Spatial Statistics 66 (2025) 100885

Contents lists available at ScienceDirect

Spatial Statistics

journal homepage: www.elsevier.com/locate/spasta

Derivative-based spatial mediation with INLA-SPDE

Claudio Rubino [a], Chiara Di Maria [a,*], Antonino Abbruzzo [a], Gioacchino Bono [b], Germana Garofalo [b], Giacomo Milisenda [c], Giada Adelfio [a]

[a] Department of Economics, Business and Statistics, University of Palermo, Viale delle Scienze, Building 13, Palermo, 90128, Italy
[b] Institute for Marine Biological Resources and Biotechnology, National Research Council, Via Vaccara 61, Mazara del Vallo, Trapani, 91026, Italy
[c] Integrative Marine Ecology Department, Zoological Station Anton Dohrn, Lungomare Cristoforo Colombo (complesso Roosevelt), Palermo, 90149, Italy

ARTICLE INFO

Keywords:
Geostatistical data
Mediation analysis
INLA-SPDE
Spatial correlation

ABSTRACT

In many applied fields, it may be of interest to evaluate mediational mechanisms occurring in spatial domains. The approaches proposed so far in the literature to address this issue deal with areal data and often consider linear models. In this paper, we propose an approach to assess mediation in the presence of geostatistical data by combining the integrated nested Laplace approximation (INLA) with a derivative-based approach for mediation analysis, which allows one to estimate indirect effects also in the case of nonlinear models. We investigate the effect of ignoring spatial processes in the mediator and the outcome models through a simulation study, focusing also on the case of correlated processes. To show the usefulness of our approach, we also provided an ecological application.

1. Introduction

Mediation analysis is commonly used in various disciplines like biology, social sciences and epidemiology to assess the indirect effect of an exposure on an outcome through an intermediate variable called *mediator*. In those fields, data with inherent spatial structure is encountered quite frequently. The presence of spatial correlation, which may occur at the level of either the mediator or the outcome, acts as an unmeasured confounder. In an associative mediational setting, this can introduce bias in the estimation of mediational effects. Furthermore, within a causal framework, this violates the foundational assumptions of causal models, preventing their identifiability, i.e. the ability to express them in terms of observed data (VanderWeele, 2015; Pearl, 2001).

Recently, some studies have combined mediation and spatial analyses in different ways. Wang et al. (2023) and Lu and Zhu (2022) analyzed the direct and indirect effects of some spatial variables like the presence of infrastructures or digital innovation, on economic development using a Durbin spatial model based on spatial weight matrices. Bustamante et al. (2023) investigated the mediating factors between area-level poverty and *Staphilococcus aureus* infection in California using a multilevel model with spatial-dependent random effects, while Mougeni et al. (2024) modeled malaria prevalence among 2-to −10 years old children in Gabon using different spatial models. Smith et al. (2023) and Jarvis et al. (2019) provided a decomposition of spatial effects into direct and indirect effects in the context of cluster randomized trials.

All these approaches deal with areal data. In addition, the indirect effect is not always formally defined and, when it is, it relies on the classical product or difference methods (MacKinnon, 2008; Baron and Kenny, 1986), which, however, cannot be applied in the case of more complex models with link function different from identity. Indeed, the literature about mediation analysis in linear

* Corresponding author.
E-mail address: chiara.dimaria@unipa.it (C. Di Maria).

https://doi.org/10.1016/j.spasta.2025.100885
Received 23 October 2024; Received in revised form 23 January 2025; Accepted 7 February 2025

Available online 24 February 2025
2211-6753/© 2025 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

settings is as rich as the one about nonlinear mediation is piecewise and fragmented. This is mainly due to the fact that the path analytic approach (Wright, 1934) on which the product method relies cannot be straightforwardly generalized to nonlinear models, although some attempts have been made (Tsai et al., 2006; Schluchter, 2008).

However, the presence of mediational mechanisms in geostatistical data is frequent and relevant as well. For example, suppose to be interested in the effect of wind speed on asthma prevalence in children, mediated by air pollution. Nonetheless, literature currently lacks of approaches to deal with mediation in the case of geostatistical data. The integrated nested Laplace approximation (INLA, Rue et al., 2009), combined with the Stochastic Partial Differential Equation (SPDE) proved to be quite convenient to model this kind of data, since it allows for great flexibility in model specification and reduces the computational costs associated to the estimation of complex spatial models, as will be detailed in the next section.

The INLA-SPDE approach has been applied to geostatistical data in several fields: in epidemiology, it was used to analyze the risk of transmission of infectious diseases like malaria in Mozambique (Moraga et al., 2021), the dengue fever in Malaysia (Naeeim and Rahman, 2017), or animal tubercolosis in Spain (Gortázar et al., 2017). Environmental applications are common as well: in the context of pollution mapping, Cameletti et al. (2013) considered a hierarchical spatio-temporal model for particulate matter (PM) concentration in a North-Italian region, while, moving to ecology, Cosandey-Godin et al. (2014) analyzed point-referenced data collected in the Baffin bay to evaluate the extent of Greenland shark bycatch, and Lezama-Ochoa et al. (2020) used INLA-SPDE to model the occurrence of *Mobula mobular* in the Pacific Ocean. As regards atmospheric sciences, Lindgren et al. (2011) discuss several examples of applications of INLA in complex, non-standard settings, using, among others, data on temperatures from 7280 stations located in all continents, to show the use of non-stationary SPDE, and Fioravanti et al. (2023) proposed a method to interpolate climate variables, applying their approach to temperatures recorded in Italy from 1961 to 2020. A comprehensive review of INLA-SPDE and plenty of other examples can be found in Bakka et al. (2018).

Despite the wide range of applications of INLA-SPDE, so far, it has not been applied to mediational settings, although a recent stream of literature proposed to embed mediation analysis into a Bayesian framework to make it more flexible (Yuan and MacKinnon, 2009). Indeed, many scholars highlight the superiority of the Bayesian framework to address different issues of mediation analysis, from the traditional problem concerning the asymmetry of indirect effects, and the consequent difficulty in estimating confidence intervals (Miočević et al., 2017), to more complex settings involving latent variables (Miočević, 2019; Sun et al., 2021) or dynamic processes (Huang and Yuan, 2017). We claim that Bayesian mediation analysis can be applied also in complex spatial settings through INLA-SPDE.

The aim of this paper is to extend spatial mediation to settings with geo-referenced data possibly requiring modeling strategies involving nonlinearities. To do so, we combine the Integrated Nested Laplace Approximation (INLA), and an approach to mediation analysis based on derivatives (Stolzenberg, 1980; Geldhof et al., 2018) that allows us to estimate indirect effects also for nonlinear models. To the best of our knowledge, this is the first paper formally defining mediation analysis when the variables involved vary in space.

The paper is organized as follows. First, we briefly describe INLA-SPDE and the derivative-based approach to mediation and show how they can be combined. Then we show the results of a simulation study to assess the impact of ignoring the spatial component in a mediating setting with geostatistical data. In the last section, we present an application from ecology. Conclusions will follow.

2. Methodology

In this section, we provide the basis of our proposal by first describing the main concepts on which INLA and the derivative-based approach rely and then combining them in the context of spatial data.

2.1. INLA-SPDE

The integrated nested Laplace approximation (INLA) has been proposed by Rue et al. (2009), and further developed by Martins et al. (2013) as an alternative to Markov Chain Monte Carlo (MCMC) approaches in the context of Bayesian inference. A model can be estimated with INLA if the expected value μ of a random variable Y can be modeled through a link function $g(\cdot)$ in an additive way as

$$g(\mu_i) = \eta_i = \beta_0 + \sum_{m=1}^{M} \beta_m x_{mi} + \sum_{l=1}^{L} f_l(z_{li}), \tag{1}$$

where β_0 is the intercept, the coefficients $\boldsymbol{\beta} = \{\beta_1, \dots, \beta_M\}$ quantify the effect of the explanatory variables $\boldsymbol{x}$ on the response variable y_i and $\boldsymbol{f} = \{f_1(\cdot), \dots, f_L(\cdot)\}$ is a collection of unknown functions defined on a set of covariates $\boldsymbol{z}$, which generally express spatial or temporal dependencies between observations.

Focusing on spatial data, we will consider them as realizations of a stochastic process indexed in space (Cressie, 1993) $\{Y(\boldsymbol{s}) : \boldsymbol{s} \in D\}$, where $\boldsymbol{s}$ is the vector of spatial coordinates associated to Y in a d-dimensional euclidean space and D is the observed portion of space, with $D \in \mathbb{R}^d$. In the case of $d = 2$, generally, $\boldsymbol{s}$ contains latitude and longitude of Y. In particular, in the geostatistical approach (see, e.g., Cressie and Wickle, 2011) point-referenced data consist of a set of measurements (realizations) of Y taken at a finite set of points $(s_1, \dots, s_n)$ in D, and the spatial index $\boldsymbol{s}$ can take any value in the continuum in D.

A useful approach for the estimation of a geostatistical model is to assume that there is a spatially continuous variable underlying the observations that can be modeled using a Gaussian random field (GRF) $U(s)$, which is a random function for which it holds that, for every finite set of points $(s_1, \dots, s_n)$ (Abrahamsen, 1997),

$$\boldsymbol{u} \sim \mathcal{N}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \tag{2}$$

where $\boldsymbol{u} = \{u(s_1), \dots, u(s_n)\}$ is a realization of $U(s)$ at n locations, and $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are the mean vector and the covariance matrix of the process, respectively.

The GRF incorporates the correlation structure of the process by means of its covariance matrix $\boldsymbol{\Sigma} = (\Sigma_{i,j})$, $i, j = (1, \dots, n)$, which is constructed from a covariance function. A common choice for the specification of the covariance function, which is of main interest here, is the Matérn function (Matérn, 1960), which implies that each single element Σ_{ij} of the covariance matrix $\boldsymbol{\Sigma}$ is defined as

$$\Sigma_{ij} = \mathrm{Cov}_M(u(s_i), u(s_j)) = \frac{\sigma_u^2}{\Gamma(\nu)2^{\nu-1}}(k\|s_i - s_j\|)^\nu K_\nu(k\|s_i - s_j\|), \tag{3}$$

where σ_u^2 is the marginal variance of the process, $\nu > 0$ is the smoothing parameter, $k > 0$ is a scale parameter, $\|s_i - s_j\|$ is the euclidean distance between s_i and s_j and K_ν is the modified Bessel function of second kind and order $\nu > 0$.

Although the use of GRFs proves convenient due to their good analytical properties, parameter estimation is often problematic in practice, especially with large data sets. Indeed, inference on the parameters of Eq. (3) has a computational cost equal to $\mathcal{O}(n^3)$, as it requires factoring fully dense $n \times n$ covariance matrices (Lindgren et al., 2011). Furthermore, the fitting of a model in a Bayesian inferential paradigm is traditionally based on MCMC algorithms, which require these calculations at each iteration, making the task even more difficult (in this regard, see Banerjee et al., 2003, p. 387). Lindgren et al. (2011) developed a Bayesian approach alternative to MCMC, which proved to be more computationally efficient. The method consists of representing a GRF, which is a continuous spatial process, by approximating it to a spatial process with a discrete index, i.e. a Gaussian Markov Random Field (GMRF). Thus, thanks to the sparsity of the precision matrix of such a GMRF, which is induced by the conditional independence structure of the process, appropriate computation techniques for sparse matrices can be used (for an extensive description, see Rue and Held, 2005).

This approach relies primarily on the following two results. First, a stationary GRF $U(s)$ with zero mean and Matérn covariance function, with $\nu > 0$ and $\alpha = \nu + d/2$ integer, is the exact stationary solution of the following stochastic partial differential equation (SPDE) (Wittle, 1954):

$$(k^2 - \Delta)^{\alpha/2}(\tau_u u(s)) = W(s), \tag{4}$$

where $s \in \mathbb{R}^d$, $\Delta = \sum_{i=1}^d \partial^2/\partial s_i^2$ is the Laplacian operator, α is a smoothing parameter, k is the scaling parameter of the Matérn function, τ_u controls the variance and $W(s)$ is a Gaussian *white noise* spatial process. The connection between Eq. (4) and the parameters of the Matérn function is as follows:

$$\nu = \alpha - d/2 \tag{5}$$

$$\sigma_u^2 = \frac{\Gamma(\nu)}{\Gamma(\alpha)(4\pi)^{d/2}\, k^{2\nu}\, \tau_u^2} \propto \frac{1}{\tau_u^2}. \tag{6}$$

Second, an approximate solution of Eq. (4) an be obtained using interpolation through the finite element method (FEM) (Ciarlet, 1978). In brief, this method involves dividing the domain into a collection of non-overlapping, and possibly irregular triangles, which can vary in shape and size. The vertices of these triangles, referred to as nodes, serve as the basis points for interpolation. Each node corresponds to a basis function, and the solution of the SPDE, along with its properties, is influenced by the choice of these basis functions. The approximation is as follows:

$$\tilde{u}(s) = \sum_{j=1}^m \phi_j(s) w_j, \tag{7}$$

where m is the number of vertices, ϕ_j are deterministic functions (bases) and w_j are the weights to be estimated. The bases are such that $\phi_j = 1$ at vertex j and $\phi_j = 0$ at the other vertices. The discretized representation of the domain created by these triangles is called *mesh*. Eq. (7) can be rewritten as

$$\tilde{u}(s) = \sum_{j=1}^m A_j(s) w_j, \tag{8}$$

where $A_j(s) = \phi_j(s)$ is the generic element of the sparse matrix $\boldsymbol{A}_{n\times m}$ which maps the GMRF $\tilde{\boldsymbol{u}}$ from the m triangulation vertices to the n observation locations. It can be demonstrated that weights $\boldsymbol{w}$ in Eq. (7) can be assigned an a priori distribution such that $\tilde{u}(s) \approx u(s)$ in distribution. Under this formulation, $\tilde{u}(s)$ is a GMRF with distribution $\mathcal{N}(\boldsymbol{0}, \boldsymbol{Q}_{\boldsymbol{w}}^{-1}(\tau_u, k))$. Since $\boldsymbol{Q}_{\boldsymbol{w}}$ is a sparse matrix, the hyperparameters τ_u and k can be effectively estimated, within the Bayesian framework, using INLA (Rue et al., 2009). Generally, for better interpretability, the estimation results are expressed in terms of σ_u^2 and the range parameter r, i.e. the distance such that the spatial correlation between two points is very small (about 0.14), which is defined empirically as

$$r = \frac{\sqrt{8\nu}}{k}. \tag{9}$$

In $\mathbb{R}^2$, the computational cost of using a GMRF model is typically $O(n^{3/2})$ (Lindgren et al., 2011), which represents a substantial improvement compared to the $O(n^3)$ cost associated with a full Gaussian Field (GF) model.

Table 1
Partial derivatives of spatial mediator and outcome models for identity, log and logit link functions.

Link function	$\frac{\partial \mu_Y}{\partial M}$	$\frac{\partial \mu_M}{\partial X}$
Identity	γ_2	β_1
Log	$\gamma_2 \exp(\gamma_0 + \gamma_1 X + \gamma_2 M + u_Y(\mathbf{s}))$	$\beta_1 \exp(\beta_0 + \beta_1 X + u_M(\mathbf{s}))$
Logit	$\frac{\gamma_2 \exp(\gamma_0 + \gamma_1 X + \gamma_2 M + u_Y(\mathbf{s}))}{(1+\exp(\gamma_0 + \gamma_1 X + \gamma_2 M + u_Y(\mathbf{s})))^2}$	$\frac{\beta_1 \exp(\beta_0 + \beta_1 X + u_M(\mathbf{s}))}{(1+\exp(\beta_0 + \beta_1 X + u_M(\mathbf{s})))^2}$

2.2. The derivative-based approach to mediation analysis

Let us consider the most basic mediational setting, involving just three variables: an exposure X, a mediator M and an outcome Y. The indirect effect of X on Y is the part of the total effect transmitted by M and, in the classical regression-based framework, it can be estimated using the product method (Baron and Kenny, 1986; MacKinnon, 2008). Indeed, if both the mediator and the outcome can be modeled using standard linear models without interactions, i.e.

$$\mathbb{E}[M|X] = \mu_M = \beta_0 + \beta_1 X \tag{10}$$

$$\mathbb{E}[Y|X, M] = \mu_Y = \gamma_0 + \gamma_1 X + \gamma_2 M, \tag{11}$$

the indirect effect is given by $\beta_1\gamma_2$, that is, the product of the coefficients lying on the paths connecting X to M and M to Y. Unfortunately, this approach cannot be extended to nonlinear models and, for this reason, different definitions of indirect effect have been proposed in the literature to overcome this limitation, although most of them are specific for the cases of binary or count mediator/outcome. In this paper, we adopt the approach originally proposed by Stolzenberg (1980) and applied to nonlinear and generalized linear models by Hayes and Preacher (2010) and Geldhof et al. (2018), respectively. The main idea is that an indirect effect can be seen as the variation in Y corresponding to a variation of X through M. In other words, an indirect effect can be expressed as the derivative of Y, seen as a composite function of M, which is in turn function of X. We can then write the indirect effect as

$$\frac{\partial \mu_Y}{\partial M}\frac{\partial \mu_M}{\partial X}, \tag{12}$$

i.e. the product of the derivative of the expectation of Y with respect to M and that of M with respect to X. In the trivial case of linear models in Eqs. (10)–(11), definition (12) coincides with the product method. However, when the expectations of M and Y are linked to their linear predictors through functions different from the identity, let us call them g_1 and g_2, respectively, then the indirect effect is not unique, and depends on the values of X and M (and potential covariates, if present). For this reason, Geldhof et al. (2018) proposed to call effect (12) *conditional indirect effect* (CIE), since its value is always conditional to that of other variables. See Geldhof et al. (2018) and Di Maria et al. (2024) for a broader discussion.

2.3. The spatial conditional indirect effect

In this section, we combine the approaches just discussed to deal with spatial mediation.

In estimating mediational effects using data observed in a spatial domain, it is reasonable to consider that spatial correlation can occur at the mediator level, at the outcome level, or both. Following the notation and general framework just introduced, and assuming the presence of spatial correlation at both the mediator and the outcome level, the mediation model expressed earlier in Eqs. (10) and (11) becomes

$$g_1(\mu_M(\mathbf{s})) = \beta_0 + \beta_1 X(\mathbf{s}) + u_M(\mathbf{s}) \tag{13}$$

$$g_2(\mu_Y(\mathbf{s})) = \gamma_0 + \gamma_1 X(\mathbf{s}) + \gamma_2 M(\mathbf{s}) + u_Y(\mathbf{s}) \tag{14}$$

where X, M and Y are the exposure, the mediator and the outcome, respectively, observed at location $\mathbf{s}$ in a spatial domain $D \in \mathbb{R}^d$, $u_M(\mathbf{s})$ and $u_Y(\mathbf{s})$ are two different spatial Gaussian processes which take care of the spatial correlation at the mediator and the outcome level, respectively, g_1 and g_2 are possibly non-linear link functions, connecting the conditional expectations of the mediator and the outcome to their linear predictors, and $h_k = g_k^{-1}$, $k \in \{1, 2\}$.

In order to compute the CIE according to the derivative-based approach, it is necessary to compute the partial derivatives of the mediator and the outcome model, as expressed in Eq. (12). Table 1 shows the partial derivatives of mediator and outcome spatial models, for three link functions of common choice, i.e. identity, log and logit. To simplify the notation, from now on we omit the dependence of X, M and Y on the spatial coordinates $\mathbf{s}$.

For example, in the case of a binary outcome and a count variable as a mediator, assuming that $g_1(\cdot) = \log(\cdot)$ and $g_2(\cdot) = \text{logit}(\cdot)$, the conditional indirect effect can be computed as

$$\beta_1 \exp(\beta_0 + \beta_1 X + u_M(\mathbf{s})) \times \frac{\gamma_2 \exp(\gamma_0 + \gamma_1 X + \gamma_2 M + u_Y(\mathbf{s}))}{(1+\exp(\gamma_0 + \gamma_1 X + \gamma_2 M + u_Y(\mathbf{s})))^2}. \tag{15}$$

An estimate of the indirect effect can then be obtained as the posterior mean of a sample drawn from the posterior distribution of the indirect effect, given by combining the posterior distributions of model parameters, estimated by INLA, according to the

Algorithm 1 Algorithm to implement the derivative-based spatial mediation with INLA-SPDE.

1. Fit the mediator and the outcome models in INLA
if no spatial correlation between M and Y is assumed **then**
  Estimate two different INLA-SPDE models, one for the mediator and one for the outcome
else if a non-null correlation is assumed **then**
  stack the data for the mediator and the outcome and estimate an INLA-SPDE joint model
end if
if At least one of the model have non-identity link function **then**
  Select locations $\boldsymbol{s}'$ where indirect effect will be calculated
end if
3. Derive the posterior distributions of the models' parameters
4. Derive the posterior distribution of the indirect effect using the posterior distributions obtained at step 3 and formulas in Table 1
if At least one of the model have non-identity link function **then**
  Derive the posterior distributions of u_M and u_Y at selected locations $\boldsymbol{s}'$ and include them in the mediator and outcome models
end if
5. Summarize the posterior distribution using, for example, the mean and obtain Bayesian credibility intervals

mathematical form assumed by the indirect effect. This novel approach not only accommodates nonlinear models, but also allows for spatial effect heterogeneity. Indeed, as can be seen from Table 1, when at least one of the mediator and the outcome models is nonlinear, the indirect effect is no longer constant, but varies in the region of interest. This is quite reasonable in real-world cases, where the effect of a variable can differ across areas according to their specific characteristics.

Another aspect worthy of consideration is that the spatial processes can be correlated. We assume that $\boldsymbol{u}_M$ and $\boldsymbol{u}_Y$ are jointly Gaussian distributed, i.e.

$$\boldsymbol{u}_M, \boldsymbol{u}_Y \sim \mathcal{N}(\boldsymbol{0}_{\boldsymbol{u}_M,\boldsymbol{u}_Y}, \boldsymbol{\Sigma}_{\boldsymbol{u}_M,\boldsymbol{u}_Y}), \tag{16}$$

where $\boldsymbol{0}_{\boldsymbol{u}_M,\boldsymbol{u}_Y}$ is a zero vector of expected values of length $2n$ and $\boldsymbol{\Sigma}_{\boldsymbol{u}_M,\boldsymbol{u}_Y}$ is a $2n \times 2n$ joint (block) covariance matrix. The joint covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{u}_M,\boldsymbol{u}_Y}$ is as follows:

$$\boldsymbol{\Sigma}_{\boldsymbol{u}_M,\boldsymbol{u}_Y} = \begin{bmatrix} \boldsymbol{\Sigma}_{\boldsymbol{u}_M} & \boldsymbol{\Sigma}_{\boldsymbol{u}_M,\boldsymbol{u}_Y} \\ \boldsymbol{\Sigma}^T_{\boldsymbol{u}_M,\boldsymbol{u}_Y} & \boldsymbol{\Sigma}_{\boldsymbol{u}_Y} \end{bmatrix}, \tag{17}$$

where $\boldsymbol{\Sigma}_{\boldsymbol{u}_M}$ and $\boldsymbol{\Sigma}_{\boldsymbol{u}_Y}$ are the covariance matrices of $\boldsymbol{u}_M$ and $\boldsymbol{u}_Y$, respectively, and the generic element of the extra-diagonal blocks, that is

$$\rho \times \sqrt{\mathrm{Cov}(u_M(\boldsymbol{s}_i), u_M(\boldsymbol{s}_j)) \times \mathrm{Cov}(u_Y(\boldsymbol{s}_i), u_Y(\boldsymbol{s}_j))}, \tag{18}$$

is the covariance between the two processes for locations $\boldsymbol{s}_i$ and $\boldsymbol{s}_j$, which is governed by the parameter ρ.

INLA cannot account for ρ directly, however, it is possible to write the model in such a way that it includes an additional parameter which takes into account the correlation between u_M and u_Y (Gómez-Rubio, 2020). Indeed, adopting a joint-modeling approach, where M and Y are stacked in a unique response variable, and writing

$$u_Y = u'_Y + \lambda u_M,$$

where u'_Y is another spatial process with null expectation which is independent by u_M, it is easy to prove that the covariance between the mediator and the outcome processes is

$$\begin{aligned}\mathrm{Cov}(u_M, u_Y) &= \mathbb{E}[u_M u_Y] - \mathbb{E}[u_M]\mathbb{E}[u_Y] = \mathbb{E}[u_M(u'_Y + \lambda u_M)] \\ &= \mathbb{E}[u_M u'_Y + \lambda u_M^2] = \mathbb{E}[u_M u'_Y] + \lambda \mathbb{E}[u_M^2] \\ &= \mathrm{Cov}(u_M u'_Y) + \lambda \sigma^2_{u_M},\end{aligned}$$

from which it immediately follows that

$$\rho = \frac{\mathrm{Cov}(u_M u'_Y) + \lambda \sigma^2_{u_M}}{\sigma_{u_M}\sigma_{u_Y}}.$$

In summary, the rationale underlying the proposed approach is to exploit the computational efficiency and flexibility of INLA-SPDE to model complex, possibly correlated, spatial processes and use the posterior distributions of the model parameters to estimate the indirect effect(s). Since the mediator and/or the outcome model can be nonlinear, deriving the indirect effect is not straightforward, and it is necessary to adopt an adequate approach, as the derivative-based one discussed in the previous section. The key steps to implement our proposal are shown in Algorithm 1.

Table 2
Comparison of performances in estimating the indirect effect including/excluding the GF, varying sample size and magnitude of the indirect effect when data are simulated from models including a spatial component.

n	Ind. eff.	GF			No GF		
		Coverage	avg. len.	RMSE	Coverage	avg. len.	RMSE
	0.14	0.857	0.233	0.079	0.753	0.255	0.112
50	0.39	0.860	0.377	0.122	0.780	0.409	0.163
	0.59	0.837	0.461	0.167	0.783	0.505	0.203
	0.14	0.973	0.109	0.027	0.577	0.122	0.083
200	0.39	0.893	0.182	0.052	0.543	0.198	0.135
	0.59	0.930	0.228	0.059	0.510	0.246	0.174

3. Simulation study

In this section, we discuss the results of a simulation study to evaluate the impact of including/excluding the GFs from the mediator and the outcome models on the estimation of the indirect effect. We will first address the case where the two spatial processes are uncorrelated, then we will move to the case of correlated ones. In both settings, data are generated either from linear models (identity link function) and from a linear mediator model and a logistic outcome model. The former specification allows us to have a unique value of the indirect effect not depending on specific values X, while in the second, there are several indirect effects, depending on the value of X and also on the spatial locations simulated.

3.1. Uncorrelated processes

Let us start by considering the linear case. We simulated data from six different scenarios, obtained by varying the sample size, fixed either to 50 or 200, and the magnitude of the indirect effect, considering small, medium and large effect sizes. Following a relevant stream of literature about mediation analysis (MacKinnon et al., 2002, 2004; Li et al., 2007; Williams and MacKinnon, 2008), these values were set to 0.14, 0.39 and 0.59, respectively. For each scenario, we generated $K = 300$ data sets according to models (13)–(14), where the mediator and the outcome are assumed to be both Normally distributed with expected values as in (13)–(14) and variance fixed to 4, and the spatial processes are GFs with variance 25. The spatial range is set to 0.8 in all scenarios, and the GF of the mediator and the outcome are uncorrelated. We fitted the mediator and the outcome models either ignoring the spatial components and taking them into account, estimating the indirect effect in both cases, evaluating the root mean square error (RMSE), as well as the coverage rates and the average length of confidence and highest density intervals (HDIs, Hoff, 2009). The results are shown in Table 2. As can be noticed, ignoring the GFs in the estimation process leads to lower coverage rates (always below the nominal level 0.95) and a larger average length of the indirect effect's confidence intervals, as well as larger RMSE values. This holds true also when increasing the sample size, where the coverage still reduces. Thus, as expected, ignoring the spatial processes has a remarkable impact on the estimates.

Moving to the nonlinear case, at each iteration, the mediator was generated as before, while the outcome is from a Bernoulli with expectation μ_Y as in Eq. (14) and $g_2 = \text{logit}(\cdot)$. As already mentioned, in this case there is not a unique indirect effect, which depends both on explanatory variables and the values assumed by the spatial processes in the different locations. In real-world settings, one would choose some points of particular interest in correspondence of which estimating the indirect effects for some values of X and other possible covariates. In this light, we kept the coordinates of the points fixed across iterations and selected three locations where we estimated the indirect effects. The locations were selected arbitrarily according to the density of points, which we hypothesize could affect the precision of the estimates; then, we have a point in a low-density area, one in an average-density and another in a high-density area. A graphical representation of the simulated points and the selected ones are shown in Fig. 1. All indirect effects, taking the form

$$\beta_1\gamma_2 \times \frac{\exp(\gamma_0 + \gamma_1 X + \gamma_2 M + u_Y(\mathbf{s}))}{(1 + \exp(\gamma_0 + \gamma_1 X + \gamma_2 M + u_Y(\mathbf{s})))^2},$$

were estimated for $X = \bar{x}$, i.e. the mean of X, and the corresponding value of the mediator, obtained using model (13). This leads to a specific value of the indirect effect at each iteration, but these values are quite similar across replicates. The performance was evaluated using the same measures as before as well as the relative bias. The results are shown in Table 3.

When $n = 50$, we can note that the indirect effects are underestimated in all locations, although the magnitude of bias is smaller when GF is taken into account. The average length of the HDIs is larger when spatial components are considered than in the case without GF, but this difference vanishes increasing the sample size. The RMSE values do not differ between the models with and without GF and reduces as sample size increases. The coverage rates are higher than the nominal level in both cases, although this can be due to the fact that the estimated indirect effects are close to zero and most intervals contain zero.

We also simulated data in the opposite way, i.e. from mediator and outcome models without spatial components u_M and u_Y, and also, in this case, we fitted two models, one including GFs and another one ignoring it. The results are reported in Tables 4 and 5 for the linear and nonlinear cases, respectively. In the former, for $n = 50$, fitting the correct models without GF leads to a better performance in terms of confidence intervals' coverage rates and RMSE than the models including GFs, while the average length of the intervals is approximately equal in either cases. Increasing the sample size to $n = 200$ produces basically the same results across

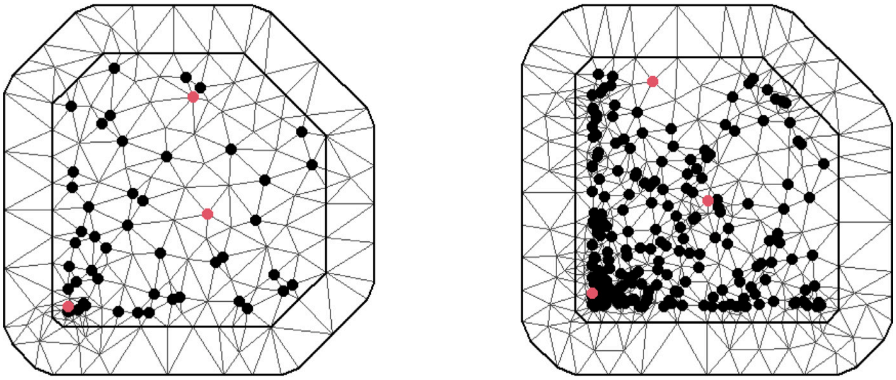

Fig. 1. Set of points used for the simulations in the nonlinear case, when $n = 50$ (left) and $n = 200$ (right). Selected locations where we estimated the indirect effects are highlighted in red.

Table 3
Comparison of performances in estimating the indirect effect including/excluding the GF, varying sample size and spatial locations where the indirect effect is estimated. Data are simulated from nonlinear models including a spatial component. p. dens denotes the density of the points around the one where the indirect effect was estimated.

n	p. dens	GF				No GF			
		rel. bias	Coverage	avg. len.	RMSE	rel. bias	Coverage	avg. len	RMSE
50	Low	−0.680	0.797	0.111	0.020	−1.223	0.937	0.080	0.026
50	Medium	−0.644	0.973	0.178	0.024	−1.221	0.980	0.076	0.021
50	High	−0.469	0.923	0.310	0.047	−1.190	0.843	0.078	0.027
200	Low	0.084	0.963	0.045	0.009	−0.314	0.900	0.031	0.009
200	Medium	0.887	0.983	0.030	0.008	1.080	0.923	0.032	0.009
200	High	−0.021	0.943	0.030	0.007	−0.133	0.957	0.031	0.008

Table 4
Comparison of performances in estimating the indirect effect including/excluding the GF, varying sample size and magnitude of the indirect effect when data are simulated without spatial components.

n	Ind. eff.	GF			No GF		
		Coverage	avg. len.	RMSE	Coverage	avg. len.	RMSE
	0.14	0.893	0.228	0.064	0.923	0.227	0.060
50	0.39	0.917	0.371	0.103	0.953	0.370	0.092
	0.59	0.920	0.458	0.126	0.940	0.455	0.116
	0.14	0.933	0.109	0.029	0.933	0.109	0.029
200	0.39	0.950	0.183	0.048	0.947	0.182	0.047
	0.59	0.920	0.224	0.063	0.930	0.223	0.062

Table 5
Comparison of performances in estimating the indirect effect including/excluding the GF, varying sample size and spatial locations where the indirect effect is estimated. Data are simulated from nonlinear models without a spatial component. p. dens denotes the density of the points around the one where the indirect effect was estimated.

n	p. dens.	GF				No GF			
		rel. bias	Coverage	avg. len.	RMSE	rel. bias	Coverage	avg. len.	RMSE
50	Low	−0.030	0.960	0.126	0.027	−0.575	0.837	0.051	0.016
50	Medium	−0.061	0.953	0.125	0.027	−0.560	0.840	0.051	0.016
50	High	−0.054	0.947	0.133	0.030	−0.574	0.833	0.052	0.016
200	Low	0.034	0.970	0.049	0.012	−0.475	0.607	0.023	0.011
200	Medium	−0.009	0.967	0.048	0.012	−0.494	0.573	0.023	0.011
200	High	0.004	0.927	0.048	0.013	−0.485	0.557	0.023	0.012

the two models, suggesting that in larger samples, including a spatial component when the true data generating mechanisms do not contain it, does not affect the estimates or, in other words, the spatial component is basically estimated as null.

As regards the non-linear case, contrary to what expected, the coverage rates of the model not including the GF are lower than those of the models with GF, probably because of the smaller width of HDIs than those obtained from models with GF. Also the relative bias is larger in the setting without GF, while RMSE values are comparable across the two models when $n = 200$, those obtained without GF are almost a half of those obtained including GF when $n = 50$. The fact that models including GF outperform

Table 6
Comparison of performances in estimating the indirect effect when data are simulated from linear models, varying sample size and the magnitude of the correlation between the spatial process of the mediator and the outcome, including/excluding the GF and accounting or not for such correlation.

n	ρ	No GF			GF			GF corr		
		Coverage	avg. len.	RMSE	Coverage	avg. len.	RMSE	Coverage	avg. len.	RMSE
50	0	0.900	0.260	0.134	0.925	0.239	0.110	1.000	0.499	0.111
50	0.2	0.790	0.524	0.205	0.888	0.483	0.156	0.996	0.991	0.159
50	0.4	0.743	0.524	0.234	0.856	0.484	0.170	0.989	0.987	0.172
50	0.6	0.610	0.539	0.294	0.761	0.492	0.201	0.993	0.996	0.202
50	0.8	0.408	0.551	0.362	0.605	0.496	0.274	0.945	0.982	0.272
200	0	0.451	0.400	0.296	0.712	0.363	0.211	0.914	0.607	0.219
200	0.2	0.505	0.250	0.180	0.923	0.228	0.064	0.929	0.262	0.081
200	0.4	0.399	0.261	0.216	0.882	0.231	0.077	0.927	0.268	0.099
200	0.6	0.091	0.268	0.339	0.619	0.239	0.121	0.921	0.376	0.102
200	0.8	0.084	0.268	0.344	0.630	0.238	0.118	0.930	0.935	28.795

Table 7
Comparison of performances in estimating the indirect effect when the mediator is from a linear model and the outcome from a binary one, varying sample size and the magnitude of the correlation between the spatial process of the mediator and the outcome, including/excluding the GF and accounting or not for such correlation. p. dens denotes the density of the points around the one where the indirect effect was estimated.

n	ρ	p. dens	No GF				GF				GF corr			
			rel. bias	Coverage	avg. len.	RMSE	rel. bias	Coverage	avg. len.	RMSE	rel. bias	Coverage	avg. len.	RMSE
50	0.2	Low	−0.736	0.983	0.059	0.015	−0.398	0.980	0.135	0.021	−0.325	0.970	0.155	0.021
		Medium	−0.633	0.993	0.059	0.013	−0.090	0.967	0.113	0.018	−0.067	0.973	0.132	0.019
		High	−0.767	0.993	0.058	0.012	−0.293	0.813	0.090	0.021	−0.309	0.850	0.100	0.023
	0.4	Low	0.285	0.953	0.094	0.023	0.310	0.953	0.183	0.028	0.311	0.943	0.189	0.026
		Medium	−0.529	0.947	0.095	0.025	−0.215	0.947	0.241	0.034	−0.249	0.920	0.247	0.035
		High	11.255	0.937	0.094	0.025	3.670	0.963	0.063	0.013	3.770	0.940	0.068	0.014
	0.6	Low	−0.413	0.943	0.079	0.021	0.103	0.970	0.211	0.041	0.072	0.960	0.221	0.040
		Medium	−0.199	0.933	0.097	0.025	0.515	0.947	0.240	0.048	0.394	0.953	0.251	0.040
		High	2.943	0.910	0.095	0.026	2.214	0.960	0.122	0.031	2.215	0.957	0.122	0.030
	0.8	Low	0.429	0.937	0.095	0.026	0.976	0.967	0.239	0.048	0.819	0.957	0.242	0.047
		Medium	−0.424	0.953	0.095	0.024	0.361	0.967	0.280	0.049	0.389	0.947	0.292	0.056
		High	−0.434	0.930	0.082	0.024	0.647	0.953	0.228	0.064	0.522	0.960	0.231	0.057
200	0.2	Low	−0.900	0.623	0.016	0.007	−0.225	0.860	0.021	0.005	−0.039	0.933	0.027	0.006
		Medium	−0.923	0.457	0.016	0.008	0.374	0.983	0.035	0.008	0.546	0.983	0.040	0.009
		High	−0.963	0.083	0.017	0.016	−0.354	0.793	0.032	0.009	−0.279	0.863	0.036	0.009
	0.4	Low	−0.264	0.903	0.037	0.011	0.036	0.947	0.052	0.012	−0.057	0.937	0.049	0.012
		Medium	0.271	0.950	0.038	0.009	0.569	0.993	0.044	0.011	0.511	0.997	0.044	0.010
		High	−0.333	0.907	0.037	0.011	0.279	0.957	0.057	0.016	0.102	0.960	0.054	0.014
	0.6	Low	−0.364	0.900	0.041	0.013	0.079	0.967	0.058	0.012	−0.066	0.960	0.054	0.011
		Medium	0.941	0.883	0.042	0.013	2.339	0.933	0.060	0.021	1.896	0.967	0.056	0.018
		High	0.804	0.903	0.042	0.012	0.028	0.950	0.024	0.006	0.206	0.990	0.030	0.006
	0.8	Low	2.136	0.833	0.045	0.016	1.058	0.997	0.039	0.009	1.205	1.000	0.041	0.010
		Medium	2.769	0.817	0.046	0.016	1.914	0.980	0.037	0.011	2.188	0.987	0.042	0.011
		High	0.048	0.940	0.045	0.011	0.826	0.937	0.066	0.020	0.352	0.990	0.056	0.014

those without GF even in the case when data truly do not present spatial processes lets us think that the presence of GF helps in catching the underlying complexity of the nonlinear model, leading to superior performances.

3.2. Correlated processes

To evaluate how taking into account the correlation between the two processes impacts the indirect effect estimates, we ran another simulation study, where data were generated from two models as in the previous section, however, in the linear case, the indirect effect is fixed at 0.6 and the covariance between u_M and u_Y ranges from 0 to 0.8, with a step of 0.2. We assessed the coverage rates and the average length of HDIs for the indirect effect and the RMSE by fitting models without GFs, models including GFs but ignoring the correlation between the two processes and models including GFs and accounting for the correlation. Results are shown in Tables 6 and 7.

In the linear case, it can be noted that the coverage rates of HDIs are always below the nominal level for both the model without GF and the model including a GF but not accounting for the correlation between u_M and u_Y, and they decrease as the sample size increases, as it is particularly evident for the model without GF. In contrast, the coverage rates obtained using a model with GF and taking into account ρ are close to the nominal level in every scenario. This may also be due to a higher average length of these intervals compared to that of the intervals obtained from the other models, although when the sample size is 200, this difference attenuates. Also, the RMSE of the indirect effect estimates in the most complex model is comparable with those of the other models, except for the case of $n = 200$ and $\rho = 0.8$ when the RMSE is incredibly high. This may be due to a computational problem of INLA, for which it is quite difficult to disentangle the components of u_Y given the high correlation with u_M. It is also worth remarking that, despite the increase in ρ, the spatial model without the correlation appears to maintain better performance compared to the non-spatial model, as indicated by higher coverage of credibility intervals and lower average length of intervals for most combinations of n and ρ.

In the nonlinear case, the relative bias of estimates obtained without GF generally has a larger magnitude (in absolute value) than that of estimates from models including GF, although, for $n = 200$ and $\rho = 0.8$, all models show very high values of relative bias. The average length of HDIs and the RMSE are smaller for models without GF when $n = 50$, but the differences across models reduces as the sample size increases. Coverage rates are almost always close to the nominal level, but those relative to models including GF are generally higher than those obtained without GF. Taking into account the correlation between the mediator and the outcome spatial processes yields a better performance in terms of relative bias, coverage rates and RMSE, while the average length of HDIs is pretty similar between the models including GF, with and without correlation. In general, the density of the points in the neighborhood of the location where the indirect effect was estimated affects the estimates, but not in a linear way.

In general, when data are generated from nonlinear models (with or without correlated processes), the results of simulations are more difficult to read and models with GF do not always perform better than those without GF. We checked if this issue may be due to the magnitude of the range, so we replicated the simulations in the nonlinear setting fixing the range to 0.2, but the results do not differ from those already observed, see the Supplementary material.

4. An application to real data

In this section, we provide an application of our proposal to a real data set, to show its usefulness in applied problems and its applicability in more complex settings.

4.1. Motivation of the study

One of the areas where spatial mediation can prove particularly useful is the analysis of ecological data, which often exhibits spatial correlation along with a complex structure of association among the factors involved. In order to illustrate an application on real ecological data (albeit as a small example rather than a comprehensive study), we investigate the effect of sea depth on the abundance of the Pandora (*pagellus erythrinus*) fish species, as mediated by temperature.

The Pagellus erythrinus is a fish species widely distributed in the Mediterranean Sea, the Black Sea, and the Atlantic waters adjacent to southern Europe. Its distribution is closely related to abiotic factors such as temperature and water depth. Studies have shown that Pagellus erythrinus prefers environments with temperatures between 12 and 24 °C, exhibiting a strong preference for coastal waters and rocky or mixed bottoms (Moranta et al., 1998). Understanding its distribution patterns is crucial for the development of species management and conservation strategies, considering also the impact of fishing activities and climate change on its habitat (Dulčić and Grbec, 2000).

4.2. The data

The data are georeferenced biomass indices of fish within the demersal trawl surveys MEDITS (Mediterranean International Trawl Survey program, Bertrand et al., 2002), performed in the study area between 1995 and 2018. The MEDITS survey is carried out annually in late spring-early summer, providing a long-term dataset of fishery-independent data relating to demersal species abundance, demographic structure, and spatial distribution. Sampling followed a random design stratified by depth (depth strata: 10–50 m, 51–100 m, 101–200 m, 201–500 m, 501–800 m) with the number of haul per stratum proportional to each stratum surface. Specifically, we analyzed $n = 56$ catches of Pandora, considering its abundance (biomass) measured in grams/km^2, the Bottom Sea Temperature (BST) measured in Celsius degrees and the depth of the catch measured in meters. The depth of the catches ranges between 18 m (minimum observed depth in the dataset) and 75 m (the thermocline, beyond which temperature no longer decreases with increasing depth), considering only the most recent available period from 2015 to 2018. Fig. 2 shows the location of Pandora catches and their corresponding depth in the study area.

4.3. An INLA-SPDE spatial mediation model for Pandora

To analyze the effect of depth on the abundance of Pandora as mediated by temperature, we propose the following mediation model:

$$\mu_M(s) = \beta_0 + \beta_1 X(s) + u_M(s) \tag{19}$$

$$\log(\mu_Y(s)) = \gamma_0 + \gamma_1 X(s) + \gamma_2 M(s) + u_Y(s), \tag{20}$$

where X is the depth, M is the temperature, Y is the biomass of Pandora, s is the vector of spatial coordinates (latitude and longitude) of the catch, and $u_M(s)$ and $u_Y(s)$ are two Matérn Gaussian fields with range and standard deviation r_{u_M}, σ_{u_M} and r_{u_Y}, σ_{u_Y}, respectively. We assumed that $M \sim \mathcal{N}(\mu_M, \sigma^2_M)$; also, considering the non-normal distribution of biomass, we assumed

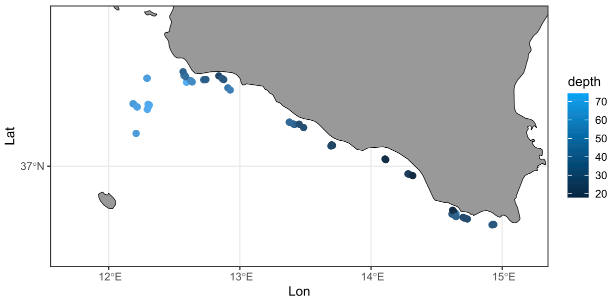

Fig. 2. Location of Pandora fish catches in the study area and their corresponding depth.

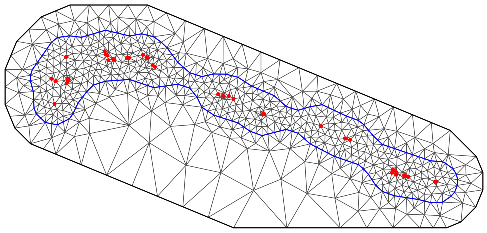

Fig. 3. Triangulation of the study region and catch locations (red dots).

that $Y \sim$ Gamma with expected value μ_Y and variance σ_Y^2, linked to its linear predictor through log link function, which requires nonlinear mediation analysis.

The models expressed in Eqs. (19) and (20) were estimated using the INLA-SPDE approach. To choose the priors for the hyperparameters of the SPDE component, we assumed a joint Normal prior for hyperparameters λ_1 and λ_2 with expected values $\mu_{\lambda_1} = -2.64$ and $\mu_{\lambda_2} = 1.38$, precision $\tau_{\lambda_1} = \tau_{\lambda_2} = 0.1$ and covariance zero. Fig. 3 shows the triangulation of the study region (mesh). Maximum triangle side length is 0.08 degree for the study region and 0.4 degree for the outer region, and minimum triangle side length is equal to 0.008 degrees.

After drawing samples from the (approximate) posterior distributions of the latent effects and the hyperparameters, a sample of S draws from the posterior distribution of the indirect effect can be computed, following the derivative-based approach, as follows:

$$CIE^p(\boldsymbol{s}') = \boldsymbol{\beta}_1^p \odot \boldsymbol{\gamma}_2^p \odot \exp\left(\boldsymbol{\gamma}_0^p + \boldsymbol{\gamma}_1^p x(\boldsymbol{s}') + \boldsymbol{\gamma}_2^p \odot \boldsymbol{m}^p(\boldsymbol{s}') + \sum_{j=1}^{m} A^*_{js'} \boldsymbol{w}_j^p\right), \tag{21}$$

where $\boldsymbol{s}'$ is a prediction location of interest, $\odot$ indicates the Hadamard product, $\boldsymbol{\beta}_1^p$, $\boldsymbol{\gamma}_0^p$, $\boldsymbol{\gamma}_1^p$, $\boldsymbol{\gamma}_2^p$, $\boldsymbol{w}^p$ are vectors of posterior samples of length S, $x(\boldsymbol{s}')$, $\boldsymbol{m}(\boldsymbol{s}')$ are measured depth and posterior predictive BST, respectively, for location $\boldsymbol{s}'$, and $A^*_{js'}$ is a generic element of the projection matrix $\boldsymbol{A}^*_{L\times m}$ which maps the GMRF from the m triangulation vertices to the L prediction locations of interest. The indirect effect was estimated both with and without the spatial component at the mediator and response levels.

4.4. Results

Table 8 reports posterior statistics for parameters and hyperparameters of the INLA-SPDE models expressed in Eqs. (19) and (20). The inclusion of the spatial components leads to a decrease in DIC for both models (from 165.06 to 110.12 for the mediator model and from 1191.45 to 1177.28 for the outcome model). The posterior means of the SDs of the spatial components ($\sigma_{u_M} = 0.85$ °C and $\sigma_{u_Y} = 0.92$ °C) suggest that their inclusion into the mediation model allowed to capture some spatial heterogeneity, unexplained by covariates. Also, considering the posterior means of the range parameters ($r_{u_M} = 1.08$ and $r_{u_Y} = 0.04$), the spatial correlation for BST appears to decrease more rapidly with distance than that for biomass, suggesting a relatively higher variability in the spatial distribution of BST compared to biomass.

High-density intervals of the effects β_1, γ_1, and γ_2 provide evidence of their significance in a Bayesian context. Bottom sea temperature appears to decrease with increasing depth, as expected, while the abundance of Pandora fish decreases with increasing temperature (posterior means of β_1 and γ_2 are -0.05 and -0.42, respectively). The significance of γ_1 suggests that depth could affect Pandora abundance not only indirectly through a change in temperature, but also directly.

Fig. 4 shows the mean and standard deviation of the posterior distribution of the CIE, computed for each point in a grid of $L = 10^4$ prediction locations of interest $P = \{\boldsymbol{s}_1, \dots, \boldsymbol{s}_L\}$ in the study region.

Considering the 95% credibility intervals based on quantiles, computed for each location in P, none of them contains the value 0. Conversely, the highest density intervals tend to be more conservative (12.64% of them do not contain zero). This difference can be

Table 8
Posterior statistics for parameters and hyperparameters of the mediation model.

	Mean	Sd	Median	Mode	HDI low	HDIhigh
γ_0	17.925	3.019	17.934	17.935	11.969	23.848
γ_2	−0.415	0.160	−0.415	−0.415	−0.731	−0.099
γ_1	−0.040	0.012	−0.040	−0.040	−0.064	−0.015
σ_Y	0.767	0.095	0.763	0.760	0.583	0.955
σ_{u_Y}	0.922	0.351	0.864	0.756	0.336	1.622
r_{u_Y}	0.040	0.034	0.030	0.018	0.003	0.103
β_0	18.581	0.982	18.603	18.597	16.468	20.599
β_1	−0.049	0.011	−0.049	−0.049	−0.071	−0.027
σ_M	0.559	0.064	0.554	0.545	0.439	0.686
σ_{u_M}	0.854	0.243	0.816	0.743	0.439	1.345
r_{u_M}	1.083	0.659	0.908	0.664	0.227	2.380

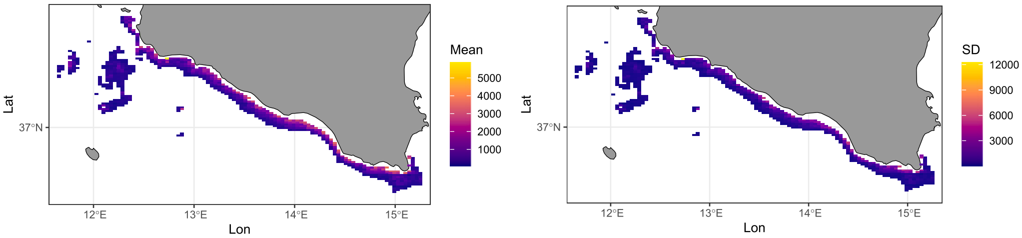

Fig. 4. Mean and SD of the posterior distribution of the indirect effect in the study area.

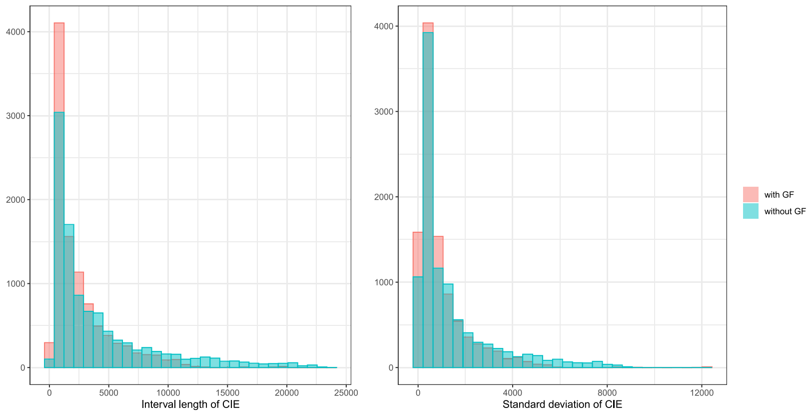

Fig. 5. Comparison of the SDs and HDI interval lengths of the posterior distribution of the CIE, for the spatial and the non-spatial model.

attributed to the asymmetry of the posterior distributions of the indirect effect. Moreover, the model without the spatial component results in an 87% highest density intervals that does not contain zero. Therefore, the inclusion of the spatial components in this case tends to weaken the significance of the indirect effect. Also, comparing with the model without the spatial component, the spatial model yields posterior distributions of the CIE that are less dispersed and with narrower HDIs, as shown in Fig. 5.

The posterior mean of the indirect effect is positive for each location: this suggests that with increasing depth, Pandora fish tends to be more abundant due to the decrease in temperature. This is in agreement with the discussed results reported in Table 8.

5. Conclusions

In this paper, we explored how to address spatial mediation by combining the INLA-SPDE approach with the derivative-based method within the Bayesian inferential framework. This was done to estimate mediational effects in the context of nonlinear mediation analysis, particularly focusing on the inclusion of spatial correlation in geostatistical data.

To evaluate the impact of considering spatial correlation when estimating indirect effects with point-referenced data, we compared a traditional mediation model with a spatial mediation model. The latter model incorporates a Gaussian spatial process at both the mediator and outcome levels. Our simulation results highlighted the potential for biased estimates when spatial correlation is neglected, especially in scenarios characterized by large sample sizes. We also showed the importance of taking into account spatial correlation, although when it is very high, it heavily impacts the estimates and their variance.

In the context of an application on ecological data, the INLA-SPDE spatial mediation model provided evidence of a positive indirect effect of depth on the abundance of Pandora fish, mediated by bottom sea temperature. Results suggest that including spatial components leads to less dispersed posterior distributions of the indirect effect with narrower credibility intervals, thereby impacting its significance. Additionally, given the asymmetry of the posterior distributions of the indirect effect, the choice of credibility intervals (quantile-based vs. HDI) influences conclusions about the significance of the CIE, with HDIs being more conservative.

Although our work presents several characters of novelty, it is not free of limitations. First of all, both in the simulation study and in the application, we did not consider the case of additional covariates. They could be spatial variables as well, and may play the role of confounders between X, M and Y. Our approach can definitely deal with such kind of variables, but effectiveness of INLA in disentangling their spatial components from those of the three main mediational variables and the role played by the correlation between their spatial processes and those of X, M and Y is yet to be investigated. Secondly, although we managed to account for the correlation between u_M and u_Y when estimating mediation models, we ignored it in the computation of the indirect effect. Consistently with the literature concerning mediation with correlated errors (le Cessie, 2016; Zhao and Luo, 2014), we believe that taking into account such a correlation primarily reflects on the bias of all coefficient estimates, leading to unbiased estimates of mediational effects. Nonetheless, we did not explore the issue of whether and how the formulas for the indirect effect change when the correlation between spatial processes (or errors more generally) is addressed. This is a rather underexplored topic in the field of mediation in general.

To conclude, our work fits in the recent stream of literature concerning spatial mediation but, in contrast to most contributions focusing on areal data, it deals with geostatistical data. We addressed a very simple setting including only an exposure, a mediator and an outcome (with the possibility of including additional covariates), but our framework can be extended to more complex settings. First of all, the inclusion of a temporal component to accommodate longitudinal data should be straightforward. Second, the case of multiple mediators, quite common in applied research, may be addressed carefully considering the spatial and temporal dependence among them. Finally, our proposal was developed in an associational framework, but providing a causal interpretation of spatial effects, as done, for example, in Smith et al. (2023), is a challenging future research direction.

CRediT authorship contribution statement

Claudio Rubino: Conceptualization, Methodology, Software, Formal analysis, Visualization, Writing – original draft, Writing – review & editing. **Chiara Di Maria:** Conceptualization, Methodology, Software, Formal analysis, Visualization, Writing – original draft, Writing – review & editing. **Antonino Abbruzzo:** Conceptualization, Writing – review & editing, Supervision. **Gioacchino Bono:** Data curation. **Germana Garofalo:** Data curation. **Giacomo Milisenda:** Data curation. **Giada Adelfio:** Conceptualization, Writing – review & editing, Supervision.

Acknowledgments

Giada Adelfio's Research is funded by the European Union – NextGenerationEU; Award Number: Project code CN−00000033, Concession Decree No. 1034 of 17 June 2022 adopted by the Italian Ministry of University and Research, CUP UNIPA B73C22000790001, Project title "National Biodiversity Future Center - NBFC" and MUR- MUR - PRIN 2022: Spatio-temporal Functional Marked Point Processes for probabilistic forecasting of earthquakes 2022BN7CJP. P. I. G.A. CUP B53C24006340006.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.spasta.2025.100885.

References

Abrahamsen, P., 1997. A Review of Gaussian Random Fields and Correlation Functions. Norsk Regnesentral/Norwegian Computing Center Oslo.

Bakka, H., Rue, H., Fuglstad, G.-A., Riebler, A., Bolin, D., Illian, J., Krainski, E., Simpson, D., Lindgren, F., 2018. Spatial modeling with R-INLA: a review. Wiley Interdiscip. Rev.: Comput. Stat. 10 (6), e1443.

Banerjee, S., Carlin, B.P., Gelfand, A.E., 2003. Hierarchical Modeling and Analysis for Spatial Data. Chapman and Hall/CRC.

Baron, R.M., Kenny, D.A., 1986. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. J. Pers. Soc. Psychol. 51 (6), 1173–1182.

Bertrand, J.A., Gil De Sola, L., Papaconstantinou, C., Relini, G., Souplet, A., 2002. The general specifications of the MEDITS surveys. Sci. Mar. 66 (S2), 9–17.

Bustamante, B.L.M., May, L., Fejerman, L., Martínez-López, B., 2023. A Bayesian multilevel analysis exploring population-level effects mediating the relationship between area-level poverty and community-acquired methicillin-resistant staphylococcus aureus (CA-MRSA) infection across California communities. Heal. Place 83, 103094.

Cameletti, M., Lindgren, F., Simpson, D., Rue, H., 2013. Spatio-temporal modeling of particulate matter concentration through the SPDE approach. AStA Adv. Stat. Anal. 97 (2), 109–131.

le Cessie, S., 2016. Bias formulas for estimating direct and indirect effects when unmeasured confounding is present. Epidemiol. 27 (1), 125–132.

Ciarlet, P.G., 1978. The Finite Element Method for Elliptic Problems. North-Holland, Amsterdam.

Cosandey-Godin, A., Krainski, E.T., Worm, B., Flemming, J.M., 2014. Applying Bayesian spatiotemporal models to fisheries bycatch in the Canadian arctic. Can. J. Fish. Aquat. Sci. 72, 1–12.

Cressie, N., 1993. Statistics for Spatial Data. Wiley, Hoboken.

Cressie, N., Wickle, C.K., 2011. Statistics for Spatiotemporal Data. Wiley, New York.

Di Maria, C., Rubino, C., Albano, A., 2024. The derivative-based approach to nonlinear mediation models: insights and applications. Qual. Quant. 58, 4383–4405.

Dulčić, J., Grbec, B., 2000. Climate change and its influence on the biology and management of the dentex (Pagellus Erythrinus) in the Adriatic sea. Clim. Change 49 (3), 285–293.

Fioravanti, G., Martino, S., Cameletti, M., Toreti, A., et al., 2023. Interpolating climate variables by using INLA and the SPDE approach. Int. J. Climatol. 43 (14), 6866–6886.

Geldhof, G.J., Anthony, K.P., Selig, J.P., Mendez-Luck, C.A., 2018. Accommodating binary and count variables in mediation: a case for conditional indirect effects. Int. J. Behav. Dev. 42 (2), 300–308.

Gómez-Rubio, V., 2020. Bayesian Inference with INLA. Chapman & Hall/CRC Press, Boca Raton, FL.

Gortázar, C., Fernández-Calle, L.M., Collazos-Martínez, J.A., Mínguez-González, O., Acevedo, P., 2017. Animal tuberculosis maintenance at low abundance of suitable wildlife reservoir hosts: A case study in northern Spain. Prev. Vet. Med. 146, 150–157. http://dx.doi.org/10.1016/j.prevetmed.2017.08.009.

Hayes, A.F., Preacher, K.J., 2010. Quantifying and testing indirect effects in simple mediation models when the constituent paths are nonlinear. Multivar. Behav. Res. 45, 627–660.

Hoff, P.D., 2009. A First Course in Bayesian Statistical Methods. Springer.

Huang, J., Yuan, Y., 2017. Bayesian dynamic mediation analysis. Psychol. Methods 22 (4), 667–686. http://dx.doi.org/10.1037/met0000073.

Jarvis, C.I., Multerer, L., Lewis, D., Binka, F., Edmunds, W.J., Alexander, N., Smith, T.A., 2019. Spatial effects of permethrin-impregnated bed nets on child mortality: 26 years on, a spatial reanalysis of a cluster randomized trial. Am. J. Trop. Med. Hyg. 101 (6), 1434.

Lezama-Ochoa, N., Pennino, M.G., Hall, M.A., Lopez, J., Murua, H., 2020. Using a Bayesian modelling approach (INLA-spde) to predict the occurrence of the spinetail devil ray (mobular mobular). Sci. Rep. 10 (1), 18822.

Li, Y., Schneider, J.A., Bennett, D.A., 2007. Estimation of the mediation effect with a binary mediator†. Stat. Med. 26 (16), 3398–3414. http://dx.doi.org/10.1002/sim.2730.

Lindgren, F., Rue, H., Lindstrom, J., 2011. An explicit link between Gaussian fields and Gaussian Markov random fields: The stochastic partial differential equation approach. J. R. Stat. Soc. Ser. B 73, 423–498.

Lu, Y., Zhu, S., 2022. Digital economy, scientific and technological innovation, and high-quality economic development: A mediating effect model based on the spatial perspective. PLoS One 17 (11), e0277245.

MacKinnon, D.P., 2008. Introduction to Statistical Mediation Analysis. Taylor & Francis Group, New York.

MacKinnon, D.P., Lockwood, C.M., Hoffman, J.M., West, S.G., Sheets, V., 2002. A comparison of methods to test mediation and other intervening variable effects. Psychol. Methods 7 (1), 83–104. http://dx.doi.org/10.1037/1082-989x.7.1.83.

MacKinnon, D.P., Lockwood, C.M., Williams, J., 2004. Confidence limits for the indirect effect: Distribution of the product and resampling methods. Multivar. Behav. Res. 39 (1), 99–128.

Martins, T.G., Simpson, D., Lindgren, F., Rue, H., 2013. Bayesian computing with INLA: new features. Comput. Statist. Data Anal. 67, 68–83.

Matérn, B., 1960. Stochastic Models and Their Application to Some Problems in Forest Surveys. Stockholm.

Miočević, M., 2019. A tutorial in Bayesian mediation analysis with latent variables. Methodology 15 (4), 137–146. http://dx.doi.org/10.1027/1614-2241/a000177.

Miočević, M., Gonzalez, O., Valente, M.J., MacKinnon, D.P., 2017. A tutorial in Bayesian potential outcomes mediation analysis. Struct. Equ. Model.: A Multidiscip. J. 25 (1), 121–136. http://dx.doi.org/10.1080/10705511.2017.1342541.

Moraga, P., Dean, C., Inoue, J., Morawiecki, P., Noureen, S.R., Wang, F., 2021. Bayesian spatial modelling of geostatistical data using INLA and SPDE methods: A case study predicting malaria risk in mozambique. Spat. Spat. Tempor. Epidemiol. 39, 100440. http://dx.doi.org/10.1016/j.sste.2021.100440.

Moranta, J., Stefanescu, C., Massutí, E., Morales-Nin, B., Lloris, D., 1998. Fish community structure and depth-related trends on the continental slope of the balearic islands (Algerian basin, western Mediterranean). Mar. Ecol. Prog. Ser. 171, 247–259.

Mougeni, F., Lell, B., Kandala, N.-B., Chirwa, T., 2024. Bayesian spatio-temporal analysis of malaria prevalence in children between 2 and 10 years of age in Gabon. Malar. J. 23 (1), 57.

Naeeim, N.S.A., Rahman, N.A., 2017. Estimating relative risk for dengue disease in peninsular Malaysia using INLA. Malays. J. Fundam. Appl. Sci. 13 (4).

Pearl, J., 2001. Direct and indirect effects. In: Breese, J., Koller, D. (Eds.), Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence. Morgan Kauffmann, San Francisco, pp. 411–420.

Rue, H., Held, L., 2005. Gaussian Markov Random Fields: Theory and Applications. Chapman & Hall, Boca Raton.

Rue, H., Martino, S., Chopin, N., 2009. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations (with discussion). J. R. Stat. Soc. Ser. B 71, 319–392.

Schluchter, M.D., 2008. Flexible approaches to computing mediated effects in generalized linear models: Generalized estimating equations and bootstrapping. Multivar. Behav. Res. 43 (2), 268–288.

Smith, M.J., Charlton, M.E., Oleson, J.J., 2023. Causal decomposition maps: An exploratory tool for designing area-level interventions aimed at reducing health disparities. Biom. J. 65 (8), 2200213.

Stolzenberg, r.m., 1980. The measurement and decomposition of causal effects in nonlinear and nonadditive models. Sociol. Methodol. 11, 459–488.

Sun, R., Zhou, X., Song, X., 2021. Bayesian causal mediation analysis with latent mediators and survival outcome. Struct. Equ. Model.: A Multidiscip. J. 28 (5), 778–790. http://dx.doi.org/10.1080/10705511.2020.1863154.

Tsai, T., Shau, W., Hu, F., 2006. Generalized path analysis and generalized simultaneous equations model for recursive systems with responses of mixed types. Struct. Equ. Model. 13 (2), 229–251.

VanderWeele, T.J., 2015. Explanation in Causal Inference: Methods for Mediation and Interaction. Oxford University Press.

Wang, D., Xu, L., Du, J., 2023. The direct and indirect spatial spillover effects of infrastructure on urban green and smart development. Front. Env. Sci. 11, 1197048.

Williams, J., MacKinnon, D.P., 2008. Resampling and distribution of the product methods for testing indirect effects in complex models. Struct. Equ. Model.: A Multidiscip. J. 15 (1), 23–51. http://dx.doi.org/10.1080/10705510701758166.

Wittle, P., 1954. On stationary processes in the plane. Biometrika 41, 434–449.

Wright, S., 1934. The method of path coefficients. Ann. Math. Stat. 5 (3), 161–215.

Yuan, Y., MacKinnon, D.P., 2009. Bayesian mediation analysis. Psychol. Methods 14 (4), 301–322.

Zhao, Y., Luo, X., 2014. Estimating causal mediation effects under correlated errors. arXiv preprint arXiv:1410.7217.