MOX-Report No. 65/2026

# Modeling group heterogeneity in spatio-temporal data via physics-informed regression

De Sanctis, M.F.; Arnone, E.; Ieva, F.; Sangalli, L.M.

MOX, Dipartimento di Matematica
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

mox-dmat@polimi.it https://mox.polimi.it

# Modeling group heterogeneity in spatio-temporal data via physics-informed regression

Marco F. De Sanctis[a], Eleonora Arnone[b],
Francesca Ieva[a], Laura M. Sangalli[a,*]

[a] *MOX, Dipartimento di Matematica, Politecnico di Milano, Piazza Leonardo Da Vinci 32, Milano, 20133, Italy*
[b] *Dipartimento di Management, Università degli Studi di Torino, Corso Unione Sovietica, 218 bis, Torino, 10134, Italy*

**Abstract**

We propose a physics-informed semiparametric framework for modeling spatio-temporal data with group structure. The approach extends classical mixed-effects regression by incorporating a nonparametric space–time component, regularized through a partial differential equation to encode the underlying physical dynamics, while random effects capture group-specific variability. Estimation is carried out via a two-step procedure based on a functional extension of the Iteratively Reweighted Least Squares algorithm. We establish asymptotic properties of both fixed and random effect estimators and we assess the performance of the method through simulation studies against state-of-the-art alternatives. The proposed framework is applied to hourly nitrogen dioxide data over Lombardy (Italy), where random effects account for measurement heterogeneity across monitoring stations with different sensor technologies, demonstrating its effectiveness in capturing both physical dynamics and group heterogeneity.

*Keywords:* mixed-effect spatial regression, smoothing with differential regularization, air quality assessment.

[*] Corresponding author: laura.sangalli@polimi.it

Published in Spatial Statistics, 2026, DOI: 10.1016/j.spasta.2026.101022. Publisher version available at: https://doi.org/10.1016/j.spasta.2026.101022

## 1. Introduction

In this work, we address the modeling of spatio-temporal data exhibiting a group structure, which may arise, for instance, from the use of different measurement technologies. Our focus is on phenomena characterized by complex spatio-temporal patterns driven by external forces, a common feature in many real-world applications, particularly in environmental sciences. An illustrative example is shown in Figure 1, which reports hourly measurements of nitrogen dioxide ($NO_2$) collected on 15 January 2019 by the Agenzia Regionale per la Protezione dell'Ambiente (ARPA) monitoring network in the Lombardy region (Italy). These observations exhibit pronounced temporal variability and sharp spatial gradients, strongly influenced by air circulation. Modeling such processes entails several challenges. First, monitoring sensors differ in technology, design, and calibration systems (see the bottom-right panel of Figure 1). These differences induce a natural grouping structure, requiring statistical models capable of disentangling group-specific variability from the underlying signal of interest. A second challenge arises from the influence of physical mechanisms, such as wind dynamics and diffusion processes, on pollutant concentrations. Figure 2 illustrates the wind field observed on the same day as the $NO_2$ measurements. Jointly accounting for group heterogeneity and physically driven dynamics is essential for an accurate characterization of the phenomenon.

In classical linear regression settings, where no spatial or temporal structure is present, group-specific effects are typically modeled through random components, leading to linear mixed-effects models (see, e.g., Pinheiro and Bates, 2000; Gałecki and Burzykowski, 2012). A substantial body of work has extended this framework to spatio-temporal regression. In this context, however, random effects are most often introduced to capture residual spatial dependence, rather than to explicitly represent grouping structures within the data. Applications of such spatial random effect models include ordinal data (Mullen and Birkeland, 2008), compositional data (Di Brisco and Migliorati, 2021), and environmental data (Smith et al., 2003), with several contributions also addressing spatial confounding (e.g., Khan and Calder, 2022). More general formulations further extend generalized linear mixed models to additive or nonparametric settings (Lin and Zhang, 1999; Karcher and Wang, 2001). In contrast, a smaller body of literature incorporates random effects in spatio-temporal models specifically to represent group structures, which is also the focus of this work. For instance, Wood

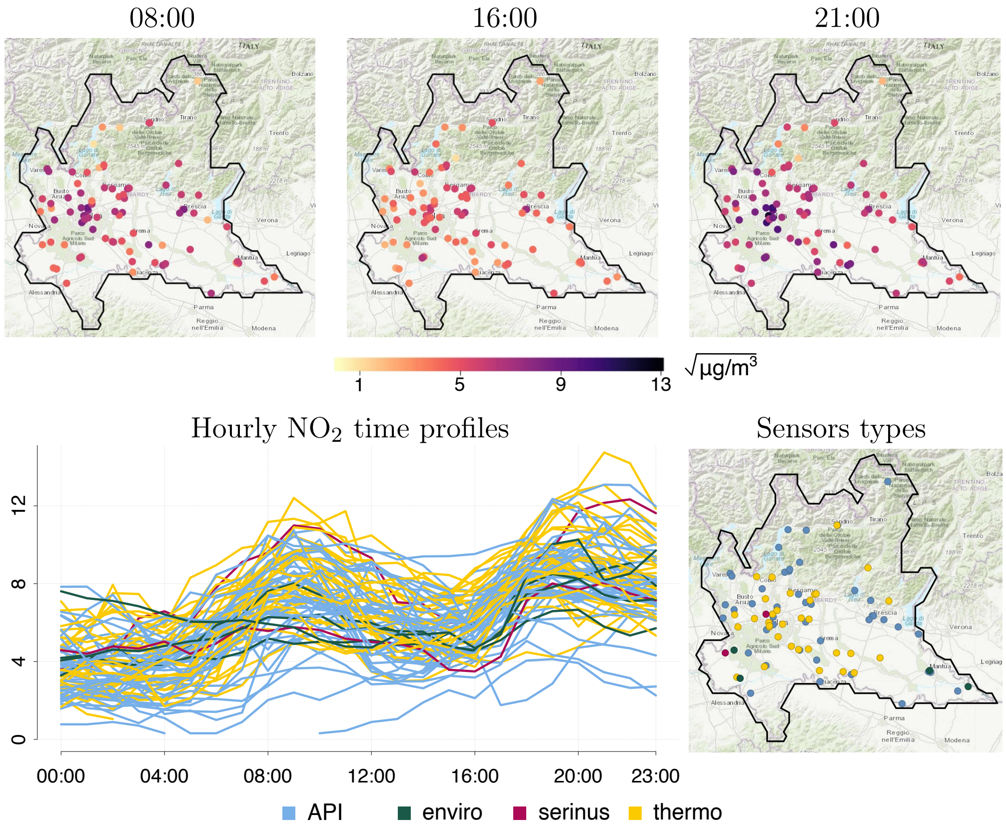


Figure 1: Top panel: spatial distribution of square-root $NO_2$ concentrations in Lombardy at three representative hours of the day (08:00, 16:00, and 21:00). Bottom panel: hourly temporal profiles of square-root $NO_2$ concentrations across ARPA monitoring stations on 15 January 2019 (left); spatial distribution of sensor technology types across the region (right).

(2006) introduced random effects to capture grouping factors while modeling spatial and temporal dependence through low-rank smooths. Yanosky et al. (2014) applied this framework to particulate matter data, including site-specific random effects to account for unobserved heterogeneity across monitoring stations, while modeling large-scale trends via smooth functions of geographic and meteorological covariates. Similarly, Sahu et al. (2006) modeled fine particulate matter concentrations using a combination of fixed effects and spatio-temporal random components distinguishing between rural and urban areas. More recently, da Matta et al. (2025) proposed a Bayesian spatio-temporal functional model in which random effects represent regional climate regimes. These contributions highlight the role of random components in flexibly capturing additional sources of variability associated with grouping structures.

A further challenge arises from the influence of complex physical mechanisms, such as wind currents and diffusion processes, on pollutant concentrations. In recent years, the integration of physical information into statistical models has received increasing attention. Among these approaches, a prominent line of research incorporates such information through Partial Differential Equations (PDEs). Early contributions in this direction include the PDE-based regularization framework for spatial regression proposed by Azzimonti et al. (2014, 2015) and further developed by Tomasetto et al. (2024), with extensions to the spatio-temporal setting introduced by Arnone et al. (2019). More recently, stochastic PDE approaches have been extended to incorporate physically interpretable features such as anisotropy and transport term (Clarotto et al., 2024; Carrizo Vergara et al., 2022), building on the seminal work of Lindgren et al. (2011) and the literature reviewed in Lindgren et al. (2022). Additional developments include the modeling of non-stationary Gaussian random fields on compact Riemannian manifolds (Pereira et al., 2022) and the construction of physics-informed covariance functions, such as those based on the exponential Boltzmann–Gibbs representation (Allard et al., 2021). Alternative approaches integrating PDEs in time-varying settings have also been proposed, including Wikle and Hooten (2010), Richardson (2017), and Hefley et al. (2017).

Despite these advances, existing approaches do not explicitly integrate physics-informed modeling with group-structured data. Building on this perspective, we propose a physics-informed semiparametric mixed-effects regression model that integrates data-driven modeling with the physical dynamics of the underlying process. The model includes fixed and random parametric components to account for covariate effects and group heterogeneity, together with a nonparametric component describing the nonlinear spatio-temporal evolution of the phenomenon. The estimation of this component is guided by a physics-informed regularization term based on a Partial Differential Equation (PDE) that, in the application to $NO_2$ data, encodes pollutant dispersion through an advection term driven by the wind field (see Figure 2). This mixed-effects formulation provides two main advantages. First, it enables the modeling of group-specific variability, distinguishing among sensor technologies within the monitoring network. Second, it allows for the separation of measurement noise, arising from instrumentation differences, from the underlying spatio-temporal signal of interest. From a methodological standpoint, the inclusion of random effects in the objective functional leads to a non-quadratic optimization problem, as the covariance matrix of the random components enters nonlinearly. As a result, iterative procedures are required to solve the estimation problem.

An additional strength of the proposed framework lies in its ability to handle missing data, a common feature of spatio-temporal environmental studies based on sensor networks. Measurements from monitoring stations are often incomplete due to temporary malfunctions, equipment failures, or maintenance interruptions. To address this issue, we formulate the model within the statistical framework of Arnone et al. (2023b), which ensures stable estimation in the presence of missing data. Moreover, the proposed approach can be applied to data defined over spatial domains with complex geometries. For the spatial discretization of the nonparametric component, we employ a finite element basis, which is well suited to represent phenomena evolving over irregular or non-convex regions, including domains with natural barriers or curved surfaces. This flexibility is particularly important when the physical dynamics of the process are influenced by the geometry of the domain, as in applications involving water bodies with irregular coastlines or biological signals observed on complex three-dimensional structures (see, e.g., Sangalli, 2021; Tomasetto et al., 2024; Castiglione et al., 2025).

We investigate the theoretical and empirical properties of the proposed framework. In particular, we derive asymptotic results for both fixed and random effect estimators. The empirical performance of the method is assessed through simulation studies, where it is compared with state-of-the-art alternatives, demonstrating its competitive advantages. Finally, the practical potential of the approach is illustrated through an application to hourly $NO_2$ measurements over Lombardy, showing its ability to capture pollutant dynamics while accounting for sensor heterogeneity.

The remainder of the paper is organized as follows. Section 2 introduces the proposed physics-informed mixed-effects model for spatio-temporal data. Section 3 presents the estimation strategy for both parametric and nonparametric components. The asymptotic properties of the fixed and random effect estimators are discussed in Section 4. Section 5 evaluates the performance of the proposed method through simulation studies against state-of-the-art alternatives. The methodology is applied to air quality data in Section 6, illustrating its ability to handle sensor heterogeneity in $NO_2$ assessment over Lombardy. Finally, Section 7 summarizes the main contributions and out-lines directions for future research.

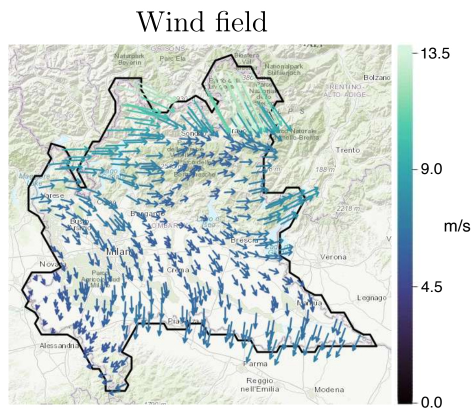


Figure 2: Daily average wind vector field over Lombardy on 15 January 2019. Wind intensity and direction are derived from data collected at 119 ARPA monitoring stations.

## 2. Physics-informed mixed-effect model for space-time data

Let $\{\mathbf{p}_i\}_{i=1}^n$ be a set of $n$ spatial locations on a bounded domain $\mathscr{D} \subset \mathbb{R}^2$, and let $\{t_j\}_{j=1}^m$ be a set of $m$ time instants in the interval $[0, T] \subset \mathbb{R}$. At these spatio-temporal locations, we partially observe $y_{ij}$, for $i = 1, \ldots, n,\ j = 1, \ldots, m$. Assume that these observations are divided into $g$ non-overlapping groups, indexed by $k = 1, \ldots, g$. For each group $k$, we define the set of indices corresponding to the observed spatio-temporal locations as

$$\mathcal{O}_k = \{(i, j) \in \{(1,1), \ldots, (n, m)\} : \text{the observation } y_{ij} \text{ belongs to group } k\}.$$

We denote by $|\mathcal{O}_k|$ its cardinality, and set $\mathcal{O} = \bigcup_{k=1}^g \mathcal{O}_k$. For each pair $(i, j) \in \mathcal{O}$, we further observe $\mathbf{x}_{ij} \in \mathbb{R}^q$ and $\mathbf{z}_{ij} \in \mathbb{R}^p$, representing space-time varying fixed effect and random effect design vectors, respectively. Note that the definition of $\mathcal{O}$ naturally accounts for missing data, with $|\mathcal{O}| \leq nm$.

For each group $k = 1, \ldots, g$, we describe the observations through the following semiparametric mixed-effects model

$$\begin{aligned} \mathbf{y}_k &= X_k \boldsymbol{\beta} + \boldsymbol{f}_k + Z_k \mathbf{b}_k + \boldsymbol{\varepsilon}_k, \\ \mathbf{b}_k &\overset{iid}{\sim} \mathcal{N}(\mathbf{0}, \Sigma_{\mathbf{b}}), \quad \boldsymbol{\varepsilon}_k \sim \mathcal{N}(\mathbf{0}, \sigma^2 I), \end{aligned} \tag{1}$$

where $\mathbf{y}_k$ is the vector collecting $\{y_{ij}\}_{(i,j)\in\mathcal{O}_k}$; $X_k \in \mathbb{R}^{|\mathcal{O}_k| \times q}$ stores, by rows, the covariates $\{\mathbf{x}_{ij}\}_{(i,j)\in\mathcal{O}_k}$, which affect the response $\mathbf{y}_k$ through the fixed effects $\boldsymbol{\beta} \in \mathbb{R}^q$, common to all observations; $Z_k \in \mathbb{R}^{|\mathcal{O}_k| \times p}$ collects the group-specific random effect design vectors $\{\mathbf{z}_{ij}\}_{(i,j)\in\mathcal{O}_k}$ associated with the random effects $\mathbf{b}_k \in \mathbb{R}^p$; $\Sigma_{\mathbf{b}}$ represents the unknown covariance matrix of the random effects $\mathbf{b}_k$; $\boldsymbol{f}_k$ is the vector of evaluations $f(\mathbf{p}_i, t_j)$, for $(i,j) \in \mathcal{O}_k$, representing a smooth function common to all groups and defined over the spatio-temporal domain $\mathscr{D} \times [0, T]$; finally, $\boldsymbol{\varepsilon}_k$ is the homoscedastic Gaussian within-group error for the $k$-th group, with variance $\sigma^2 > 0$. We assume that the fixed effect covariate matrix $X = \left[X_1, \cdots, X_g\right]^\top$ is full rank, and that the constant vector $\mathbf{1}$ does not belong to span$(X)$, to ensure ensure that the model is identifiable. Moreover, we assume that the random effects $\{\mathbf{b}_k\}_{k=1}^g$ are independent of the error terms $\{\boldsymbol{\varepsilon}_k\}_{k=1}^g$, as standard in classical mixed-effects models. The validity of these assumptions can be assessed using well-established diagnostic tools (see, e.g., Pinheiro and Bates (2000)). The model in equation (1) extends the one-level linear mixed-effects formulation in, e.g., Pinheiro and Bates (2000), by including a nonparametric component $f$, in analogy with models introduced, e.g., in Wood (2017). The nonparametric term, as detailed in Section 2.1, captures the spatio-temporal structure of the data while incorporating prior physical information about the underlying phenomenon.

In the semiparametric model (1), we aim to estimate the smooth function $f$, the fixed effects $\boldsymbol{\beta}$, and the random effect covariance matrix $\Sigma_{\mathbf{b}}$, which characterizes the variability across groups. To this end, leveraging the framework proposed in Arnone et al. (2023b), we minimize the following loss functional:

$$J(\boldsymbol{\beta}, f, \Sigma_{\mathbf{b}}) = \frac{1}{|\mathcal{O}|} \sum_{k=1}^{g} \sum_{(i,j)\in\mathcal{O}_k} \left(y_{ij} - \mathbf{x}_{ij}^\top \boldsymbol{\beta} - f(\mathbf{p}_i, t_j) - \mathbf{z}_{ij}^\top \mathbf{b}_k\right)^2 + \mathcal{P}_{\lambda_{\mathscr{D}}, \lambda_T}(f), \tag{2}$$

The first term represents a quadratic data fidelity term, as in classical linear mixed-effects regression models, while the second term, $\mathcal{P}_{\lambda_{\mathscr{D}}, \lambda_T}(f)$, is a physics-informed penalty that enforces spatio-temporal smoothness of $f$ and incorporates prior knowledge on the physical dynamics of the process, as detailed in Section 2.1. The penalty $\mathcal{P}_{\lambda_{\mathscr{D}}, \lambda_T}(f)$ is weighted by two positive parameters $\lambda_{\mathscr{D}}$ and $\lambda_T$, which control the trade-off between data fidelity, measured by the quadratic loss in the first term of (2), and the spatial and temporal regularity of the field $f$, enforced by $\mathcal{P}_{\lambda_{\mathscr{D}}, \lambda_T}(f)$. The values of these smoothing parameters are selected using the Generalized Cross-Validation criterion (see, e.g., Craven and Wahba, 1978; Wahba, 1985; Arnone et al., 2023b). Importantly, the functional (2) depends on $\Sigma_{\mathbf{b}}$ through the random coefficients $\mathbf{b}_k$. This induces a non-quadratic structure of the estimation functional, which prevents a closed-form solution and necessitates the use of an iterative estimation scheme, as described in Section 3.

### *2.1. Including physical information*

In many applications, especially in environmental settings, prior physical knowledge about the phenomenon under study is available and can be incorporated into the model through suitable regularization terms. Such information is often formalized via partial differential equations (PDEs) describing the spatio-temporal evolution of the process. Here, we incorporate such PDE-based physical information into the statistical model through the penalty term

$$\mathcal{P}_{\lambda_{\mathscr{D}},\lambda_T}(f) = \lambda_{\mathscr{D}}\mathcal{P}_{\mathscr{D}}(f) + \lambda_T\mathcal{P}_T(f),$$

where

$$\begin{aligned} \mathcal{P}_{\mathscr{D}}(f) &= \int_0^T \int_{\mathscr{D}} \left(\mathcal{L}(\mathbf{p})f(\mathbf{p},t) - u(\mathbf{p},t)\right)^2 \, d\mathbf{p}\, dt, \\ \mathcal{P}_T(f) &= \int_0^T \int_{\mathscr{D}} \left(\frac{\partial^2 f(\mathbf{p},t)}{\partial t^2}\right)^2 \, d\mathbf{p}\, dt. \end{aligned} \tag{3}$$

The operator $\mathcal{L}(\mathbf{p})$ is a linear, second-order elliptic operator defining the diffusion-advection-reaction PDE $\mathcal{L}(\mathbf{p})f(\mathbf{p},t) = u(\mathbf{p},t)$, where

$$\mathcal{L}(\mathbf{p})f(\mathbf{p},t) = -\nabla \cdot (K(\mathbf{p})\nabla f(\mathbf{p},t)) + \boldsymbol{\gamma}(\mathbf{p}) \cdot \nabla f(\mathbf{p},t) + c(\mathbf{p})f(\mathbf{p},t). \tag{4}$$

Specifically, the operator involves the gradient $\nabla = (\partial/\partial p_1, \partial/\partial p_2)^\top$; the diffusion tensor $K : \mathscr{D} \to \mathcal{S}^+$ encoding (possibly anisotropic) diffusion processes, where $\mathcal{S}^+$ denotes the space of symmetric positive definite matrices; the advection field $\boldsymbol{\gamma} : \mathscr{D} \to \mathbb{R}^2$, which models directional transport in the spatial domain $\mathscr{D}$; and the scalar reaction term $c : \mathscr{D} \to \mathbb{R}$, which controls shrinkage of the field. The PDE parameters $K(\mathbf{p})$, $\boldsymbol{\gamma}(\mathbf{p})$, and $c(\mathbf{p})$ can vary over space, modeling different forms of physically-informed non-stationarity. Finally, the space-time varying forcing term $u : \mathscr{D} \times [0,T] \to \mathbb{R}$ in the spatial penalty $\mathcal{P}_{\mathscr{D}}(f)$ represents possible exogenous inputs. In this work, for simplicity, we restrict to the homogeneous case, with $u \equiv 0$, and we refer the reader to Azzimonti et al. (2014) for an analysis of nonhomogeneous forcing terms.

The simplest instance of (3) arises when $\mathcal{L}$ reduces to the Laplace operator, that is $(K, \boldsymbol{\gamma}, c) = (I, \mathbf{0}, 0)$, as considered by Bernardi et al. (2018) in the simpler setting of spatio-temporal regression with fixed effects only. More general formulations, involving advection-diffusion dynamics, have been proposed in spatial regression without random effects by Azzimonti et al. (2014, 2015); Arnone et al. (2019) and, for the linear case, by Tomasetto et al. (2024), as well as by De Sanctis et al. (2025) and Castiglione et al. (2025) for quantile regression problems.

In general, $K$, $\boldsymbol{\gamma}$, and $c$ may depend on hyperparameters $\boldsymbol{\xi}$, which can be estimated from data using a parameter cascading algorithm (see, e.g., Ramsay et al., 2007; Xun et al., 2013), as discussed in Bernardi et al. (2018) and Tomasetto et al. (2024). For instance, in the simulation studies in Section 5, we model spatial anisotropy through a diffusion tensor $K(\boldsymbol{\xi})$, where $\boldsymbol{\xi}$ are unknown parameters controlling the direction and intensity of the anisotropy, estimated from the data via a parameter cascading approach; see Bernardi et al. (2018). In the $NO_2$ case study over Lombardy, we instead adopt a diffusion-transport regularization term,

$$-\Delta f(\mathbf{p}, t) + \xi\, \boldsymbol{\gamma}(\mathbf{p}) \cdot \nabla f(\mathbf{p}, t),$$

which combines an isotropic diffusion component, representing atmospheric dispersion, with a space-varying transport term driven by the wind field $\boldsymbol{\gamma}$, encoding both its local direction and intensity. The relative contribution of the two components is governed by an unknown parameter, which is estimated from the data via parameter cascading; see Tomasetto et al. (2024) for details.

### *2.2. Estimation problem*

In order to properly define the estimation problem, we introduce suitable functional spaces for the nonparametric component $f$. Let $H^2(\mathscr{D})$ denote the second-order Sobolev space, i.e., the space of functions in $L^2(\mathscr{D})$ with square-integrable weak derivatives up to order two. We further denote by $L^2(0, T; H^2(\mathscr{D}))$ the space of square-integrable functions on $(0, T)$ with values in $H^2(\mathscr{D})$, and by $L^2(0, T; L^2(\mathscr{D}))$ the space of square-integrable functions on $(0, T)$ with values in $L^2(\mathscr{D})$.

The regularization terms in (3) are well-defined on the space

$$V = \left\{ f \in L^2(0, T; H^2(\mathscr{D})) : \frac{\partial^2 f}{\partial t^2} \in L^2(0, T; L^2(\mathscr{D})) \right\}.$$

Let $\nu$ denote the outward unit normal vector to the boundary $\partial\mathscr{D}$. We restrict $V$ by imposing homogeneous Neumann boundary conditions on $f$, namely $\nabla f \cdot \nu = 0$, leading to the space

$$V_{BC} = \left\{ f \in V : \nabla f \cdot \nu = 0 \text{ on } \partial\mathscr{D} \times (0, T] \right\}.$$

Homogeneous Neumann boundary conditions are so-called natural conditions for the class of differential models considered here, and $V_{BC}$ provides a suitable functional embedding for the estimation of the nonparametric component $f$. Alternative boundary conditions, including Dirichlet, Robin, or mixed types, can also be considered, allowing for flexible modeling of the behavior of $f$ at the boundary of the domain $\mathscr{D}$; see, e.g., Azzimonti et al. (2015) for details.

We thus consider the estimation problem

$$\underset{\boldsymbol{\beta} \in \mathbb{R}^q,\ f \in V_{BC},\ \Sigma_{\mathbf{b}} \in \mathcal{S}^+}{\arg\min} \quad J(\boldsymbol{\beta}, f, \Sigma_{\mathbf{b}}). \tag{5}$$

We point out that, unlike simpler regression models considered, e.g., in Bernardi et al. (2017); Arnone et al. (2019); Augustin et al. (2013); Marra et al. (2012); Arnone et al. (2023b), problem (5) is not quadratic due to the presence of the covariance matrix $\Sigma_{\mathbf{b}}$. This makes the estimation problem more involved, requiring iterative strategies to approximate (5), as described in Section 3.

## 3. Model estimation

In this section, we present the iterative strategy used to solve the estimation problem in (5) and to estimate $(\boldsymbol{\beta}, f, \Sigma_{\mathbf{b}})$. To this end, we first reformulate model (1) as a fixed effects regression model with correlated errors, following the classical approach described in Wood (2017). By collecting the group-specific quantities into block vectors, model (1) can be written as

$$\underbrace{\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_g \end{bmatrix}}_{\mathbf{y}} = \underbrace{\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_g \end{bmatrix}}_{X} \boldsymbol{\beta} + \underbrace{\begin{bmatrix} \boldsymbol{f}_1 \\ \boldsymbol{f}_2 \\ \vdots \\ \boldsymbol{f}_g \end{bmatrix}}_{\boldsymbol{f}} + \underbrace{\begin{bmatrix} Z_1\mathbf{b}_1 + \boldsymbol{\varepsilon}_1 \\ Z_2\mathbf{b}_2 + \boldsymbol{\varepsilon}_2 \\ \vdots \\ Z_g\mathbf{b}_g + \boldsymbol{\varepsilon}_g \end{bmatrix}}_{\mathbf{e}},$$

which can be written in compact form as

$$\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{f} + \mathbf{e}. \tag{6}$$

The vector $\mathbf{e}$ collects both the random effect contributions and the measurement errors, and has a non-diagonal covariance structure. The formulation in (6) shows that the proposed mixed-effects model can be equivalently expressed as a fixed effects regression model with heteroskedastic and correlated errors, where the random components $Z_k\mathbf{b}_k$ and the noise terms $\boldsymbol{\varepsilon}_k$ are combined into a single stochastic term $\mathbf{e}$.

Since, by assumption, $\mathbf{b}_k$ and $\boldsymbol{\varepsilon}_k$ are independent Gaussian vectors for each $k$, it follows that

$$\mathbf{e} \sim \mathcal{N}(\mathbf{0}, \sigma^2\Sigma_{\mathbf{e}}),$$

where $\Sigma_e$ is a block diagonal matrix with blocks $\Sigma_{e,k} = \frac{1}{\sigma^2} Z_k \Sigma_{\mathbf{b}} Z_k^\top + I_{|\mathcal{O}_\mathrm{k}|}$, for $k = 1, \ldots, g$, and $I_{|\mathcal{O}_\mathrm{k}|}$ denotes the identity matrix of dimension $|\mathcal{O}_\mathrm{k}|$. The penalized negative log-likelihood of this model is then given by

$$\frac{1}{2\,|\mathcal{O}|} ||(\sigma^2\Sigma_{\mathbf{e}})^{-1/2}(\mathbf{y} - X\boldsymbol{\beta} - \boldsymbol{f})||^2 + \frac{\log\left(\sqrt{(2\pi)^{|\mathcal{O}|}\sigma^2 \det(\Sigma_{\mathbf{e}})}\right)}{|\mathcal{O}|} + \mathcal{P}_{\lambda_{\mathscr{D}},\lambda_T}(f), \quad (7)$$

where $\det(\Sigma_{\mathbf{e}})$ denotes the determinant of $\Sigma_{\mathbf{e}}$.

Therefore, solving (5) is equivalent to minimizing the penalized negative log-likelihood (7). The resulting objective functional is non-quadratic, due to the dependence of $\Sigma_{\mathbf{e}}$ on the unknown covariance matrix $\Sigma_{\mathbf{b}}$, which induces a nonlinear coupling between the parameters and prevents a closed-form solution. This motivates the use of iterative estimation schemes. In particular, this formulation naturally leads to a Functional Penalized Iterative Reweighted Least Squares (FPIRLS) algorithm, extending the classical PIRLS method (O'sullivan et al., 1986) to the functional setting. In the context of physics-informed penalized regression for spatially dependent data, related approaches have been proposed by Wilhelm and Sangalli (2016) for generalized linear models and by Castiglione et al. (2025) for quantile regression. The estimation problem considered here is more involved, as it requires the joint estimation of the fixed effects $(\boldsymbol{\beta}, f)$ and the variance-covariance matrix of the random effects $\Sigma_{\mathbf{b}}$. In particular, (7) cannot be minimized jointly with respect to $(\boldsymbol{\beta}, f, \Sigma_{\mathbf{e}})$. We therefore adopt an iterative two-step procedure, as discussed in Wood (2017) for the simpler purely parametric case. At each iteration, the algorithm alternates between the following two steps. In the first step, we estimate the fixed effects $(\boldsymbol{\beta}, f)$ for a given covariance matrix $\Sigma_{\mathbf{b}}$, as detailed in Section 3.1. In the second step, we update the random effects covariance matrix $\Sigma_{\mathbf{b}}$ for given $(\boldsymbol{\beta}, f)$, as described in Section 3.2.

*3.1. Estimation of fixed effects*

We now describe the estimation procedure for the fixed effects $(\boldsymbol{\beta}, f)$, for a given covariance matrix of the random effects $\Sigma_{\mathbf{b}}$. By neglecting the terms in (7) that do not depend on $(\boldsymbol{\beta}, f)$, the problem reduces to the following penalized least-squares formulation:

$$(\hat{\boldsymbol{\beta}}, \hat{f}) = \underset{\boldsymbol{\beta}\in\mathbb{R}^q,\ f\in V_{BC}}{\arg\min} \left\{ \frac{1}{|\mathcal{O}|} \|\Sigma_{\mathbf{e}}^{-\frac{1}{2}}(\mathbf{y} - X\boldsymbol{\beta} - \boldsymbol{f})\|^2 + \mathcal{P}_{\lambda_{\mathscr{D}},\lambda_T}(f) \right\}. \tag{8}$$

Problem (8) can be solved using the techniques described in Sangalli (2021). In particular, to handle the PDE penalty and the possibly irregular geometry of the spatial domain $\mathscr{D}$, we adopt a numerical discretization scheme based on finite element bases in space and cubic B-spline bases in time. Let $\{\psi_1(\mathbf{p}), \dots, \psi_N(\mathbf{p})\}$ be a set of $N$ finite element basis functions (i.e., piecewise polynomial functions) defined on a triangulation $\mathscr{D}_\tau$ of the domain $\mathscr{D}$, and let $\{\varphi_1(t), \dots, \varphi_M(t)\}$ be a set of $M$ cubic B-spline basis functions defined on the time interval $[0, T]$. Let $\Psi = \{\psi_\ell(\mathbf{p}_i)\}_{i,\ell} \in \mathbb{R}^{n\times N}$ be the matrix of spatial evaluations of the $N$ finite element bases at the $n$ locations $\{\mathbf{p}_1, \dots, \mathbf{p}_n\}$, and $\Phi = \{\varphi_r(t_j)\}_{j,r} \in \mathbb{R}^{m\times M}$ be the matrix of temporal evaluations of the $M$ spline bases at the $m$ time instants $\{t_1, \dots, t_m\}$. We then represent the nonparametric term $f$ by the basis expansion

$$f(\mathbf{p}, t) = \sum_{\ell=1}^{N} \sum_{r=1}^{M} f_{\ell r}\, \psi_\ell(\mathbf{p})\, \varphi_r(t),$$

and collect the coefficients of the basis expansion in the vector of coefficients $\mathbf{f} \in \mathbb{R}^{NM}$. To account for the missingness pattern of the data, we define $B$ as the sub-matrix of $\Phi \otimes \Psi$ obtained by removing the $(i+nj)$-th row whenever the datum at $(\mathbf{p}_i, t_j)$ is not observed. Additionally, we define the following matrices:

$$H = X(X^\top \Sigma_{\mathbf{e}}^{-1} X)^{-1} X^\top \Sigma_{\mathbf{e}}^{-1}, \quad Q = \Sigma_{\mathbf{e}}^{-1}(I - H).$$

Following the same arguments as in Arnone et al. (2023b), we can estimate the fixed effects estimator $(\check{\boldsymbol{\beta}}, \check{\mathbf{f}}) \in \mathbb{R}^q \times \mathbb{R}^{NM}$, at each iteration of the FPIRLS algorithm, as:

$$\begin{aligned} \check{\boldsymbol{\beta}} &= \left(X^\top \Sigma_{\mathbf{e}}^{-1} X\right)^{-1} X^\top \Sigma_{\mathbf{e}}^{-1}(\mathbf{y} - B\check{\mathbf{f}}), \\ \check{\mathbf{f}} &= \frac{1}{|\mathcal{O}|} \left( \frac{1}{|\mathcal{O}|} B^\top Q B + P \right)^{-1} B^\top Q \mathbf{y}, \end{aligned} \tag{9}$$

where $P$ is the discrete counterpart of the penalty $\mathcal{P}_{\lambda_{\mathscr{D}},\lambda_T}(f)$ in (3), as detailed in Section S1.5 of the supplementary material.

*3.2. Covariance estimation of random effects*

In this section, we describe the procedure to estimate the covariance matrix of the random effects $\Sigma_{\mathbf{b}}$, for given $(\boldsymbol{\beta}, f)$, obtained according to (9).

We define $D = \Sigma_{\mathbf{b}}/\sigma^2$ as the relative precision matrix of the model, and we denote by $\Delta$ a relative precision factor, that is, a matrix such that $D^{-1} = \Delta^\top \Delta$. Such a matrix $\Delta$ always exists, although it is not unique in general. We estimate $D$ via the Expectation-Maximization (EM) algorithm, which guarantees a monotone increase of the likelihood at each iteration. To this end, we introduce the following pseudo-data matrices:

$$\tilde{\mathbf{y}}_k = \begin{bmatrix} \mathbf{y}_k \\ \mathbf{0} \end{bmatrix}, \quad \tilde{X}_k = \begin{bmatrix} X_k \\ \mathbf{0} \end{bmatrix}, \quad \tilde{\hat{\boldsymbol{f}}}_k = \begin{bmatrix} \hat{\boldsymbol{f}}_k \\ \mathbf{0} \end{bmatrix}, \quad \tilde{Z}_k = \begin{bmatrix} Z_k \\ \Delta \end{bmatrix}, \quad \forall k = 1, \ldots, g.$$

We denote by $R_k$ the upper triangular matrix in the QR decomposition of $\tilde{Z}_k$. We then define matrix $L \in \mathbb{R}^{(g+1)p \times p}$ as

$$L = \begin{bmatrix} \hat{\mathbf{b}}_1^\top/\sigma \\ (R_1^{-1})^\top \\ \vdots \\ \hat{\mathbf{b}}_g^\top/\sigma \\ (R_g^{-1})^\top \end{bmatrix}. \tag{10}$$

where $\hat{\mathbf{b}}_k$ denotes the conditional maximum likelihood estimate of $\mathbf{b}_k$, for $k = 1, \ldots, g$. Finally, we denote by $A$ the triangular factor in the QR decomposition of $L$.

Within the EM algorithm, the E-step replaces the unobserved random effects $\mathbf{b}_k$ with their conditional expectations, approximated here by $\hat{\mathbf{b}}_k$. In the M-step, these estimates are plugged into the conditional likelihood of model (6), which is then maximized with respect to $D$ (Bates and Pinheiro, 1998).

The following proposition provides the analytical expression of the estimator of the relative precision matrix $D$.

**Proposition 3.1.** *For a given pair of fixed effects $(\hat{\boldsymbol{\beta}}, \hat{f})$, the maximizer of the conditional likelihood of model* (6) *is given by*

$$\hat{D} = \frac{AA^\top}{g}.$$

The proof is deferred to Section S1.1 of the supplementary material.

## 4. Asymptotic distribution

Consistently with the estimation strategy introduced in Section 3, we derive asymptotic results for the fixed effects, for a fixed value of $\Sigma_{\mathbf{b}}$. Similarly, we study the asymptotic properties of the covariance estimator $\hat{\Sigma}_{\mathbf{b}}$ for fixed values of the fixed effect components.

Let $(\breve{\boldsymbol{\beta}}_{|\mathcal{O}|}, \breve{\mathbf{f}}_{|\mathcal{O}|}) \in \mathbb{R}^q \times \mathbb{R}^{NM}$ denote the vector of discretized estimators of $(\boldsymbol{\beta}, f)$ at FPIRLS convergence, where the notation $(\cdot)_{|\mathcal{O}|}$ indicates dependence on the sample size $|\mathcal{O}|$. We define the matrices

$$\Omega_{|\mathcal{O}|} = |\mathcal{O}|(B^\top Q B)^{-1}, \quad \Xi_{|\mathcal{O}|} = \frac{X^\top \Sigma_{\mathbf{e}}^{-1} X}{|\mathcal{O}|}.$$

The following two propositions state the asymptotic distributions of the fixed effect estimators $(\breve{\boldsymbol{\beta}}_{|\mathcal{O}|}, \breve{\mathbf{f}}_{|\mathcal{O}|})$.

**Proposition 4.1.** *Assume that the limits*

$$\lim_{|\mathcal{O}| \to +\infty} \Omega_{|\mathcal{O}|} = \Omega, \quad \lim_{|\mathcal{O}| \to +\infty} \Xi_{|\mathcal{O}|} = \Xi,$$

*exist and are non-singular. If* $\lambda_{\mathscr{D}} \sqrt{|\mathcal{O}|} \to \overline{\lambda}_{\mathscr{D}}$ *and* $\lambda_T \sqrt{|\mathcal{O}|} \to \overline{\lambda}_T$, *for some finite values* $\overline{\lambda}_{\mathscr{D}}$ *and* $\overline{\lambda}_T$, *then* $\breve{\mathbf{f}}_{|\mathcal{O}|}$ *has asymptotic distribution*

$$\sqrt{|\mathcal{O}|}(\breve{\mathbf{f}}_{|\mathcal{O}|} - \mathbf{f}) \xrightarrow{d} \mathcal{N}_{NM}(\mathbf{0}, \sigma^2 \Omega).$$

*Moreover,* $\breve{\mathbf{f}}_{|\mathcal{O}|}$ *is consistent for* $\mathbf{f}$, *that is,* $\breve{\mathbf{f}}_{|\mathcal{O}|}$ *converges to* $\mathbf{f}$ *in probability.*

**Proposition 4.2.** *Let* $\{\breve{\mathbf{f}}_{|\mathcal{O}|}\}$ *be a sequence of consistent estimators for* $\mathbf{f}$. *Under the same assumptions stated in Proposition 4.1, the estimator* $\breve{\boldsymbol{\beta}}_{|\mathcal{O}|}$ *has asymptotic distribution*

$$\sqrt{|\mathcal{O}|}(\breve{\boldsymbol{\beta}}_{|\mathcal{O}|} - \boldsymbol{\beta}) \xrightarrow{d} \mathcal{N}_q \left( \mathbf{0},\ \sigma^2 \left( \Xi^{-1} + \frac{1}{|\mathcal{O}|^2} \Xi^{-1} (\Sigma_{\mathbf{e}}^{-1} X)^\top B \Omega B^\top (\Sigma_{\mathbf{e}}^{-1} X) \Xi^{-1} \right) \right).$$

The following proposition provides the asymptotic distribution of the variance estimators $(\hat{\sigma}, \hat{\Sigma}_{\mathbf{b}})$ of the model, in the case of independent random effects.

**Proposition 4.3.** *Assume that the random effects* $\mathbf{b}_k$ *are independent, that is,* $\Sigma_{\mathbf{b}}$ *is a diagonal matrix with diagonal* $(\sigma_{b,1}^2, \ldots, \sigma_{b,p}^2)$. *Then, for* $|\mathcal{O}| \to +\infty$,

$$\begin{bmatrix} \log\hat{\sigma} \\ \hat{\sigma}_{b,1} \\ \vdots \\ \hat{\sigma}_{b,p} \end{bmatrix} \xrightarrow{d} \mathcal{N}\left(\begin{bmatrix} \log\sigma \\ \sigma_{b,1} \\ \vdots \\ \sigma_{b,p} \end{bmatrix}, \mathcal{I}_{\sigma,\Sigma_{\mathbf{b}}}^{-1}\right), \tag{11}$$

*where* $\mathcal{I}_{\sigma,\Sigma_{\mathbf{b}}}$ *denotes the empirical information matrix, given by*

$$\mathcal{I}_{\sigma,\Sigma_{\mathbf{b}}} = \begin{bmatrix} \frac{2}{\hat{\sigma}^2}\sum_{k=1}^{g}\|\boldsymbol{\varepsilon}_k\|^2 & 0 & 0 & 0 \\ 0 & \frac{3}{\hat{\sigma}_{b,1}^4}\sum_{k=1}^{g}\|\mathbf{b}_k\|^2 - \frac{g}{\hat{\sigma}_{b,1}^2} & 0 & 0 \\ \vdots & \vdots & \vdots & 0 \\ 0 & 0 & \cdots & \frac{3}{\hat{\sigma}_{b,p}^4}\sum_{k=1}^{g}\|\mathbf{b}_k\|^2 - \frac{g}{\hat{\sigma}_{b,p}^2} \end{bmatrix}.$$

The proofs of Propositions 4.1, 4.2, and 4.3 are deferred to Sections S1.2, S1.3, and S1.4 of the supplementary material, respectively.

## 5. Simulation study

In this section, we evaluate the performance of the proposed model, denoted as Mixed-Effect Spatio-Temporal Regression with Partial Differential Equation regularization (`MEST-PDE`), against state-of-the-art methods for spatio-temporal regression with mixed-effects. The proposed approach is implemented in the `fdaPDE` library (Palummo et al., 2025).

We randomly sample $n = 100$ spatial locations over the unit square domain $\mathscr{D} = [0,1]^2$, and partition them into $g = 6$ groups. For the time dimension, we consider $m = 11$ equispaced instants in the unit interval $[0,1]$. We consider the model

$$\mathbf{y}_k = X_k\boldsymbol{\beta} + \boldsymbol{f}_k + Z_k b_k + \boldsymbol{\varepsilon}_k, \quad k = 1,\ldots,6, \tag{12}$$

where $Z_k = (1,\ldots,1)^\top$, that is, the model includes a random intercept only, as motivated by the $NO_2$ case study in Section 6. For the true spatio-temporal nonparametric term $f$, we consider a realization of a Gaussian field with Matérn covariance, generated using the `spate.sim` function from the R package `spate` (Sigrist et al., 2015), specifying anisotropy with intensity 8 and angle $\pi/4$. For each time instant, we generate the two independent fixed effect covariates $X_k = (\mathbf{x}_{1,k}, \mathbf{x}_{2,k})$ as Gaussian random fields using the R function `grf` from the package `geoR` (Ribeiro Jr and Diggle, 2025). Further details about the generation of the fixed effect covariates, together with their spatio-temporal visualization, are reported in Section S2.1 of the Supplementary Material. Finally, we set $\boldsymbol{\beta} = (1, -1)$, $b_k \sim \mathcal{N}(0, \sigma^2_{\mathbf{b}})$, and $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \sigma^2 I_{|\mathcal{O}|})$, and we set $\sigma^2 = 1/4^2$ and $\sigma_{\mathbf{b}}$ so that the ratio $\sigma^2_{\mathbf{b}}/(\sigma^2 + \sigma^2_{\mathbf{b}})$ equals 0.30. The generation of the noise $\boldsymbol{\varepsilon}$ is repeated 100 times, obtaining 100 independent replicas of the data.

In this setting, we consider the proposed `MEST-PDE` model, together with its isotropic variant, denoted as `MEST-ISO`, which assumes isotropic diffusion in the estimation process. Both `MEST-PDE` and `MEST-ISO` are implemented using a regular triangulation of the square domain with 476 nodes and linear finite elements. The `MEST-PDE` model includes a purely diffusive differential regularization term with unknown hyperparameters, introduced to capture anisotropy in the data. These hyperparameters, which determine the magnitude and orientation of the anisotropy, are estimated through the parameter cascading algorithm, as described in Section 2.

We compare these two methods with alternative techniques available in the literature, focusing on those implemented in widely used software. Specifically, we consider the Generalized Additive Mixed Model introduced in Wood (2006) and implemented via the `gamm` function in the `mgcv` package (Wood and Wood, 2015). As in the proposed `MEST-PDE` model, we adopt B-spline bases for the temporal discretization of all competing methods. For the spatial discretization, these approaches rely on thin plate spline bases (Wahba, 1990) and on soap film smoothing (Wood et al., 2008). We refer to these two alternatives as `TPS` and `SOAP`, respectively. We also consider the corresponding implementations based on the `gamm4` function (Wood et al., 2017), which relies on the `lme4` framework (Bates et al., 2015) instead of `nlme` (Pinheiro et al., 2017), as in `gamm`. We denote these variants as `TPS4` and `SOAP4`, respectively. For `TPS` and `TPS4`, we employ 75 and 50 spatial basis functions, respectively. For both `SOAP` and `SOAP4`, we use 8 spatial basis functions, along with 8 additional basis functions to define the boundary-interpolating soap film (see Wood et al., 2008). This configuration represents the maximum number of basis functions ensuring numerical stability, as larger values lead to execution failure.

For the temporal dimension, all methods use 10 cubic B-spline basis functions. Finally, smoothness selection is performed via Restricted Maximum Likelihood for `TPS4` and `SOAP4`, while for the other methods it is based on

Generalized Cross-Validation.

We have also explored the R package `sdmTMB` (Anderson et al., 2022), which is based on R-INLA (Rue et al., 2009). However, this model produces unstable estimates of the nonparametric component $f$, which affect the overall results. For this reason, we exclude `sdmTMB` from the following discussion.

Model accuracy is assessed using the Root Mean Squared Error (RMSE) as a global measure of reconstruction error, which, for a given nonparametric term $f$, is defined as

$$RMSE(\hat{f}) = \sqrt{\int_{\mathscr{D}} \int_0^T \left( f(\mathbf{p}, t) - \hat{f}(\mathbf{p}, t) \right)^2 d\mathbf{p}\, dt},$$

where $\hat{f}$ represents the estimated field. The RMSE is approximated over a fine spatio-temporal grid of 27,500 points.

Figure 3 displays the observed data, together with the true and estimated nonparametric fields, at selected time instants. The `MEST-PDE` model more accurately captures the anisotropic features of the field, particularly at later time instants, where competing methods tend to oversmooth the central region of the domain. This behavior is reflected in the RMSE of the nonparametric component, reported in the left panel of Figure 4, where `MEST-PDE` achieves the best performance, highlighting the benefit of incorporating anisotropic penalization. The last two panels of Figure 4 show the estimates of the fixed effect regression coefficients. All methods exhibit comparable accuracy in estimating $\beta_1$, although a slight negative bias is observed, particularly for `TPS4`. In contrast, for $\beta_2$, the `MEST-PDE` model provides more accurate estimates than the competing approaches.

Regarding the estimation of the random component, Figure 5 shows the boxplots of the variance terms $\sigma_{\mathbf{b}}$ and $D$. All methods exhibit comparable performance in estimating $\sigma_{\mathbf{b}}$, with similar levels of variability. For the precision factor $D$, the `MEST-PDE` and `MEST-ISO` models provide more accurate estimates, which can be attributed to a more reliable reconstruction of the noise variance $\sigma^2$.

Section S2 of the Supplementary Material reports additional simulation results comparing accuracy and computational cost under increasing sample sizes, highlighting the favorable scalability of the proposed MEST-PDE. These results show that the method is substantially faster than `TPS4` and `SOAP4`, while `TPS` and `SOAP` are computationally cheaper but significantly

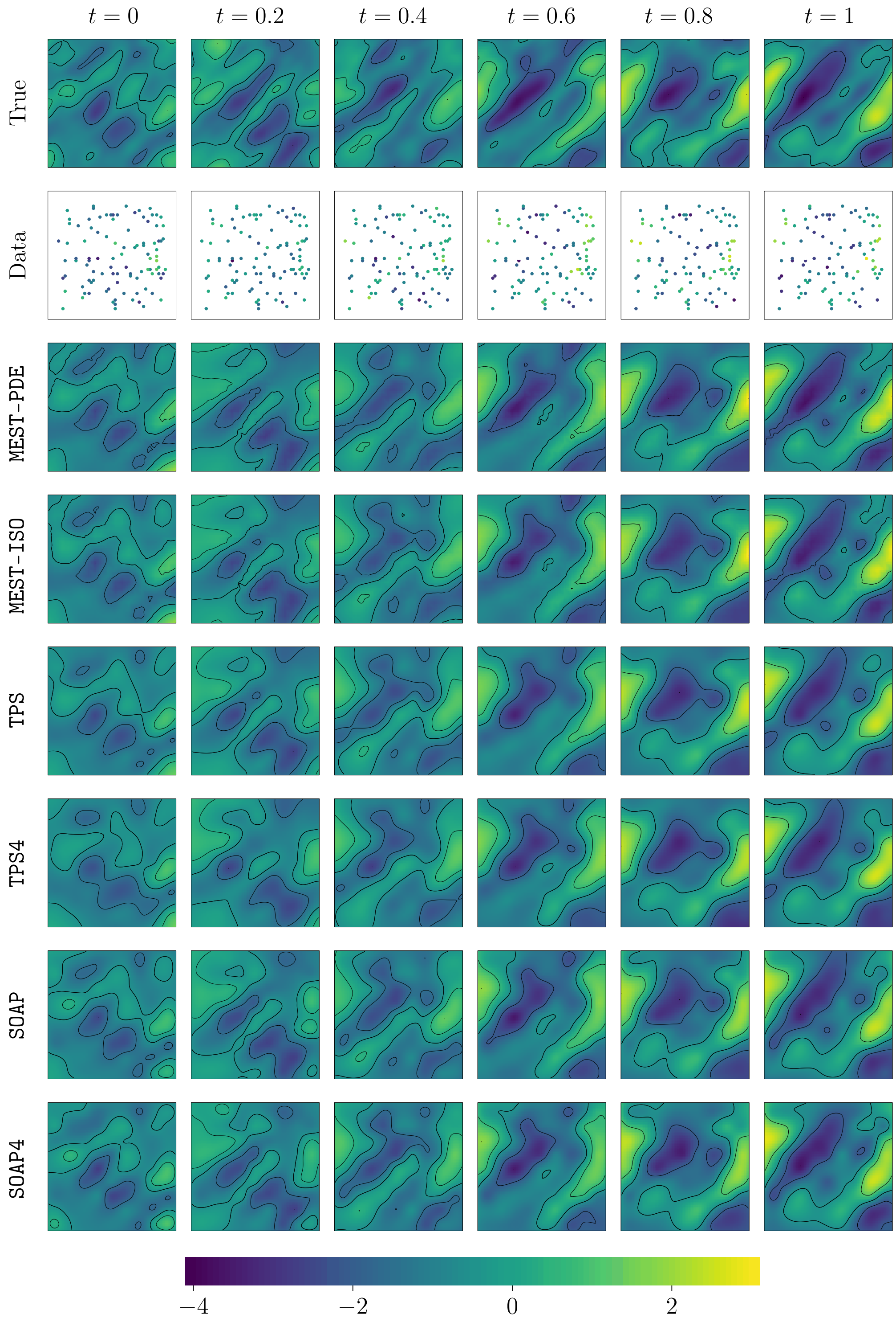


Figure 3: True and estimated nonparametric terms at selected time instants. The first row shows the true field, the second row displays the observed data for a fixed replica, while the subsequent rows report the estimated fields (averaged over the 100 replicates) for each of the competing methods: the proposed Mixed-Effect Spatio-Temporal Regression with Partial Differential Equation regularization (`MEST-PDE`); its isotropic counterpart (`MEST-ISO`); Thin Plate Spline models based on `nlme` (`TPS`) and `lme4` (`TPS4`); and Soap film smoothing based on `nlme` (`SOAP`) and `lme4` (`SOAP4`).

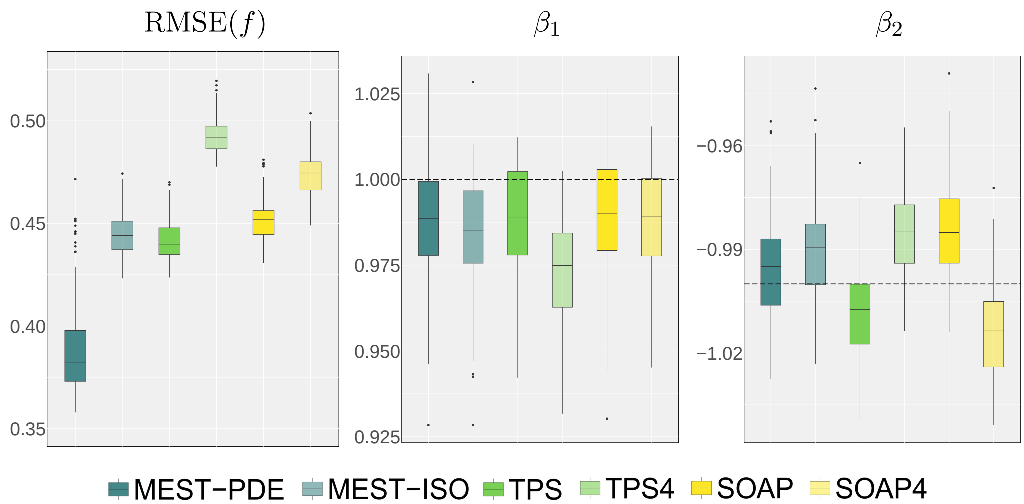


Figure 4: Accuracy comparison of the fixed effect estimates obtained by the competing methods: the proposed Mixed-Effect Spatio-Temporal Regression with Partial Differential Equation regularization (`MEST-PDE`); its isotropic counterpart (`MEST-ISO`); thin plate spline models based on `nlme` (`TPS`) and `lme4` (`TPS4`); and soap film smoothing based on `nlme` (`SOAP`) and `lme4` (`SOAP4`). Left panel: RMSE of the nonparametric field $f$. Central panel: estimates of $\beta_1$. Right panel: estimates of $\beta_2$.

less accurate. In particular, increasing the number of spatial basis functions for `TPS` and `SOAP` is not feasible in practice, as it leads to execution failure, thereby preventing a comparison of computational costs at comparable levels of accuracy with `MEST-PDE`. Additionally, `MEST-PDE` shows a comparable or better asymptotic computational complexity than `TPS` and `SOAP`, and a significantly better asymptotic computational complexity than `TPS4` and `SOAP4`.

## 6. Dealing with sensor heterogeneity in air quality assessment

In this section, we analyze the spatio-temporal concentration of nitrogen dioxide ($NO_2$) over Lombardy. This pollutant is primarily emitted from combustion processes, such as road traffic and industrial activities; elevated concentrations are associated with adverse effects on both human health and ecosystems (Chen et al., 2024). Within a single day, hourly $NO_2$ concentrations can vary substantially due to human-related factors such as traffic flows and heating demand. While these high-frequency measurements provide a detailed representation of air quality dynamics, they also pose addi-

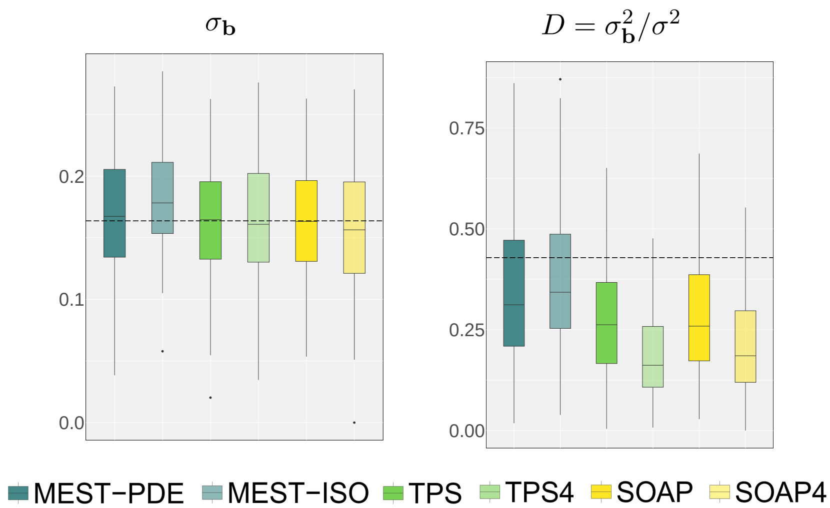


Figure 5: Estimated variance components obtained by the competing methods: the proposed Mixed-Effect Spatio-Temporal Regression with Partial Differential Equation regularization (`MEST-PDE`); its isotropic counterpart (`MEST-ISO`); thin plate spline models based on `nlme` (`TPS`) and `lme4` (`TPS4`); and soap film smoothing based on `nlme` (`SOAP`) and `lme4` (`SOAP4`). Left panel: estimated standard deviation of the random effects. Right panel: estimated relative precision factor.

tional modeling challenges due to stronger variability compared to daily or monthly aggregates. Nevertheless, capturing such short-term fluctuations is essential for exposure assessment, as even brief $NO_2$ peaks may have significant health impacts (see, e.g., World Health Organization, 2024). Figure 1 illustrates $NO_2$ measurements on 15 January 2019, provided by ARPA and publicly available from Open Data Lombardia, Trasformazione Digitale in Lombardia (2024). The data exhibit a clear diurnal pattern, with higher concentrations in the morning and evening, reflecting traffic intensity and commuting behavior. The dataset contains a small proportion of missing observations (approximately 2.41% of the total), which, as discussed in the previous sections, can be naturally handled within our modeling framework. Additionally, analyzing the daily pattern of $NO_2$ concentrations allows us to incorporate information about the day's wind field, which plays a crucial role in pollutant dispersion. In particular, the wind field, shown in Figure 2, is introduced in the regularizing term as an appropriate transport term, as described below.

As anticipated in Section 1, we consider distinct sources of variability in

$NO_2$ concentrations, in order to disentangle the contributions of physical, morphological, anthropogenic, and technological factors influencing air pollution levels. To account for the heterogeneity of measurement technologies across the ARPA monitoring network, we include in the statistical model a random effect, specified as a station-specific random intercept $b_k$. This term captures the variability introduced by the measurement process of the instrumentation, allowing the model to separate sensor-related noise from the underlying spatio-temporal dynamics of the phenomenon of interest. Note that a purely fixed effect model including dummy variables would not capture the variability induced by the measurement process. The spatial distribution of the different sensor technologies, namely *API*, *enviro*, *serinus*, and *thermo*, is illustrated in the bottom-right panel of Figure 1.

Another source of variability arises from both geographical and anthropogenic characteristics of the territory. According to recent literature, natural and human-driven factors are key drivers of $NO_2$ formation (see, e.g., Salama et al., 2022; California Air Resources Board, 2023). From a morphological perspective, the Po Valley is subject to frequent thermal inversion phenomena, which limit air circulation and contribute to the persistence of pollutants in lowland areas (Trinh et al., 2019). To account for this effect, we include altitude as a fixed effect covariate, distinguishing between lowland zones, where thermal inversion tends to trap pollutants near the surface, and mountainous areas, where cleaner air and stronger circulation mitigate pollution levels. Elevation data are obtained from a Digital Elevation Model provided by the National Institute of Geophysics and Vulcanology (Istituto Nazionale di Geofisica e Vulcanologia (INGV), 2024), and are displayed in the central panel of Figure 6. Human activities such as traffic, residential heating, and industrial production also represent major sources of $NO_2$. These factors are closely related to population density, which is therefore included in the model as a second fixed effect covariate. Population density data, shown in the right panel of Figure 6, are provided by the Istituto Nazionale di Statistica and sourced from Regione Lombardia (2024).

A third and final component we consider in modeling $NO_2$ concentrations is wind dynamics, which drive the transport of the pollutant. For this reason, we regularize the nonparametric component $f$ through an advection-diffusion PDE, which combines an isotropic diffusion term, modeling the dispersion of the pollutant in the air, with a transport term, accounting for advection

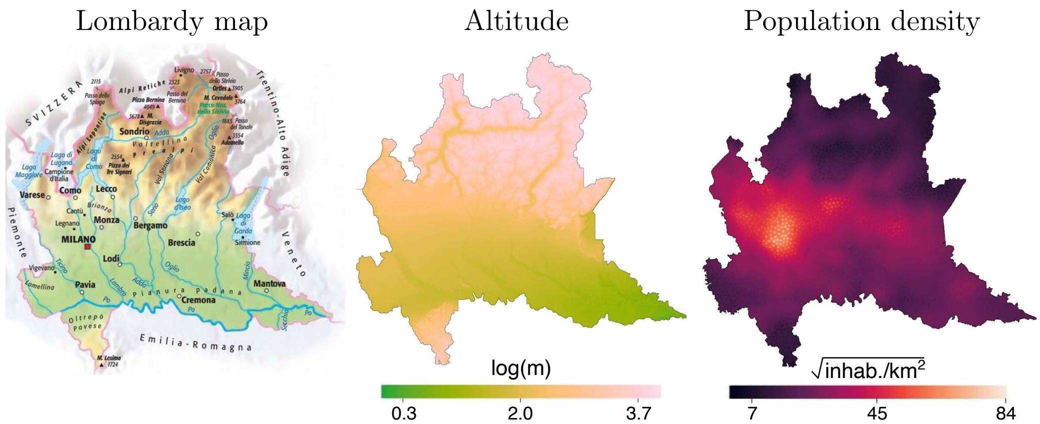


Figure 6: Left panel: geographical map of the Lombardy region (source: cartinadatieuropa.it). Central panel: logarithm of the altitude data derived from the Digital Elevation Model. Right panel: square root of the population density computed from ISTAT census data.

driven by the wind field. The resulting spatial penalty is given by

$$\mathcal{P}_{\mathscr{D}}(f) = \int_0^T \int_{\mathscr{D}} \big( -\nabla \cdot (I \nabla f(\mathbf{p}, t)) + \xi \, \boldsymbol{\gamma}(\mathbf{p}) \cdot \nabla f(\mathbf{p}, t) \big)^2 \, d\mathbf{p} \, dt, \tag{13}$$

where $\boldsymbol{\gamma}$ is the space-varying daily average wind vector, shown in Figure 2, obtained from data at 119 ARPA meteorological stations distributed across the region and collected on the same day as the $NO_2$ measurements. The unknown parameter $\xi$, which balances the relative contribution of the diffusion term $-\nabla \cdot (I \nabla f)$ and the transport term $\boldsymbol{\gamma} \cdot \nabla f$, is estimated from the data via a parameter cascading strategy, as described in Tomasetto et al. (2024). It is important to note that the transport field $\boldsymbol{\gamma}(\mathbf{p})$ considered in the present framework is spatially varying but assumed to be time-invariant. This assumption is appropriate when analyzing intra-day dynamics, as the wind effect can be reasonably approximated by its daily average. However, it becomes restrictive when extending the analysis to multiple days, since wind conditions may vary substantially over time and cannot be consistently represented by a single transport field. Extending the model to incorporate time-varying transport dynamics is an active direction of ongoing research, as discussed in the final section.

Using the mixed-effects model with the specifications described above, we estimate the spatio-temporal concentration levels of $NO_2$ over the Lombardy region. The results are shown in Figure 7 for three representative hours of the day, namely 08:00, 16:00, and 21:00. The estimated maps highlight the typical diurnal cycle of $NO_2$. In the morning, concentrations increase markedly due to traffic and domestic activities, with pronounced peaks in major urban centers such as Milano and Brescia. In the afternoon (16:00), levels decrease slightly, reflecting reduced mobility during working hours and a temporary stabilization of emissions. Nevertheless, urban hotspots remain clearly visible, indicating the persistent impact of $NO_2$ in metropolitan areas. In the evening (21:00), elevated concentrations are observed across much of the Po Valley, as limited air circulation favors the accumulation of pollutants over the course of the day, particularly in lowland and urban areas, while mountainous regions remain less affected. These patterns are consistent with the expected behavior of $NO_2$ under typical urban and meteorological conditions.

Moreover, contrary to what would be expected under a regular diurnal cycle, $NO_2$ concentrations at 00:00 are substantially lower than those observed at the end of the day. This behavior can be attributed to strong wind conditions recorded in the days preceding the observations, namely the 13$^{\text{th}}$ and 14$^{\text{th}}$ of January. In particular, a weather alert was issued on the 13$^{\text{th}}$, when wind gusts reached up to 70 km/h, which is highly unusual for the region. Finally, we observe that the area surrounding the city of Sondrio, together with the mountain towns in the northwestern part of Lombardy, consistently exhibits lower $NO_2$ concentrations throughout the day. This pattern can be explained by the Alpine morphology of the region. Indeed, low population density and limited anthropogenic activities reduce local emissions, while stronger wind circulation, characteristic of mountainous areas, enhances pollutant dispersion.

To illustrate the impact of the random effects on the estimated fields, we analyze the temporal evolution of $NO_2$ concentrations at a fixed location, corresponding to the city of Milano. Figure 8 shows the estimated hourly average profile (dashed black line), together with the sensor-specific curves associated with the 4 technological categories considered in this study (solid colored lines). The average curve reproduces the main diurnal pattern, characterized by a pronounced morning peak, a moderate decline in the afternoon, and persistent pollutant levels in the evening, consistent with the spatial patterns previously observed. The sensor-specific curves highlight the effect of the measurement technology on the estimated signal. The correction associated with the measurement technology, captured by the random component, is modest in magnitude, as expected for well-calibrated sensors. However,

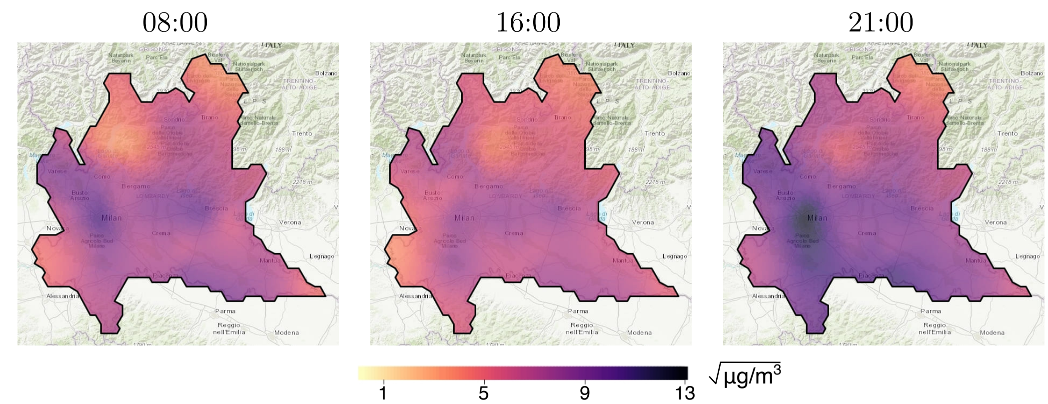


Figure 7: Estimated spatial fields at three representative hours of the day: 08:00 (left panel), 16:00 (central panel), and 21:00 (right panel).

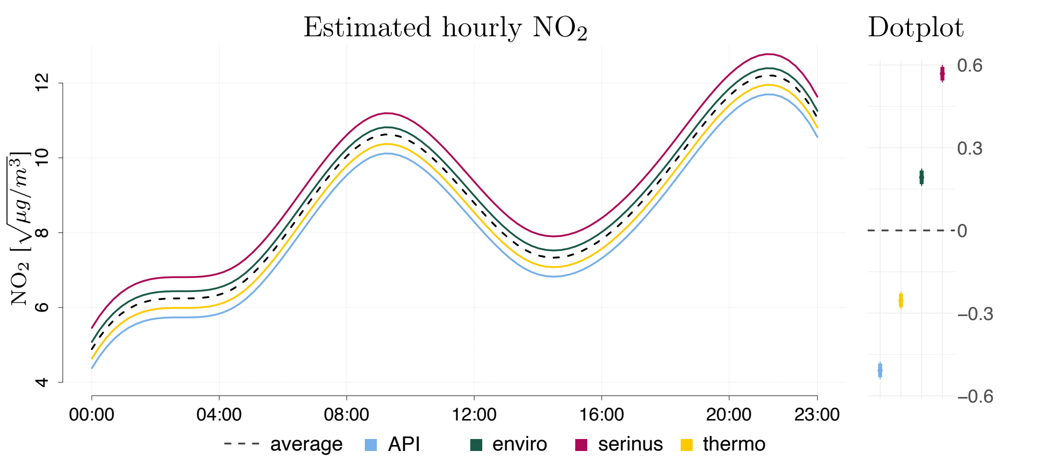


Figure 8: Left panel: estimated hourly $NO_2$ time profile for the city of Milano (dashed black curve), together with the effect of the measurement-specific technologies on the average signal (solid colored curves). Right panel: estimated 99% confidence bands for the measurement-specific random intercepts. The intervals are ordered according to the sign of the pointwise estimates $\hat{\mathbf{b}}_k$.

the 99% confidence bands of the estimated $b_k$ in the right panel of Figure 8 indicate that these sensor-specific effects, although small, are significantly different from zero. Modeling sensor technology as a random effect allows us to estimate the variance $\sigma^2_{\mathbf{b}}$ and directly quantify the portion of variability attributable to the measurement devices. This is summarized by the ratio $\text{PVRE} = \hat{\sigma}^2_{\mathbf{b}}/(\hat{\sigma}^2_{\mathbf{b}} + \hat{\sigma}^2) = 13.52\%$, confirming that the random component captures a non-negligible share of the total variability. In summary, the proposed model enables a clear separation between the spatio-temporal field, which reflects the underlying pollutant dynamics, and the random component, which accounts for systematic deviations due to sensor characteristics.

Section S3 reports a leave-one-station-out comparison of the proposed approach against `TPS` and `SOAP`, demonstrating the higher accuracy of `MEST-PDE`, for instance in terms of the Root Mean Squared Error.

## 7. Discussion and further developments

In this work, we have introduced a novel physics-informed semiparametric mixed-effects model for the analysis of spatio-temporal data. The proposed framework extends the PDE-regularized regression approach of Sangalli et al. (2013) and Azzimonti et al. (2014, 2015) by incorporating random effects into the statistical model. The inclusion of random effects broadens the applicability of the framework to settings with grouping structures, enabling a flexible representation of group-specific variability.

This work can be extended in several directions. First, one could consider introducing additional levels of random effects, extending the proposed framework to multilevel modeling in order to represent complex nested hierarchical structures, which are often encountered in real-world applications. Furthermore, extending the class of physics-informed penalties is also of interest. For instance, allowing for time-varying PDE parameters would enable the modeling of phenomena in which spatial anisotropy and stationarity evolve over time. In the case of $NO_2$ measurements discussed in Section 6, incorporating a time-varying wind field would allow the analysis to be extended to longer time periods. Such a generalization poses relevant theoretical, methodological, and computational challenges, which are currently under investigation. Another important direction concerns spatial confounding, which arises when covariates exhibit spatial autocorrelation. While recent work in semiparametric regression has addressed spatial confounding for fixed effect covariates (Dupont et al., 2022), analogous developments for mixed-effects semiparametric models of the type considered here are still lacking.

Additionally, from a computational perspective, the efficiency of the estimation procedure could be improved in the final stages of the optimization. Indeed, the EM algorithm typically converges rapidly in the initial iterations, but tends to slow down near convergence. A possible solution, proposed by Bates and Pinheiro (1998), consists of a hybrid strategy that combines EM iterations in the early phase of the algorithm with a Newton-type optimization in the final phase, thereby accelerating convergence without compromising stability. Finally, the proposed framework naturally extends to more complex spatial domains, including Riemannian manifolds (Ettinger et al., 2016; Lila et al., 2016), three-dimensional regions with complex geometries (Arnone et al., 2023a), and graphs (Clemente et al., 2026), where accounting for the domain geometry is essential for an accurate representation of spatial dependence.

**Acknowledgments**

We are grateful to the Editor and Reviewers for their insightful comments and suggestions, which have significantly improved the quality and clarity of the manuscript. We thank Alessandro Melchionda for his preliminary work on this topic in his Master Thesis at Politecnico di Milano. This work has been supported by the project GRINS - Growing Resilient, INclusive and Sustainable (GRINS PE00000018 – CUP D43C22003110001), funded by the European Union - NextGenerationEU programme. The views and opinions expressed are solely those of the authors and do not necessarily reflect those of the European Union, nor can the European Union be held responsible for them. M.F. De Sanctis, F. Ieva and L.M. Sangalli acknowledge MUR research project Dipartimento di Eccellenza 2023 - 2027, Dipartimento di Matematica, Politecnico di Milano. The simulations discussed in this work were performed on the HPC Cluster of the Department of Mathematics of Politecnico di Milano which was funded by MUR grant Dipartimento di Eccellenza 2023-2027. L.M. Sangalli also acknowledges the PRIN 2022 project CoEnv - Complex Environmental Data and Modeling (CUP 2022E3RY23), funded by the European Union – NextGenerationEU programme, and by the Italian Ministry for University and Research.

## References

Allard, D., Hristopulos, D.T., Opitz, T., 2021. Linking physics and spatial statistics: A new family of boltzmann-gibbs random fields. Electronic Journal of Statistics 15, 4085–4116.

Anderson, S.C., Ward, E.J., English, P.A., Barnett, L.A., 2022. sdmtmb: an R package for fast, flexible, and user-friendly generalized linear mixed effects models with spatial and spatiotemporal random fields. BioRxiv , 2022–03.

Arnone, E., Azzimonti, L., Nobile, F., Sangalli, L.M., 2019. Modeling spatially dependent functional data via regression with differential regularization. Journal of Multivariate Analysis 170, 275–295. doi:https://doi.org/10.1016/j.jmva.2018.09.006.

Arnone, E., Negri, L., Panzica, F., Sangalli, L.M., 2023a. Analyzing data in complicated 3d domains: Smoothing, semiparametric regression, and functional principal component analysis. Biometrics 79, 3510–3521. doi:https://doi.org/10.1111/biom.13845.

Arnone, E., Sangalli, L.M., Vicini, A., 2023b. Smoothing spatio-temporal data with complex missing data patterns. Statistical Modelling 23, 327–356. doi:https://doi.org/10.1177/1471082X211057959.

Augustin, N., Trenkel, V., Wood, S., Lorance, P., 2013. Space-time modelling of blue ling for fisheries stock management. Environmetrics 24, 109–119.

Azzimonti, L., Nobile, F., Sangalli, L.M., Secchi, P., 2014. Mixed finite elements for spatial regression with PDE penalization. SIAM/ASA Journal on Uncertainty Quantification 2, 305–335. doi:http://dx.doi.org/10.1137/130925426.

Azzimonti, L., Sangalli, L.M., Secchi, P., Domanin, M., Nobile, F., 2015. Blood flow velocity field estimation via spatial regression with PDE penalization. Journal of the American Statistical Association 110, 1057–1071. doi:https://doi.org/10.1080/01621459.2014.946036.

Bates, D., Mächler, M., Bolker, B., Walker, S., 2015. Fitting linear mixed-effects models using lme4. Journal of Statistical Software 67, 1–48.

Bates, D.M., Pinheiro, J.C., 1998. Computational methods for multilevel modelling. University of Wisconsin, Madison, WI , 1–29.

Bernardi, M.S., Carey, M., Ramsay, J.O., Sangalli, L.M., 2018. Modeling spatial anisotropy via regression with partial differential regularization. Journal of Multivariate Analysis 167, 15–30. doi:https://doi.org/10.1016/j.jmva.2018.03.014.

Bernardi, M.S., Sangalli, L.M., Mazza, G., Ramsay, J.O., 2017. A penalized regression model for spatial functional data with application to the analysis of the production of waste in venice province. Stochastic environmental research and risk assessment 31, 23–38. doi:https://doi.org/10.1007/s00477-016-1237-3.

California Air Resources Board, 2023. California environmental protection agency. URL: https://ww2.arb.ca.gov/.

Carrizo Vergara, R., Allard, D., Desassis, N., 2022. A general framework for spde-based stationary random fields. Bernoulli 28, 1–32.

Castiglione, C., Arnone, E., Bernardi, M., Farcomeni, A., Sangalli, L., 2025. PDE-regularised spatial quantile regression. Journal of Multivariate Analysis 205. doi:https://doi.org/10.1016/j.jmva.2024.105381.

Chen, X., Qi, L., Li, S., Duan, X., 2024. Long-term $NO_2$ exposure and mortality: a comprehensive meta-analysis. Environmental Pollution 341, 122971.

Clarotto, L., Allard, D., Romary, T., Desassis, N., 2024. The spde approach for spatio-temporal datasets with advection and diffusion. Spatial Statistics 62, 100847.

Clemente, A., Arnone, E., Mateu, J., Sangalli, L.M., 2026. Nonparametric estimators over metric graphs. Biometrika , asag029doi:https://doi.org/10.1093/biomet/asag029.

Craven, P., Wahba, G., 1978. Smoothing noisy data with spline functions: estimating the correct degree of smoothing by the method of generalized cross-validation. Numerische mathematik 31, 377–403.

De Sanctis, M.F., Di Battista, I., Arnone, E., Castiglione, C., Palummo, A., Bernardi, M., Ieva, F., Sangalli, L.M., 2025. Exploring nitrogen dioxide spatial concentration via physics-informed multiple quantile regression. Environmental and Ecological Statistics, 32, 855–892. doi:https://doi.org/10.1007/s10651-025-00664-8.

Di Brisco, A.M., Migliorati, S., 2021. A spatial mixed-effects regression model for electoral data. Statistical Methods & Applications 30, 543–571.

Dupont, E., Wood, S.N., Augustin, N.H., 2022. Spatial+: a novel approach to spatial confounding. Biometrics 78, 1279–1290.

Ettinger, B., Perotto, S., Sangalli, L.M., 2016. Spatial regression models over two-dimensional manifolds. Biometrika 103, 71–88. doi:https://doi.org/10.1093/biomet/asv069.

Gałecki, A., Burzykowski, T., 2012. Linear mixed-effects model, in: Linear mixed-effects models using R: a step-by-step approach. Springer, pp. 245–273.

Hefley, T.J., Hooten, M.B., Russell, R.E., Walsh, D.P., Powell, J.A., 2017. When mechanism matters: Bayesian forecasting using models of ecological diffusion. Ecology Letters 20, 640–650.

Istituto Nazionale di Geofisica e Vulcanologia (INGV), 2024. National Institute of Geophysics and Volcanology. URL: https://www.ingv.it/.

Karcher, P., Wang, Y., 2001. Generalized nonparametric mixed effects models. Journal of Computational and Graphical Statistics 10, 641–655.

Khan, K., Calder, C.A., 2022. Restricted spatial regression methods: Implications for inference. Journal of the American Statistical Association 117, 482–494.

Lila, E., Aston, J.A., Sangalli, L.M., 2016. Smooth principal component analysis over two-dimensional manifolds with an application to neuroimaging. The Annals of Applied Statistics doi:10.1214/16-AOAS975.

Lin, X., Zhang, D., 1999. Inference in generalized additive mixed models by using smoothing splines. Journal of the Royal Statistical Society Series B: Statistical Methodology 61, 381–400.

Lindgren, F., Bolin, D., Rue, H., 2022. The SPDE approach for gaussian and non-gaussian fields: 10 years and still running. Spatial Statistics 50, 100599.

Lindgren, F., Rue, H., Lindst, J., 2011. An explicit link between gaussian fields and gaussian markov random fields: The SPDE approach. Journal of the Royal Statistical Society. Series B. Statistical Methodology 73.

Marra, G., Miller, D.L., Zanin, L., 2012. Modelling the spatiotemporal distribution of the incidence of resident foreign population. Statistica Neerlandica 66, 133–160.

da Matta, D.H., Motta, M.R., Garcia, N.L., Heinemann, A.B., 2025. A bayesian spatial-temporal functional model for data with block structure and repeated measures. arXiv preprint arXiv:2501.01269 .

Mullen, R.S., Birkeland, K.W., 2008. Mixed effect and spatial correlation models for analyzing a regional spatial dataset, in: Proceedings of the 2008 International Snow Science Workshop, Whistler, British Columbia, pp. 421–425.

Open Data Lombardia, Trasformazione Digitale in Lombardia, 2024. Regione Lombardia. URL: https://www.dati.lombardia.it/stories/s/Meteo-inquinamento-aria-e-acqua-e-altri-dati-da-AR/auv9-c2sj.

O'sullivan, F., Yandell, B.S., Raynor Jr, W.J., 1986. Automatic smoothing of regression functions in generalized linear models. Journal of the American Statistical Association 81, 96–103.

Palummo, A., Arnone, E., Clemente, A., Sangalli, L.M., Ramsay, J., Formaggia, L., 2025. fdaPDE: Physics-informed spatial and functional data analysis. GitHub, https://github.com/fdaPDE/fdaPDE.

Pereira, M., Desassis, N., Allard, D., 2022. Geostatistics for large datasets on riemannian manifolds: a matrix-free approach. arXiv preprint arXiv:2208.12501 .

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., Heisterkamp, S., Van Willigen, B., Maintainer, R., 2017. Package 'nlme'. Linear and nonlinear mixed effects models, version 3, 274.

Pinheiro, J.C., Bates, D.M., 2000. Mixed-effects models in S and S-PLUS. Springer.

Ramsay, J.O., Hooker, G., Campbell, D., Cao, J., 2007. Parameter estimation for differential equations: a generalized smoothing approach. Journal of the Royal Statistical Society. Series B. Statistical Methodology 69, 741–796.

Regione Lombardia, 2024. Geoportale lombardia. URL: https://www.geoportale.regione.lombardia.it/.

Ribeiro Jr, P.J., Diggle, P.J., 2025. geoR: Analysis of Geostatistical Data. URL: https://CRAN.R-project.org/package=geoR, doi:10.32614/CRAN.package.geoR. R package version 1.9-6.

Richardson, R.A., 2017. Sparsity in nonlinear dynamic spatiotemporal models using implied advection. Environmetrics 28, e2456.

Rue, H., Martino, S., Chopin, N., 2009. Approximate bayesian inference for latent gaussian models by using integrated nested laplace approximations. Journal of the Royal Statistical Society Series B: Statistical Methodology 71, 319–392.

Sahu, S.K., Gelfand, A.E., Holland, D.M., 2006. Spatio-temporal modeling of fine particulate matter. Journal of Agricultural, Biological, and Environmental Statistics 11, 61–86.

Salama, D.S., Yousif, M., Gedamy, Y., Ahmed, H.M., Ali, M., Shoukry, E.M., 2022. Satellite observations for monitoring atmospheric $NO_2$ in correlation with the existing pollution sources under arid environment. Modeling Earth Systems and Environment 8, 4103–4121.

Sangalli, L.M., 2021. Spatial regression with partial differential equation regularisation. International Statistical Review 89, 505–531. doi:https://doi.org/10.1111/insr.12444.

Sangalli, L.M., Ramsay, J.O., Ramsay, T.O., 2013. Spatial spline regression models. Journal of the Royal Statistical Society. Series B. Statistical Methodology 75, 681–703. doi:https://doi.org/10.1111/rssb.12009.

Sigrist, F., Künsch, H.R., Stahel, W.A., 2015. spate: An R package for spatio-temporal modeling with a stochastic advection-diffusion process. Journal of Statistical Software 63, 1–23. doi:10.1111/rssb.12061.

Smith, R.L., Kolenikov, S., Cox, L.H., 2003. Spatiotemporal modeling of $PM_{2.5}$ data with missing values. Journal of Geophysical Research: Atmospheres 108.

Tomasetto, M., Arnone, E., Sangalli, L.M., 2024. Modeling anisotropy and non-stationarity through physics-informed spatial regression. Environmetrics doi:https://doi.org/10.1002/env.2889.

Trinh, T.T., Trinh, T.T., Le, T.T., Nguyen, T.D.H., Tu, B.M., 2019. Temperature inversion and air pollution relationship, and its effects on human health in hanoi city, vietnam. Environmental geochemistry and health 41, 929–937.

Wahba, G., 1985. A comparison of GCV and GML for choosing the smoothing parameter in the generalized spline smoothing problem. The annals of statistics , 1378–1402.

Wahba, G., 1990. Spline Models for Observational Data. Society for Industrial and Applied Mathematics.

Wikle, C.K., Hooten, M., 2010. A general science-based framework for dynamical spatio-temporal models. TEST 19, 417–451.

Wilhelm, M., Sangalli, L.M., 2016. Generalized spatial regression with differential regularization. Journal of Statistical Computation and Simulation 86, 2497–2518. doi:https://doi.org/10.1080/00949655.2016.1182532.

Wood, S., Scheipl, F., Wood, M.S., 2017. Package 'gamm4'. Am Stat 45, 0–2.

Wood, S., Wood, M.S., 2015. Package 'mgcv'. R package version 1, 729.

Wood, S.N., 2006. Low-rank scale-invariant tensor product smooths for generalized additive mixed models. Biometrics 62, 1025–1036.

Wood, S.N., 2017. Generalized additive models: an introduction with R. chapman and hall/CRC.

Wood, S.N., Bravington, M.V., Hedley, S.L., 2008. Soap film smoothing. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 70, 931–955.

World Health Organization, 2024. Ambient (outdoor) air pollution 2024. https://www.who.int/news-room/fact-sheets/detail/ambient-%28outdoor%29-air-quality-and-health.

Xun, X., Cao, J., Mallick, B., Maity, A., Carroll, R.J., 2013. Parameter estimation of partial differential equation models. Journal of the American Statistical Association 108, 1009–1020.

Yanosky, J.D., Paciorek, C.J., Laden, F., Hart, J.E., Puett, R.C., Liao, D., Suh, H.H., 2014. Spatio-temporal modeling of particulate air pollution in the conterminous united states using geographic and meteorological predictors. Environmental Health 13, 63.

## MOX Technical Reports, last issues

Dipartimento di Matematica
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

**64/2026** Radisic, I.; Tirotta, R.; Zingaro, A.; Pagani, S.; Dede', L.
*Physics-Informed Neural Networks for the High-Resolution Reconstruction of Flow Measurement Indicators in Fluid Dynamics*

**62/2026** Ciaramella, G.; Gong, W.; Kwok, F.; Tan, Z.
*On the Uniform Convergence Analysis of the Schwarz Alternating Method for Optimal Control Problems*

**63/2026** Ciaramella, G.; Gong, W.; Kwok, F., Tan, Z.
*Uniform Convergence of the Schwarz Alternating Method for Optimal Control Problems*

**60/2026** Ciaramella, G.; Gander, M.J.; Van Criekingen, S.; Vanzan, T.
*Algebraic and Two-Level Parallel Substructured Schwarz Methods*

**58/2026** Mapelli, A.; Massi, M.C.; Cuccuru, G.; Di Angelantonio, E.; Ieva, F.
*Prior-informed conditional Gaussian graphical models: an application to protein interaction network reconstruction*

**57/2026** Fontana, N.; Secchi, P.; Di Angelantonio, E.; Ieva, F.
*Modeling time-varying genetic effects on binary disease risk via functional Mendelian Randomization*

**56/2026** Botta, P.; Vitullo, P.; Ventimiglia, T.; Linninger, A.; Zunino, P.
*Physics-Informed Learning of Microvascular Flow Models using Graph Neural Networks*

**55/2026** Beirao da Veiga, L.; Canuto, C.; Nochetto, R.H.; Vacc, G ; Verani, M.
*A Virtual Element Method for elliptic problems on trimmed background meshes*

**54/2026** Antonietti, P. F.; Corti, M.; Leimer Saglio, C. B.; Pagani, S.
*The lymph 2.0 library: p-adaptive algorithms and parallel assembly strategies for polytopal DG methods*

**51/2026** Bellezza P.; Ciaramella G.; Macchini C.; Mazzieri I.; Verani M.
*ParaFlow: Parareal Acceleration of Gradient-Flow Minimization*