured on the beam at systolic peak time. *Bottom right*: estimate of the blood-flow velocity field in the carotid section. Figures from [4]

observations; see Fig. 4, bottom right panel. Moreover the estimates accurately highlight important features of the blood flow, such as eccentricity, asymmetry and reversion of the fluxes, that are of interest to physicians, in order to understand how the local hemodynamics influences atherosclerosis pathogenesis. See [4].

4 Modeling Data over Manifold Domains

In [12], the models are extended to non-planar domains. Few methods are available in literature to deal with data on non-planar domains (see, e.g., [5, 7, 14, 16–18, 26]); to the best of our knowledge, none of these methods is currently devised to handle the data structures we are considering, with variables of interest together with space-varying covariates, both distributed over general bi-dimensional Riemannian manifolds.

Consider n fixed data locations $\{\mathbf{x}_i = (x_{1i}, x_{2i}, x_{3i}) : i = 1, \ldots, n\}$ lying on a general bi-dimensional Riemannian manifold Γ, and let the variable of interest z_i and covariates $\mathbf{w}_i$ be observed at location $\mathbf{x}_i$, for $i = 1, \ldots, n$. Likewise in (1), assume the model

$$z_i = \mathbf{w}_i^t \boldsymbol{\beta} + f(\mathbf{x}_i) + \epsilon_i, \qquad i = 1, \ldots, n \tag{13}$$

where the spatial field f is now defined on the manifold Γ, $f : \Gamma \to \mathbb{R}$. In analogy to (2), the vector of regression coefficients $\boldsymbol{\beta}$ and the surface or spatial field f are estimated by minimizing the penalized sum-of-square-error functional

$$J_{\Gamma,\lambda}(\boldsymbol{\beta}, f) = \sum_{i=1}^{n} \left(z_i - \mathbf{w}_i^t \boldsymbol{\beta} - f(\mathbf{x}_i)\right)^2 + \lambda \int_{\Gamma} (\Delta_\Gamma f(\mathbf{x}))^2 \, d\mathbf{x} \tag{14}$$

where $\Delta_\Gamma f$ denotes the Laplace-Beltrami operator, a generalization of the standard Laplacian to functions $f = f(x_1, x_2, x_3)$, defined on a non-planar domain Γ, with $(x_1, x_2, x_3) \in \Gamma$. The definition of the Laplace-Beltrami operator requires the computation of the gradient operator ∇_Γ and of the divergence operator div_Γ over the non planar domain; see, e.g., [10]. The Laplace-Beltrami operator of f is then defined as

$$\Delta_\Gamma f(\mathbf{x}) = \mathrm{div}_\Gamma \, \nabla_\Gamma f(\mathbf{x}).$$

Similar to the standard Laplacian, the Laplace-Beltrami operator provides a simple measure of the local curvature of the function f, as defined on the curved domain Γ. Likewise to the standard Laplacian, the Laplace-Beltrami is invariant with respect to rigid transformations (rotations, translations and reflections) of the spatial coordinates of the non-planar domain. Hence, the employment of the Laplace-Beltrami operator as a roughness penalty ensures that the concept of smoothness does not depend on the orientation of the coordinate system or on the orientation of the domain Γ itself.

In [12] the estimation problem (14) is recast over a planar domain, via a conformal reparametrization of the non-planar domain Γ. The reparametrization is obtained by a continuously differentiable map

$$\begin{aligned} X &: \Omega \to \Gamma \\ \mathbf{p} &= (p_1, p_2) \mapsto \mathbf{x} = (x_1, x_2, x_3) \end{aligned} \tag{15}$$

where Ω is an open, convex and bounded set in $\mathbb{R}^2$ and the boundary of Ω, denoted $\partial\Omega$, is piecewise C^∞. The map X essentially provides a change of variable between the planar coordinates $\mathbf{p} = (p_1, p_2)$ and the non-planar coordinates $\mathbf{x} = (x_1, x_2, x_3)$, and is unique up to dilations, rotations and translations. Consider the first order partial derivatives of X with respect to the planar coordinates p_1 and p_2, $\frac{\partial X}{\partial p_1}(\mathbf{p})$ and $\frac{\partial X}{\partial p_2}(\mathbf{p})$, that are column vectors in $\mathbb{R}^3$. We denote by $\langle \cdot, \cdot \rangle$ the Euclidean scalar product of two vectors and by $\| \cdot \|$ the corresponding norm. Consider the (space-dependent) metric tensor

$$G(\mathbf{p}) := \nabla X(\mathbf{p})' \nabla X(\mathbf{p}) = \begin{pmatrix} \left\| \frac{\partial X}{\partial p_1}(\mathbf{p}) \right\|^2 & \langle \frac{\partial X}{\partial p_1}(\mathbf{p}), \frac{\partial X}{\partial p_2}(\mathbf{p}) \rangle \\ \langle \frac{\partial X}{\partial p_1}(\mathbf{p}), \frac{\partial X}{\partial p_2}(\mathbf{p}) \rangle & \left\| \frac{\partial X}{\partial p_2}(\mathbf{p}) \right\|^2 \end{pmatrix}.$$

Let $\mathscr{W}(\mathbf{p}) := \sqrt{det(G(\mathbf{p}))}$, and define the matrix $\mathbf{K}(\mathbf{p}) = \mathscr{W}(\mathbf{p})\, G^{-1}(\mathbf{p})$, where $G^{-1}(\mathbf{p})$ denotes the inverse of $G(\mathbf{p})$. Then for $f \circ X \in \mathscr{C}^2(\Omega)$, the Laplace-Beltrami operator be re-expressed in terms of the planar coordinates $\mathbf{p}$ as

$$\Delta_\Gamma f(\mathbf{x}) = \frac{1}{\mathscr{W}(\mathbf{p})} \mathrm{div}(\mathbf{K}(\mathbf{p}) \nabla f(X(\mathbf{p}))) \tag{16}$$

where div and ∇ denote the standard divergence and gradient operators over planar domains, defined in Sect. 2. Thus, by considering the map X and the corresponding planar parametrization (16) for the Laplace-Beltrami operator, after setting $\mathbf{u}_i = X^{-1}(\mathbf{x}_i)$, we can reformulate the estimation problem (14) over the manifold Γ as an equivalent problem over the planar domain Ω as follows: find $\boldsymbol{\beta}$ and the function $f \circ X$, defined on Ω, that minimizes

$$\begin{aligned} J_{\Omega,\lambda}(f \circ X) = &\sum_{i=1}^{n} \left(z_i - \mathbf{w}_i' \boldsymbol{\beta} - f(X(\mathbf{p}_i)) \right)^2 \\ &+ \lambda \int_\Omega \frac{1}{\mathscr{W}(\mathbf{p})} \left(\mathrm{div}\big(\mathbf{K}(\mathbf{p}) \nabla f(X(\mathbf{p}))\big) \right)^2 d\Omega. \end{aligned} \tag{17}$$

As previously remarked, the map X and its corresponding planar domain Ω are unique only up to dilations, rotations and translations. However, the terms $\mathbf{K}(\mathbf{u})$ and $\mathscr{W}(\mathbf{u})$, that account for the change of variable from the original non-planar coordinates to the planar ones, adjust for each considered map X and for the corresponding planar domain Ω. Consequently this leads to different planar parameterizations of the function f as well as of the estimation problem (17), all equivalent to problem (14) on the original manifold Γ.

The conformal reparametrization is computed resorting to non-planar finite elements. These are defined likewise planar finite elements, but over non-planar meshes. See [12] for details. The discretization of the conformal map via non-planar finite elements provides a planar mesh for the corresponding planar domain Ω, and the discretization of the terms $\mathscr{W}$ and $\mathbf{K}$ in (17). These terms, together with the planar mesh, are then used for the solution of the equivalent estimation problem (17) over the planar domain. The discretization of problem (17) with planar finite elements follows the same lines described in Sect. 2.3, with the important difference that the matrices R_0 and R_1 are now defined as

$$R_0 := \int_{\Omega_{\mathscr{T}}} \mathscr{W}(\boldsymbol{\psi}\boldsymbol{\psi}^t) \qquad R_1 := \int_{\Omega_{\mathscr{T}}} \nabla\boldsymbol{\psi}^t \mathbf{K} \nabla\boldsymbol{\psi}.$$

This change is due to the different penalty in (17) with respect to (2), with the terms $\mathscr{W}(\mathbf{p})$ and $\mathbf{K}(\mathbf{p})$ accounting for the change of variable from the original non-planar coordinates to the planar ones. The estimator properties are as in Sect. 2.4.

4.1 Applied Illustrative Problems: Modeling Hemodynamical Stresses on Cerebral Arteries and Studying Cortical Surface Data

The extension to manifold domains has fascinating fields of applications. In [12] the models are applied to the study of hemodynamical forces exerted by blood flow on the wall of cerebral arteries affected by aneurysms; see Fig. 5. These data come from three-dimensional angiographies and computational fluid dynamics simulations, and belong to the AneuRisk project, a scientific endeavor 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; see http://mox.polimi.it/it/progetti/aneurisk/. In [8] we discuss instead an application in the neurosciences by analyzing cerebral cortex thickness data; see Fig. 6.

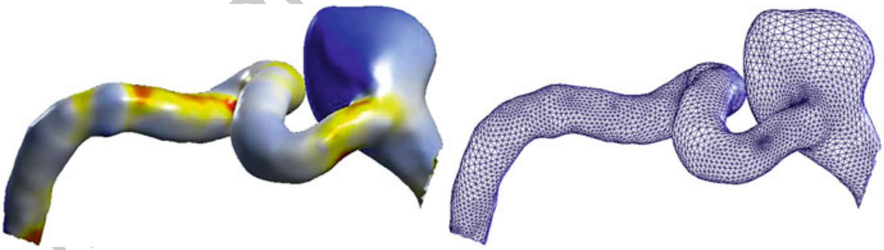

Fig. 5 *Left*: estimate of shear stress (modulus of shear stress at systolic peak) exerted by blood flow on the wall of an internal carotid artery affected by an aneurysm. *Right*: triangular mesh reconstruction of the wall of the internal carotid artery in *left panel*. From [12]

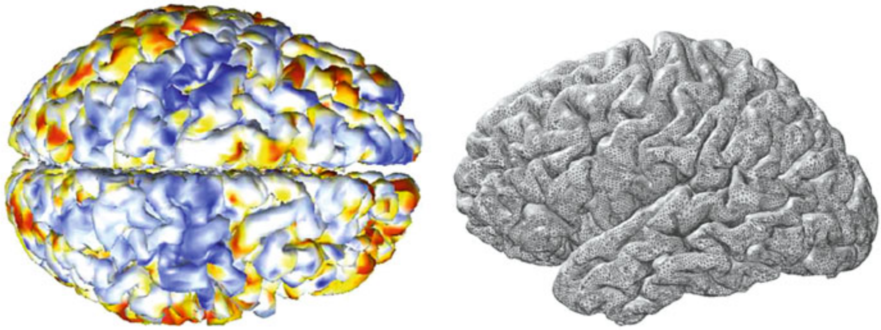

Fig. 6 *Left*: estimate of cortical thickness. *Right*: triangular mesh reconstruction of cortical surface

5 Discussion

We are currently extending this approach in various directions. One important generalization concerns for instance the modeling of space-time data, both over planar and over non-planar domains, which is of interest in several applications, including those briefly mentioned in Sects. 2.1, 3.1 and 4.1. Moreover, via a generalized linear framework, we are extending the models in order to handle outcomes having general distributions within the exponential family, including binomial, Poisson and gamma outcomes, further broadening the applicability of the proposed models. Combining all these features may in fact create a class of models that aims at handling data structures for which no statistical modeling currently exists.

Acknowledgements This paper reviews joint works with Laura Azzimonti, Bree Ettinger, Fabio Nobile, Simona Perotto, Jim Ramsay and Piercesare Secchi. This research has been funded by the research program Dote Ricercatore Politecnico di Milano—Regione Lombardia, project "Functional data analysis for life sciences", and by the starting grant *FIRB Futuro in Ricerca*, MIUR Ministero dell'Istruzione dell'Università e della Ricerca, research project "Advanced statistical and numerical methods for the analysis of high dimensional functional data in life sciences and engineering" (http://mox.polimi.it/users/sangalli/firbSNAPLE.html).

References

1. Augustin, N.H., Trenkel, V.M., Wood, S.N., Lorance, P.: Space-time modelling of blue ling for fisheries stock management issue. Environmetrics **24**(Part 2), 109–119 (2013)
2. Azzimonti, L.: Blood flow velocity field estimation via spatial regression with PDE penalization. Ph.D. Thesis, Politecnico di Milano (2013)
3. Azzimonti, L., Nobile, F., Sangalli, L.M., Secchi, P.: Mixed finite elements for spatial regression with PDE penalization. SIAM/ASA J. Uncertain. Quantif. (2013)

AQ1

4. Azzimonti, L., Sangalli, L.M., Secchi, P., Domanin, M., Nobile, F.: Blood flow velocity field estimation via spatial regression with PDE penalization. Technical Report MOX 19/2013, Dipartimento di Matematica, Politecnico di Milano (2013)
5. Baramidze, V., Lai, M.-J., Shum, C.K.: Spherical splines for data interpolation and fitting. SIAM J. Sci. Comput. **28**, 241–259 (2006)
6. Buja, A., Hastie, T., Tibshirani, R.: Linear smoothers and additive models. Ann. Stat. **17**, 453–555 (1989)
7. Chung, M.K., Robbins, S.M., Dalton, K.M., Davidson, R.J., Alexander, A.L., Evans, A.C.: Cortical thickness analysis in autism with heat kernel smoothing. NeuroImage **25**, 1256–1265 (2005)
8. Dassi, F., Ettinger, B., Perotto, S., Sangalli, L.M.: A mesh simplification strategy for a spatial regression analysis over the cortical surface of the brain. Technical Report MOX 31/2013, Dipartimento di Matematica, Politecnico di Milano (2013)
9. D'Elia, M., Perego, M., Veneziani, A.: A variational data assimilation procedure for the incompressible Navier-Stokes equations in hemodynamics. SIAM J. Sci. Comput. **52**(Part 4), 340–359 (2012)
10. Dierkes, U., Hildebrandt, S., Sauvigny, F.: Minimal Surfaces, vol. 1, 2nd edn. Springer, Heidelberg (2010)
11. Ettinger, B., Guillas, S., Lai, M.-J.: Bivariate splines for ozone concentration forecasting. Environmetrics **23**(Part 4), 317–328 (2012)
12. Ettinger, B., Perotto, S., Sangalli, L.M.: Spatial regression models over two-dimensional manifolds. Technical Report MOX 54/2012, Dipartimento di Matematica, Politecnico di Milano (2012)
13. Formaggia, L., Quarteroni, A., Veneziani, A.: Cardiovascular mathematics: modeling and simulation of the circulatory system. Springer, Berlin (2009)
14. Gneiting, T.: Strictly and non-strictly positive definite functions on spheres. Bernoulli **19**, 1087–1500 (2013)
15. Guillas, S., Lai, M.: Bivariate splines for spatial functional regression models. J. Nonparametr. Stat. **22**(Part 4), 477–497 (2010)
16. Hagler, D.J., Saygin, A.P., Sereno, M.I.: Smoothing and cluster thresholding for cortical surface-based group analysis of fMRI data. NeuroImage **33**, 1093–1103 (2006)
17. Jun, M.: Non-stationary cross-covariance models for multivariate processes on a globe. Scand. J. Stat. **38**, 726–747 (2011)
18. Lindgren, F., Rue, H., Lindström, J.: 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 (2011)
19. Marra, G., Miller, D., Zanin, L.: Modelling the spatiotemporal distribution of the incidence of resident foreign population. Stat. Neerl. **66**, 133–160 (2012)
AQ2
20. Quarteroni, A.: Numerical Models for Differential Problems. Springer, Berlin (2013).
21. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna (2011)
22. Ramsay, T.: Spline smoothing over difficult regions. J. R. Stat. Soc. Ser. B Stat. Methodol. **64**, 307–319 (2002)
23. Ramsay, J.O., Silverman, B.W.: Functional Data Analysis, 2nd edn. Springer, Berlin (2005)
24. Sangalli, L.M., Ramsay, J.O., Ramsay, T.O.: Spatial spline regression models. J. R. Stat. Soc. Ser. B Stat. Methodol. **75**(Part 4), 681–703 (2013)
25. Stuart, A.: Inverse problems: a Bayesian perspective. Acta Numer. **19**, 451–559 (2010)
26. Wahba, G.: Spline interpolation and smoothing on the sphere. SIAM J. Sci. Stat. Comput. **2**, 5–16 (1981)
27. Wood, S.N., Bravington, M.V., Hedley, S.L.: Soap film smoothing. J. R. Stat. Soc. Ser. B Stat. Methodol. **70**, 931–955 (2008)

AUTHOR QUERIES

AQ1. Please provide volume number and page range for [3].
AQ2. Ref. [20] is not cited in the text. Please provide the citation or delete it from the list.