Università degli Studi di Palermo

AREA QUALITÀ, PROGRAMMAZIONE E SUPPORTO STRATEGICO
SETTORE STRATEGIA PER LA RICERCA
U. O. DOTTORATI

Dottorato in Scienze Economiche e Statistiche
Dipartimento di Scienze Economiche, Aziendali e Statistiche
SECS-S/01 - Statistica

INLA-SPDE SPATIAL MODELLING AND BAYESIAN MEDIATION ANALYSIS: INSIGHTS, ADVANCES AND APPLICATIONS

IL DOTTORE
CLAUDIO RUBINO

IL COORDINATORE
VITO M. R. MUGGEO

IL TUTOR
GIADA ADELFIO

IL CO-TUTOR
ANTONINO ABBRUZZO

CICLO XXXVI
ANNO CONSEGUIMENTO TITOLO 2024

Acknowledgements

Thanks to my supervisor, Prof. Giada Adelfio, and my co-supervisor, Prof. Antonino Abbruzzo, for their invaluable support throughout this project.

Also thanks to Massimo Attanasio, the Principal Investigator, and Andrea Priulla for granting access to the valuable data provided by the Ministero dell'Istruzione, dell'Università e della Ricerca (MIUR), PRIN 2017 project titled 'From high school to job placement: micro data life course analysis of university student mobility and its impact on the Italian North-South divide' [grant n. 2017HBTK5P].

Thanks to the reviewers, whose valuable suggestions have contributed to improving the quality of this work.

Furthermore, I want to express my gratitude to the professors, researchers and colleagues I had the pleasure of collaborating with during my Ph.D. experience, and all those who have contributed to this project and to my personal and professional growth.

Finally, a very special thanks go to my wife, Diletta, and my beloved daughter, Marta. Their presence has made every challenge more manageable and every success more meaningful.

Contents

Introduction

This doctoral thesis presents an exploration of Bayesian statistical methods, focusing on two primary areas: spatial modeling using the INLA-SPDE approach and mediation analysis. The study also investigates a combined approach to estimate mediational effects in presence of spatial correlation.

The first Chapter of the thesis introduces an application of the INLA-SPDE approach in the context of marine ecology, exploring the effect of the environmental temperature on the faunal composition of the Central Mediterranean Sea. We exploit the modeling flexibility of this approach, using the SPDE approximation to Gaussian fields along with stochastic smoothing techniques to provide insights into the potential impacts of climate change on demersal fish communities.

The second Chapter of the thesis shifts the focus to nonlinear mediation analysis. Mediation models are commonly used in biology, social sciences and epidemiology to assess the indirect effects of an exposure on an outcome through an intermediate variable, which is called mediator. In many real-world applications, the presence of either interaction terms in the models or not normally distributed variables, which requires link functions different from the identity, implies the non-applicability of techniques typical of traditional mediation to estimate mediational effects. This issue has primarily been addressed in literature using non-parametric counterfactual-based approaches, but very few solutions have been proposed to deal with nonlinear models in a path-analytic framework.

First, in the context of a healthcare research, we use a two-step approach combining graphical models and counterfactual-based nonlinear mediation analysis to investigate the effect of socio-economic status on cancer mortality as mediated by out-of-region mobility, focusing on lung and colon cancer patients in Sicily, Italy. The study highlights the role of extra-regional mobility as a mediator and its implications for health policy.

Than, we move into the path-analytic framework offering a comprehensive discussion of the derivative-based method for nonlinear mediation analysis proposed by Geldhof et al. (2018), addressing some of its potential issues through novel solutions.

We also present a real data application, exploiting some of our developments.

In the third Chapter it is addressed, in the context of a work-in-progress, the possibility to combine the INLA-SPDE method and the derivative-based approach to mediation analysis for estimating non-linear mediational effects with spatially-correlated data, enabling the computation of a spatial conditional indirect effect. With very few exceptions of recent explorations into spatial variability in causal inference (Reich et al., 2021), the literature on mediation analysis predominantly lacks consideration for spatial heterogeneity. With simulations and an application on real data, we illustrate the advantages of considering spatial correlation in mediation analysis, particularly in improving the accuracy of the estimation of the indirect effect.

Chapter 1

Applying the INLA-SPDE approach to explore the effects of temperature on the faunal composition of the Central Mediterranean Sea

1.1 Introduction

In this chapter, we initially introduce the Integrated nested Laplace approximation technique for the estimation of Bayesian models. Than, we illustrate the concept of employing Gaussian fields for spatial data analysis, and how they can be approximated using the Stochastic Partial Differential Equation (SPDE) method. Subsequently, we present a modellistic approach, which exploits the modeling flexibility of INLA-SPDE within the realm of marine ecology.

1.2 Integrated nested Laplace approximation

Within the Bayesian inferential framework, prior beliefs about unknown parameters are updated using information from observed data, obtaining the parameters' posterior distribution. From the posterior distribution, relevant statistics related to the parameters of interest are computed, like marginal distributions, mean values, variances, quantiles, among others (for an extensive treatment of the Bayesian approach to statistical inference see, for example, Gelman et al., 2013).

However, in real-world applications, the main obstacle lies in the fact that com-

puting the posterior distribution analytically is rarely feasible, hence computational procedures to obtain the posterior distribution must be employed. Several approaches are available, with the most renowned being using Markov chain Monte Carlo (MCMC) techniques (Robert et al., 1999). This method involves drawing a sample from the joint posterior distribution and conducting inference based on this posterior sample. However, MCMC methods are computationally demanding, and generally require a substantial amount of time to reach convergence.

As an appealing alternative to MCMC methods, the integrated nested Laplace approximation (INLA) has been proposed by Rue et al. (2009), and further developed by Martins et al. (2013). The INLA method has been successfully applied to various types of analyses, such as spatio-temporal models for meteorological data (Lindgren et al., 2011), epidemiology (Bisanzio et al., 2011), pollution risk mapping (Cameletti et al., 2013), disease mapping and spread (Schrodle and Held, 2011; Schrodle et al., 2012), fishing practices (Cosandey-Godin et al., 2014), econometric models (Gómez-Rubio et al., 2015a), mixed generalized linear models (Fong et al., 2010), log-Gaussian Cox processes (Gómez-Rubio et al., 2015b; Illian and Sørbye, Rue, H. et al., 2012) and survival analysis (Martino et al., 2011). An extensive list of applications of INLA can be found in Rue et al. (2017).

A model can be estimated with INLA if for a random variable Y its expected value μ can be modeled through a link function $g(\cdot)$ in an additive way as

$$g(\mu_i) = \eta_i = \beta_0 + \sum_{m=1}^{M} \beta_m x_{mi} + \sum_{l=1}^{L} f_l(z_{li}), \tag{1.1}$$

where β_0 is the intercept, the coefficients $\boldsymbol{\beta} = \{\beta_1, \ldots, \beta_M\}$ quantify the effect of the explanatory variables $\boldsymbol{x}$ on the response variable y_i and $\boldsymbol{f} = \{f_1(\cdot), \ldots, f_L(\cdot)\}$ is a collection of unknown functions defined on a set of covariates $\boldsymbol{z}$. Given that the terms $f_l(\cdot)$ can take different forms, the class of models which could be expressed as in Equation (1.1) is very large (Fahrmeir and Lang, 2001).

INLA assumes Gaussian priors to the vector of parameters $\boldsymbol{\theta} = \{\boldsymbol{\eta}, \beta_0, \boldsymbol{\beta}, \boldsymbol{f}\}$. In particular, INLA applies properly only if $\boldsymbol{\theta}$ can be written as a Gaussian Markov Random Field (GMRF), because the conditional independence structure of GMRFs allows to use appropriate efficient computational techniques (Rue and Held, 2005). The vector $\boldsymbol{\theta}$ may depend on some hyperparameters $\boldsymbol{\psi}$ (e.g. variances or correlation parameters) with, in general, $\dim(\boldsymbol{\psi}) \ll \dim(\boldsymbol{\theta})$.

The marginal posterior distributions of the parameters are given by

$$p(\theta_i \mid \boldsymbol{y}) = \int p(\theta_i, \boldsymbol{\psi} \mid \boldsymbol{y})\, d\boldsymbol{\psi} = \int p(\theta_i \mid \boldsymbol{\psi}, \boldsymbol{y}) p(\boldsymbol{\psi} \mid \boldsymbol{y})\, d\boldsymbol{\psi}, \tag{1.2}$$

while those of hyperparameters are given by

$$p(\psi_k \mid \boldsymbol{y}) = \int p(\boldsymbol{\psi} \mid \boldsymbol{y})\, d\boldsymbol{\psi}_{-k}, \tag{1.3}$$

which cannot be solved in closed form. Rue et al. (2009) proposed to approximate $p(\boldsymbol{\psi} \mid \boldsymbol{y})$ as

$$p(\boldsymbol{\psi} \mid \boldsymbol{y}) \approx \left. \frac{p(\boldsymbol{y} \mid \boldsymbol{\theta}, \boldsymbol{\psi}) p(\boldsymbol{\theta} \mid \boldsymbol{\psi}) p(\boldsymbol{\psi})}{\tilde{p}(\boldsymbol{\theta} \mid \boldsymbol{\psi}, \boldsymbol{y})} \right|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*(\boldsymbol{\psi})} \equiv \tilde{p}(\boldsymbol{\psi} \mid \boldsymbol{y}), \tag{1.4}$$

where $\tilde{p}(\boldsymbol{\theta} \mid \boldsymbol{\psi}, \boldsymbol{y})$ is the Gaussian approximation of $p(\boldsymbol{\theta} \mid \boldsymbol{\psi}, \boldsymbol{y})$, applied according to the Laplace method (Tierney and Kadane, 1986), and $\boldsymbol{\theta}^*(\boldsymbol{\psi})$is the mode of $p(\boldsymbol{\theta} \mid \boldsymbol{\psi}, \boldsymbol{y})$ for a given $\boldsymbol{\psi}$. The key point is that the Gaussian approximation is applied to a quantity that is generally close to Gaussian, resulting in a high degree of accuracy in the majority of cases.

The approximated marginal distributions $\tilde{p}(\psi_k \mid \boldsymbol{y})$ are obtained, after choosing a grid of values ψ_k, using a numerical integration technique (Martins et al., 2013). For the marginal distributions of the parameters θ_i, Rue et al. (2009) propose three different approaches to compute an approximation of $p(\theta_i \mid \boldsymbol{\psi}, \boldsymbol{y})$: 1) Gaussian approximation; 2) Laplace approximation, and; 3) simplified Laplace approximation. The Gaussian approximation is the easiest to be obtained but it provides poorer results. At the cost of being computationally expensive, Laplace approximation produces better results. The simplified Laplace approximation is a simplification of the last approach and it gives satisfactory results with a good computational time. After choosing one of these methods, the posterior marginal distributions can be computed as

$$\tilde{p}(\theta_i \mid \boldsymbol{y}) \approx \sum_j \tilde{p}(\theta_i \mid \boldsymbol{\psi}^{(j)}, \boldsymbol{y}) \tilde{p}(\boldsymbol{\psi}^{(j)} \mid \boldsymbol{y}) \Delta_j, \tag{1.5}$$

where $\boldsymbol{\psi}^{(j)}$ are specific integration points and Δ_j are the corresponding weights.

In summary, the accuracy and the computational efficiency of INLA are obtained by reformulating the problem in order to use the nested Laplace approximation to a quantity which is generally close to Gaussian, by employing efficient computational techniques for GMRF, and by using appropriate integration methods to retrieve the marginal posterior distributions of interest.

1.3 Gaussian fields in spatial analysis

In general, spatial data can be considered as realizations of a stochastic process indexed in space (Cressie, 1993):

$$\{Y(\boldsymbol{s}) : \boldsymbol{s} \in D\}, \tag{1.6}$$

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 consists of a set of measurements (realizations) of Y taken at a finite set of points $(\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n)$ in D, and the spatial index $\boldsymbol{s}$ can take any value in the continuum in D. The interest, typically, is in inferring the main characteristics of the spatial process such as its mean and variability and predicting values of Y at unobserved points in space using information derived from the analysis of observed data $\boldsymbol{y}$.

An 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) $U(\boldsymbol{s})$, which is a random function for which holds that, for every finite set of points $(\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n)$ (Abrahamsen, 1997),

$$\boldsymbol{u} \sim \mathcal{N}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \tag{1.7}$$

where $\boldsymbol{u} = \{u(\boldsymbol{s}_1), \ldots, u(\boldsymbol{s}_n)\}$ is a realization of $U(\boldsymbol{s})$ 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 $\boldsymbol{\Sigma} = (\Sigma_{i,j})$, $i, j = (1, \ldots, n)$, 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

$$\Sigma_{ij} = \mathrm{Cov}_M(u(\boldsymbol{s}_i), u(\boldsymbol{s}_j)) = \frac{\sigma_u^2}{\Gamma(\nu)2^{\nu-1}}(k||\boldsymbol{s}_i - \boldsymbol{s}_j||)^\nu K_\nu(k||\boldsymbol{s}_i - \boldsymbol{s}_j||), \tag{1.8}$$

where σ_u^2 is the marginal variance of the process, $\nu > 0$ is the smoothing parameter, $k > 0$ is a scale parameter, $||\boldsymbol{s}_i - \boldsymbol{s}_j||$ is the euclidean distance between $\boldsymbol{s}_i$ and $\boldsymbol{s}_j$ and K_ν is the modified Bessel function of second kind and order $\nu > 0$. Instead of the parameter k, for better interpretability we generally consider the range parameter r, i.e. the distance such that the spatial correlation between two points is very

small (about 0.14). The range is defined empirically by the following relationship: (Lindgren and Rue, 2015):

$$r = \frac{\sqrt{8\nu}}{k}. \tag{1.9}$$

1.4 The SPDE approach

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 the Equation (1.8) 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 Markov chain Monte Carlo algorithms, which require these calculations at each iteration, making the task even more difficult (in this regard, see Banerjee et al., 2003, p. 387). An approach to overcome this problem, which is the one that was applied in this study, has been developed by Lindgren et al. (2011). An extended list of applications of this method in various fields of research can be found in Bakka et al. (2018).

The method consists in representing a GRF, which is a continuous spatial process, by approximating it to a spatial process with discrete index (i.e. a 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 is mainly based on the following two results.

First, a stationary GRF $U(\boldsymbol{s})$ 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):

$$(k^2 - \Delta)^{\alpha/2}(\tau_u u(\boldsymbol{s})) = W(\boldsymbol{s}), \tag{1.10}$$

where $\boldsymbol{s} \in \mathbb{R}^d$, $\Delta = \sum_{i=1}^{d} \delta^2/\delta \boldsymbol{s}_i^2$ is the Laplacian operator, α is a smoothing parameter, k is the scaling parameter of the Matérn function, τ_u controls the variance and $W(\boldsymbol{s})$ is a Gaussian *white noise* spatial process. The connection between Equation (1.10) and the parameters of the Matérn function is as follows:

$$\nu = \alpha - d/2 \tag{1.11}$$

$$\sigma_u^2 = \frac{\Gamma(\nu)}{\Gamma(\alpha)(4\pi)^{d/2}\, k^{2\nu}\, \tau_u^2} \propto \frac{1}{\tau_u^2}. \tag{1.12}$$

Second, an approximate solution of Equation (1.10) can be calculated by interpolation using the finite element method (FEM) (Ciarlet, 1978). In short, it consists in dividing the domain into a set of non-intersecting triangles, which can also be irregular; the vertices of the triangles are called *nodes*, and each node corresponds to a base. The solution of the SPDE and its properties depend on the choice of bases. The approximation is as follows:

$$\tilde{u}(\boldsymbol{s}) = \sum_{j=1}^{m} \phi_j(\boldsymbol{s}) w_j, \tag{1.13}$$

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. This discretization of the study region is called *mesh*. The Equation (1.13) can be rewritten as

$$\tilde{u}(\boldsymbol{s}) = \sum_{j=1}^{m} A_j(\boldsymbol{s}) w_j, \tag{1.14}$$

where $A_j(\boldsymbol{s}) = \phi_j(\boldsymbol{s})$ 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. The projector matrix $\boldsymbol{A}$ contains the basis function value for each basis, with one basis function at each column. For example, if the mesh includes only observation locations as mesh vertices, the matrix $\boldsymbol{A}$ has one nonzero value (equal to 1) for each row. Conversely, if a spatial location $\boldsymbol{s}_i$ is placed inside a triangle, matrix $\boldsymbol{A}$ has three nonzero elements on the ith row whose sum is equal to 1.

It can be shown that it is possible to specify an a priori distribution for the weights $\boldsymbol{w}$ in Equation (1.13) such that $\tilde{u}(\boldsymbol{s}) \approx u(\boldsymbol{s})$ in the distribution; for example, for $\alpha = 2$ such an a priori distribution is

$$\boldsymbol{w} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\mathcal{Q}}_{\boldsymbol{w}}^{-1}) \tag{1.15}$$

$$\boldsymbol{\mathcal{Q}}_{\boldsymbol{w}} = \tau_u^2(k^4\boldsymbol{C} + 2k^2\boldsymbol{G} + \boldsymbol{G}\boldsymbol{C}^{-1}\boldsymbol{G}), \tag{1.16}$$

where $C_{ii} = \int \phi_i(\boldsymbol{s})\, d(\boldsymbol{s})$ and $G_{ij} = \int \nabla\phi_i(\boldsymbol{s})\nabla\phi_j(\boldsymbol{s})\, d(\boldsymbol{s})$ (∇ is the gradient).
So, $\tilde{u}(\boldsymbol{s})$ is a GMRF with distribution $\mathcal{N}(\boldsymbol{0}, \boldsymbol{\mathcal{Q}}_{\boldsymbol{w}}^{-1}(\tau_u, k))$. Since $\boldsymbol{\mathcal{Q}}_{\boldsymbol{w}}$ is a sparse matrix, the hyperparameters τ_u and k can be effectively estimated, within the Bayesian framework, using INLA (Rue et al., 2009). For a GMRF model in $\mathbb{R}^2$, the computational cost is typically $O(n^{3/2})$ (Lindgren et al., 2011), which is a significant improvement over the $O(n^3)$ cost of the Gaussian field (GF). Generally, for better interpretability, the estimation results are expressed in terms of σ_u^2 and r. The

approach just described provides a discrete representation (GMRF) of a continuous process (GRF), while preserving the continuity of the correlation structure and taking into account the exact location of the points in space.

For instance, let $y(\boldsymbol{s})$ denote the observed value of the response y measured at location $\boldsymbol{s}$. We can assume that

$$y(\boldsymbol{s}) = \beta_0 + \boldsymbol{x}(\boldsymbol{s})\boldsymbol{\beta} + u(\boldsymbol{s}) + \epsilon(\boldsymbol{s}), \tag{1.17}$$

where β_0 is a regression intercept, $\epsilon(\boldsymbol{s}) \sim \mathcal{N}(0, \sigma_\epsilon^2)$ is the measurement error, $\boldsymbol{x}(\boldsymbol{s})$ is a vector of covariates of interest with corresponding vector of regression coefficients $\boldsymbol{\beta}$, and $u(\boldsymbol{s})$ is the value of the spatial field $\boldsymbol{u} \mid (\sigma_u^2, r)$ at location $\boldsymbol{s}$. The spatial field $\boldsymbol{u}$ is represented by its GMRF approximation $\tilde{\boldsymbol{u}}(\boldsymbol{s})$ as described in Section 1.4. We assume a Gaussian prior distribution for parameters β_0 and $\boldsymbol{\beta}$, and the GMRF prior in Equation (1.15) is assigned to $\boldsymbol{u}$ and $\boldsymbol{v}$. Hence, the latent field $\boldsymbol{\theta} = (\beta_0, \boldsymbol{\beta}, \boldsymbol{u})$ is jointly Gaussian distributed with hyperparameters vector $\boldsymbol{\psi_1} = (\sigma_u, r)$. The observations y are assumed to be independent given $\boldsymbol{\theta}$ and $\boldsymbol{\psi_2} = \sigma_\epsilon$. Denoting the vector of all the hyperparameters as $\boldsymbol{\psi} = (\boldsymbol{\psi_1}, \boldsymbol{\psi_2})$, the joint posterior distribution is then

$$p(\boldsymbol{\theta}, \boldsymbol{\psi} \mid \boldsymbol{y}) \propto \prod_{i=1}^{n} p(y_i \mid \boldsymbol{\theta}, \boldsymbol{\psi}) \times p(\boldsymbol{\theta} \mid \boldsymbol{\psi}) \times p(\boldsymbol{\psi}), \tag{1.18}$$

where index i denotes the generic spatial point $\boldsymbol{s}_i$. The model just described is a three-level hierarchical model with a latent Gaussian structure, hence the posterior marginal distributions for each component of $\boldsymbol{\theta}$ and $\boldsymbol{\psi}$ can be efficiently estimated using INLA (for further details see, e.g., Rue et al., 2009, 2017).

1.5 An INLA-SPDE analysis of temperature effects on Central Mediterranean ecosystem

In this section a novel application of the INLA-SPDE modeling approach is described, which is aimed at investigating the effect of climate change on the faunal composition of the Central Mediterranean Sea area, using trawl surveys biomass indices. Examples of applications of the INLA-SPDE approach in the field of marine ecology can be found in Bleuel et al. (2021); Cosandey-Godin et al. (2014); Lezama-Ochoa et al. (2020); Munoz et al. (2013); Pennino et al. (2014). To the best of our knowledge, this is the first time that this methodological approach is applied to investigate the relation between climate change and thermal affinity of marine

communities.

The content of this section was first introduced in Rubino et al. (2021) and than extended in Rubino et al. (2024).

1.5.1 Motivation of the study

Human activities are deeply impacting the global climate with strong consequences to the biotic and abiotic components of the ocean ecosystems (Hoegh-Guldberg and Bruno, 2010). Also, climatic phenomena and global warming are recognized to be the main drivers for sea temperature increase (Levitus et al., 2000). Species may respond to ocean warming by modifying their vertical distribution (Dulvy et al., 2008) and/or latitudinal range (Perry et al., 2005). Over the years, all of these changes due to ocean warming led to profound and documented impacts on fisheries due to an expected increase of warm-water species in fish communities and, in turn, in fisheries catch (Baptista et al., 2014; Cheung et al., 2013, 2010).Therefore, assessing the effects of climate change on fisheries is one of the challenges for their sustainable management (Leitão et al., 2014). The mean temperature of the catch (MTC) is a well-recognized index to assess the effects of climate change, especially in terms of warming, on fisheries catches (Cheung et al., 2013). Particularly, the MTC, computed from the average inferred temperature preference of exploited species, weighted by their biomass in the catch, was shown to correlate with trend in Sea Surface Temperature (SST) (Cheung et al., 2013). It was recently used to assess the effect of ocean warming both in large marine ecosystems and at local level: e.g. Eastern Mediterranean (Keskin and Pauly, 2014; Tsikliras et al., 2015; Tsikliras and Stergiou, 2014), Caribbean (Maharaj et al., 2018), Adriatic Sea (Fortibuoni et al., 2015). The Mediterranean Sea has a long history of fishing activity that has led to overexploitation most of the Mediterranean fish stocks (Colloca et al., 2013) increasing their vulnerability to climate variability. In fact, SST is increasing at a higher rate than the global average leading to fast changes in the catch composition of fisheries. Therefore, understanding the spatio-temporal relationship between environmental changes and community variations could be of help for correct resource management. Understanding how climatic variability is affecting MTC is relevant also to predict future changes in fish communities and related fisheries catches under different climatic scenarios. In this regard is important to consider that the response of a target variable, like MTC, to the explanatory variables is frequently not immediate but rather exhibits a delay (Legendre and Legendre, 1998; Olden and Neff, 2001).

1.5.2 The data

The Strait of Sicily is a transition area in the south-central Mediterranean Sea connecting the Western and Eastern Mediterranean sectors. Along the southern coast of Sicily (south Italy), the continental shelf is characterized by two wide and shallow (100 m depth) banks in the western (Adventure Bank) and eastern sectors (Malta Bank), separated by a narrow shelf in the middle part. Recent studies highlighted that this area is a biodiversity hotspot in the Mediterranean Sea including a high diversity and biomass of demersal communities over the offshore detritic bottoms of the Adventure bank (Consoli et al., 2016; Di Lorenzo et al., 2018).

We collected 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.

In the present work, only the hauls located on the continental shelf (depth < 200 m) are considered, assuming that the organisms that inhabit this area have a greater probability of being influenced by changes in the sea temperature (Figure 1.1).

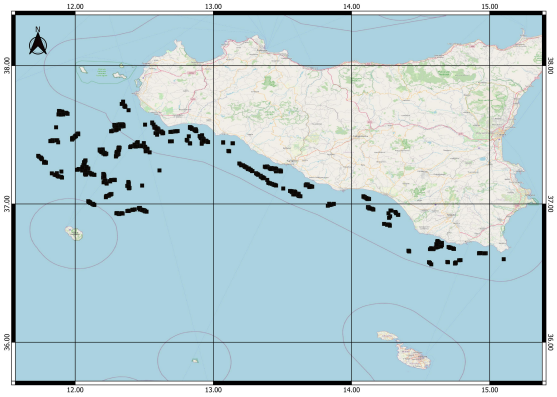


Figure 1.1: Locations of the catch points in the study area.

At each trawl station, fish species are sorted, weighted, counted and measured, and their relative abundance is expressed as kg/km^2. Each species' preferred temperature (median, 25th and 75th percentile) is acquired from the on-line database FishBase (http://www.fishbase.org). The MTC was then calculated for each haul as the average of the temperature preference of all the exploited fish species weighted

by their annual MEDITS catch, that is

$$\mathrm{MTC}_{ht} = \frac{\sum_{i=1}^{n} T_i C_{hti}}{\sum_{i=1}^{n} C_{hti}}, \tag{1.19}$$

where C_{hti} are the catches of species i for year t in haul h, T_i is the median temperature preference of species i and n is the total number of species in the annual catch. An increase of the level of MTC in an area indicates a change in the thermofilic composition of the MEDITS catch, which suggest an increase in dominance of warm-water species in that area (Cheung et al., 2013).

In order to assess the temporal and spatial changes of the MTC in the Strait of Sicily, several factors, which are assumed to be related to the MTC level, were also considered: the annual mean of the SST and the annual mean of the Bottom Sea Temperature (BST, both measured in Celsius degrees); the depth of the catch, categorized in three levels (low: [10-60m], medium: (60-100m], high: (100-200m], based on a hierarchical cluster performed on community bray Curtis dissimilarities matrix); and the spatial (latitude and longitude) and temporal (year) coordinates.

Raster annual maps of SST and BST were constructed by averaging monthly continuous digital maps (downloaded from the website http://marine.copernicus.eu - MEDSEA MULTIYEAR PHY 006 004 - 0.042° × 0.042° pixel resolution). Also, considering that a change in temperature may take time to express its effects on the faunal composition, time lags of one, two and three years were considered in modelling the effects of SST and BST on MTC. Initially, variables related to undersea currents and salinity (downloaded from the website http://marine.copernicus.eu - MEDSEA MULTIYEAR PHY 006 004 - 0.042° × 0.042° pixel resolution) were also taken into account by means of a preliminary analysis, from which, however, there was no evidence of any significant impact of them on the level of the MTC; so, they were excluded from the study.

Figure 1.2 shows the smoothed trend of BST during the time interval 1995-2018, according to each level of depth. The BST trend appears to have almost constantly increased in shallow areas, which are likely to be the most susceptible to changes in external environmental temperature.

1.5.3 The INLA-SPDE model for MTC

Let $y(\boldsymbol{s}, t)$ denote the observed value of the MTC level measured at location $\boldsymbol{s}$ and year $t = 1, \ldots, T$. We assume that

$$y(\boldsymbol{s}, t) = \beta_0 + \boldsymbol{x}(\boldsymbol{s}, t)\boldsymbol{\beta} + u(\boldsymbol{s}) + v_l(t) + \epsilon(\boldsymbol{s}, t), \tag{1.20}$$

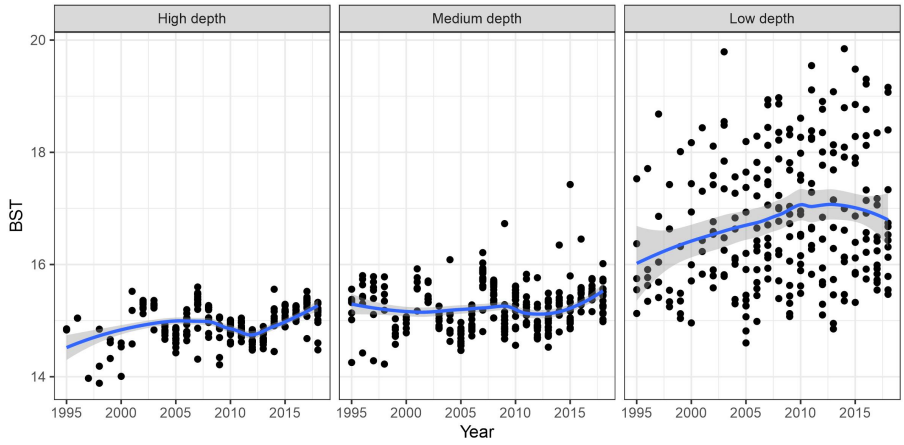


Figure 1.2: Smoothed trend of BST during the time interval 1995-2018, according to each level of depth.

where β_0 is a regression intercept, $\epsilon(\boldsymbol{s}, t) \sim \mathcal{N}(0, \sigma_\epsilon^2)$ is the measurement error, $\boldsymbol{x}(\boldsymbol{s}, t)$ is a vector of covariates of interest (namely, BST and Depth) with corresponding vector of regression coefficients $\boldsymbol{\beta}$, and $u \mid (\sigma_u^2, r)$ is a realization of the Gaussian spatial field at location $\boldsymbol{s}$, which is represented by its GMRF approximation as described in Section 1.4. Moreover, in order to investigate the overall trend of the MTC according to each level l of depth, a random walk $v_l(t)$ of order 1 has been included into the model as a smoothing latent component to take care of possible non-linear relationship, with

$$v_{t+1,l} - v_{t,l} \sim \mathcal{N}(0, \sigma_v^2), \tag{1.21}$$

where σ_v^2, together with σ_ϵ^2, controls the smoothness of the function (a detailed description of the use of random walk models for smoothing methods with INLA can be found in Wang et al., 2018).

Figure 1.3 shows the measurement locations (red dots) within the mesh in two dimensions. The vertices of the triangles represent the nodes (bases) of the Equation (1.13), the number of which is 351. The blue line delimits the region of interest. The addition of the outer edge is to avoid boundary effects (Simpson et al., 2016). Taking into account the accuracy/computational cost trade-off, a low resolution was set for the outer region (which is of no practical interest) while a higher number of nodes was allowed in the study region in order to compute an accurate approximation of the Gaussian field in that area: maximum triangle side length was 0.1 degree for the study region and 0.5 degree for the outer region. Also, in order to avoid the construction of small triangles the minimum triangle side length was set to 0.08 degrees. Details on the mesh construction criteria can be found in Krainski et al.

(2019). The α parameter of the Matérn function (see Equation 1.12) was set equal to 2.

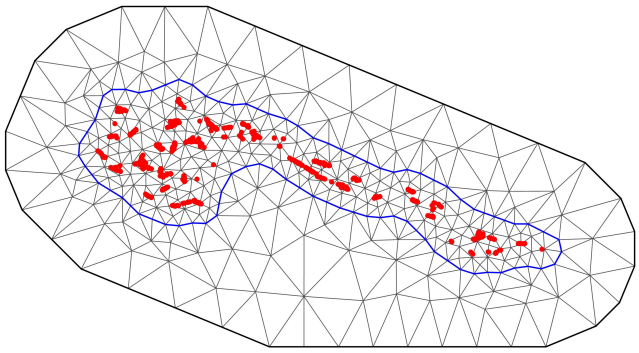


Figure 1.3: Triangulation of the study region, delimited by the blue line, and measurement locations (red dots).

Because no prior information was available, a non-informative zero-mean Gaussian prior distribution was used for the parameters β_0 and $\boldsymbol{\beta}$. Also, the GMRF prior in Equation (1.15) and the random walk prior in Equation (1.21) were assigned to $\boldsymbol{w}$ and $\boldsymbol{v}$, respectively. In practise, R-INLA employs an internal reparameterization of the hyperparameters, based on which their respective prior distributions have been assigned. The hyperparameters of the SPDE component (defined in Section 1.4) are internally reparametrized as $\lambda_1 = log(\tau_u)$ and $\lambda_2 = log(\kappa)$, for which we assumed a joint Normal prior with expected values $\mu_{\lambda_1} = -2.31$ and $\mu_{\lambda_1} = 1.05$, precision $\tau_{\lambda_1} = \tau_{\lambda_2} = 0.1$ and covariance zero. These values are chosen heuristically to provide fairly vague prior information, with the prior median for the spatial range set at a fifth of the approximate domain diameter, and the median prior variance set equal to 1. We also set a log-Gamma distribution with (a, b) parameters as a prior for the precisions $\tau_v = 1/\sigma_v^2$ and $\tau_\epsilon = 1/\sigma_\epsilon^2$, with $a = 1$ and $b = 5 \times 10^{-5}$.

The latent field $\boldsymbol{\theta} = (\beta_0, \boldsymbol{\beta}, \boldsymbol{w}, \boldsymbol{v})$ is jointly Gaussian with hyperparameter vector $\boldsymbol{\psi_1} = (\lambda_1, \lambda_2, \tau_v)$. The observations y are assumed to be Normally distributed and independent given $\boldsymbol{\theta}$ and $\boldsymbol{\psi_2} = \tau_\epsilon$. Denoting the vector of all hyperparameters by $\boldsymbol{\psi} = (\boldsymbol{\psi_1}, \boldsymbol{\psi_2})$, the joint posterior distribution is then

$$p(\boldsymbol{\theta}, \boldsymbol{\psi} \mid \boldsymbol{y}) \propto \prod_{i=1}^{n} p(y_i \mid \boldsymbol{\theta}, \boldsymbol{\psi}) \times p(\boldsymbol{\theta} \mid \boldsymbol{\psi}) \times p(\boldsymbol{\psi}). \tag{1.22}$$

The model just described is a three-level hierarchical model with a latent gaussian

structure, hence the posterior marginal distributions for each component of $\boldsymbol{\theta}$ and $\boldsymbol{\psi}$ can be efficiently estimated using INLA. All the computations were done using the INLA package in R software.

1.5.4 Results of the analysis and discussion

Posterior densities of regression coefficients and hyperparameters, estimated by the INLA-SPDE model, are shown in Figure 1.4 and 1.5, respectively. Their relevant statistics are reported in Table 1.1, along with their 95% HPD (Highest Posterior Density) intervals. In order to assess the statistical significance of an effect (in a Bayesian sense), it can be considered whether its HPD interval contains zero or not. Graphical representations of model results and checks can also be found in Figures 1.6 and 1.7.

Table 1.1: Statistics of the estimated posterior distributions of regression coefficients and hyperparameters.

Coefficient	**Mean**	**SD**	**Median**	**Mode**	**HDI low**	**HDI high**
Intercept	7.66	1.29	7.66	7.68	5.13	10.17
BST lagged (3 years)	0.27	0.08	0.27	0.27	0.10	0.43
Depth medium (ref: high)	1.07	0.19	1.07	1.07	0.69	1.46
Depth low	1.55	0.25	1.55	1.55	1.06	2.04
Hyperparameter	**Mean**	**SD**	**Median**	**Mode**	**HDI low**	**HDI high**
SD of Gaussian observations	1.14	0.03	1.14	1.15	1.08	1.21
SD of SPDE	0.60	0.11	0.59	0.57	0.41	0.82
Range of SPDE	0.22	0.08	0.21	0.19	0.09	0.37
SD of random walk	0.29	0.07	0.29	0.28	0.16	0.43

The estimated posterior mode of the SD of the SPDE random component (0.57°C) suggests that its inclusion into the model allowed to capture some spatial heterogeneity unexplained by the other covariates. This can also be seen by looking at Figure 1.6, which shows the estimated posterior mean of the SPDE component across the study region.

Figure 1.7 shows the estimated Matérn correlation as a function of distance.

Considering the dimension of the rectangle which contains the study area (about 374.6×130.8 km), spatial correlation seems to decrease fairly quickly (the posterior mode of the range is $19^\circ \approx 21.06\,\text{km}$), suggesting a relatively high variability in the spatial distribution of species. Figure 1.8 illustrates the estimated effect of the Random Walk smoothing component, which allowed to capture the non-linear

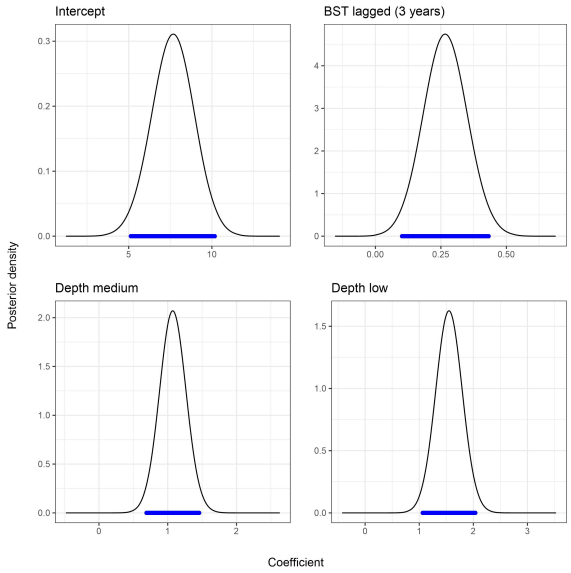


Figure 1.4: Estimated posterior densities of coefficients. Blue segments represent HD intervals.

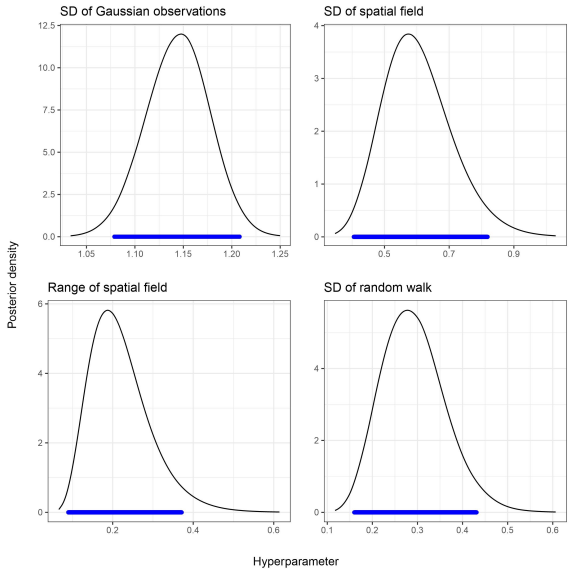


Figure 1.5: Estimated posterior densities of hyperparameters. Blue segments represent HD intervals.

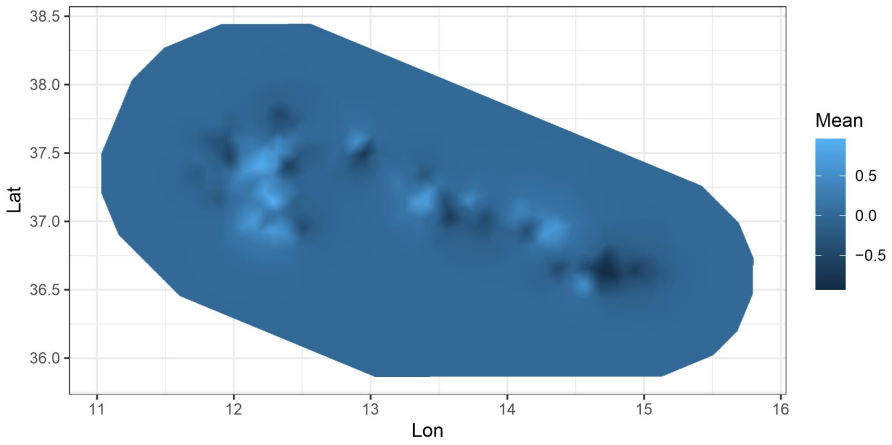


Figure 1.6: Posterior mean values of the spatial field. Spatial heterogeneity, unexplained by the other covariates, appears to be present; the estimated posterior mode of the SD of the field is 0.57. The flat area is the one outside the study region.

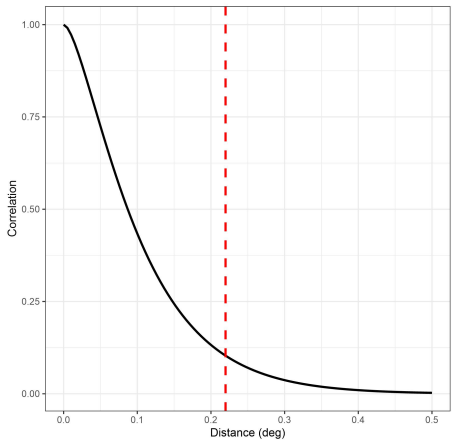


Figure 1.7: Estimated Matérn correlation function. The red dashed line represents the posterior mean value of the range.

temporal evolution of the MTC. This effect also appears to vary across different depth strata.

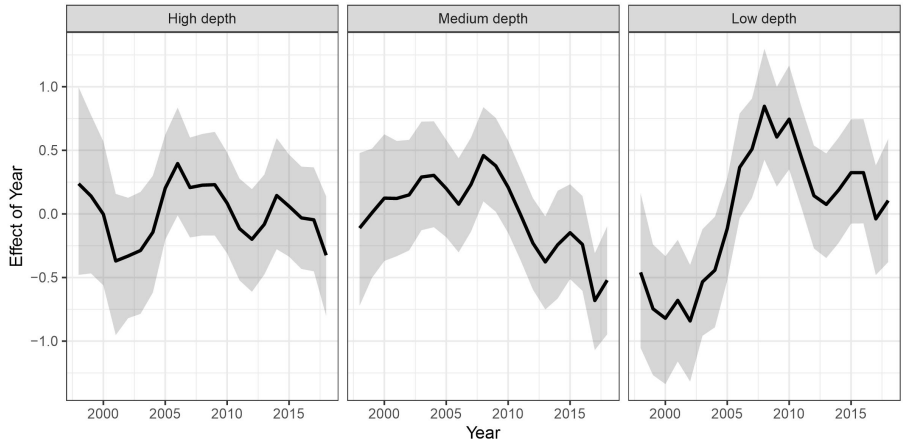


Figure 1.8: Random Walk smoothing of the MTC trend, according to depth level of the catch.

Using BST as the environmental temperature variable, compared with SST led to a decrease in Deviance Information Criterion (Spiegelhalter et al., 2002) (DIC) equal to 6.39. BST turned out to have a significant effect on the MTC level, with the best fit given by its three years lag (DIC values are 2444, 2450, 2443 and 2442 for lag 0, 1, 2, 3, respectively). We also run a 10-fold cross-validation (Hastie et al., 2009) comparing SST and the different lags of BST. We considered the Root Mean Square Error (RMSE) as the out-of-sample performance measure, computed using the differences between posterior means of the MTC and its observed values (RMSE results are 1.2154, 1.2104, 1.2193, 1.2155 and 1.2103 for SST and BST lag 0, 1, 2, 3, respectively). Although the comparison based on the results of DIC and cross-validation showed rather marginal differences, a preference emerged for the inclusion of BST lag 3 in our model. This choice is also substantiated by the understanding that changes in species abundance in response to temperature variations are not typically observable within the same or the subsequent year. This delay is due to the necessary growth period organisms require before they can be effectively captured in nets, considering each species' minimum size of retention. The lag 3 thus reflects a more realistic temporal frame to observe the impacts of temperature on species abundance, acknowledging the biological growth cycles intrinsic to the species under study. In general, the significant effect of BST suggest a positive relationship between environmental temperature variations and change in MTC.

Many marine species, including ectotherms like fish, crustaceans, and mollusks, operate near their upper thermal tolerance, making them highly susceptible to even minor temperature increases (Harley et al., 2006; Helmuth et al., 2005). These

changes can significantly affect their survival, adaptation abilities, biodiversity, and community structure (Doney et al., 2012; Pörtner and Farrell, 2008).

Considering the estimated posterior mode of the coefficient of BST, a one-degree increase of BST results in an increase of 0.27°C of MTC, on average. This value, being less than 1, means that communities that live or will live in warmer waters will be further away from their thermal optimum, and possibly this relationship will be exacerbated in the future. Studies have shown that marine communities living in warmer waters are often farther from their optimal temperature range compared to those living in colder waters (Sunday et al., 2015). This means that in warmer regions, species are often living closer to their upper thermal limits and may be more vulnerable to the impacts of warming, such as reduced metabolic rates, reduced growth, and increased susceptibility to disease and predation.

There is evidence that the response of marine communities to environmental change can have a temporal lag of several years (Poloczanska et al., 2013). Understanding this lag is critical for predicting the long-term impacts of environmental change on marine ecosystems and for developing effective conservation strategies. Our study highlights a time lag also for the central Mediterranean communities, and this could be explained by two possible hypotheses: hypothesis a) over time, also the central Mediterranean communities have adapted to the changes, acquiring a certain resilience and resistance; hypothesis b) thermophilic species are also increasing in our communities, which are more resistant to change.

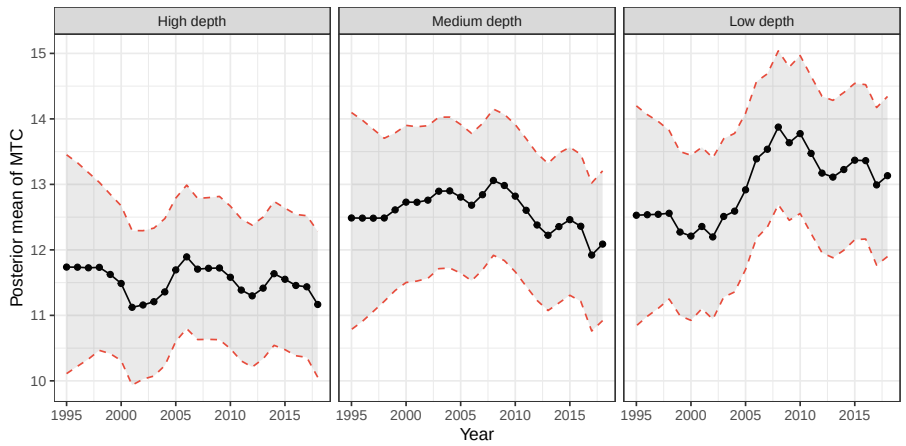


Figure 1.9: Estimated overall trend of the MTC, according to depth level of the catch.

Figure 1.9 shows the estimated overall trend of the MTC, according to each level of depth, which was computed from the joint posterior predictive distribution of the MTC, marginalizing for year and depth strata. To evaluate the reliability of the overall trend estimation, a retrospective analysis has been done by removing 1 to

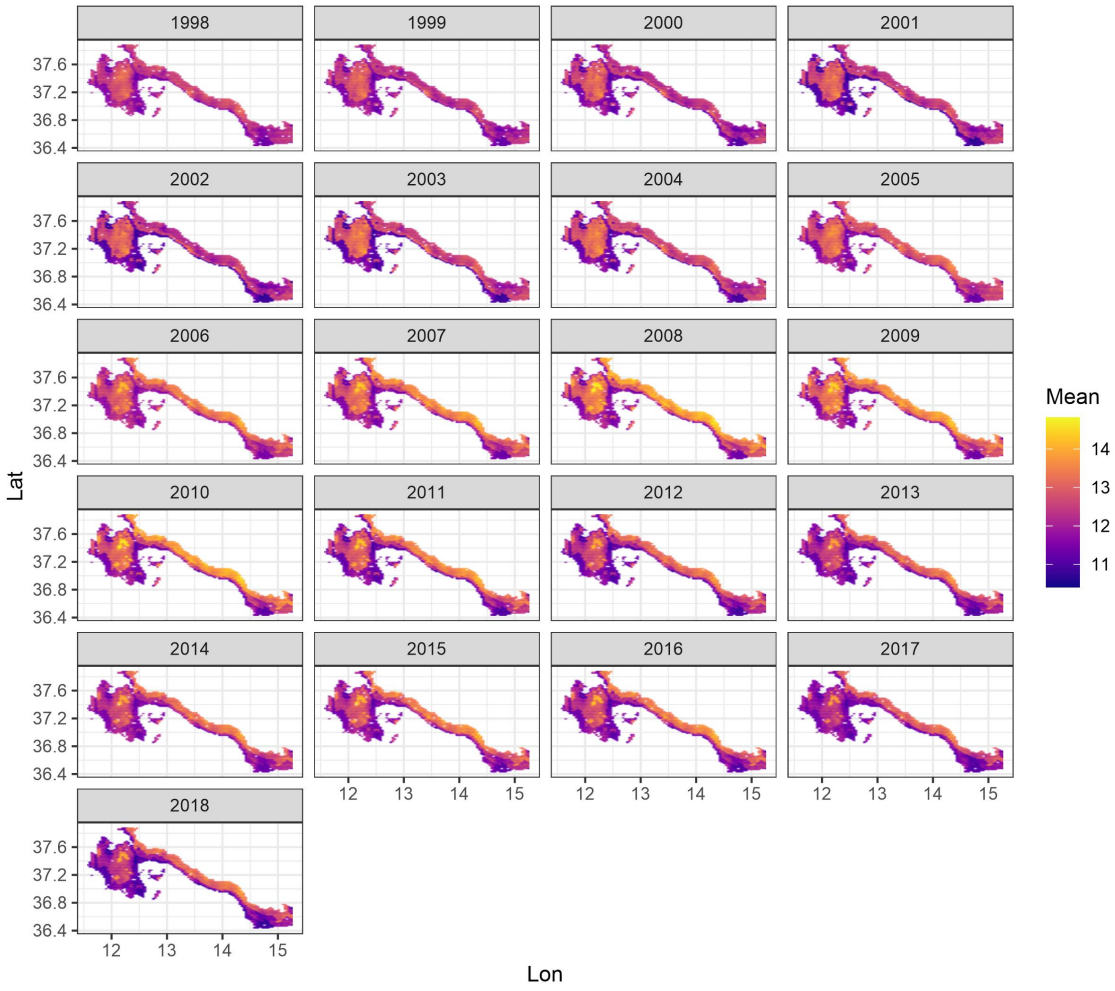


Figure 1.10: Posterior mean of the predicted MTC in the study area, for each year.

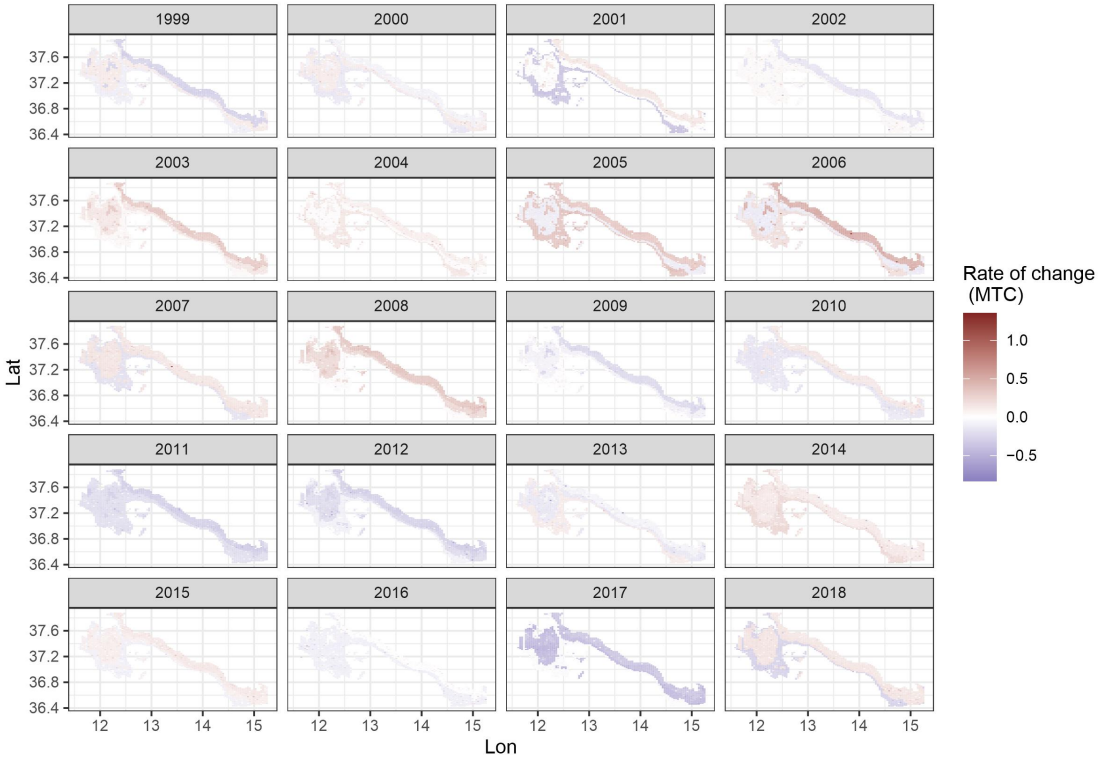


Figure 1.11: Estimated rate of change of the MTC, calculated as the difference between model predictions at t and $t-1$ time values.

4 years from the end of the study period, to see how much the trend estimate was affected by that. Figure 1.12 shows a visual comparison of the estimated curves, including that of the model estimated with the complete dataset (0 years removed), from which it can be seen that trend estimations appears to be fairly stable.

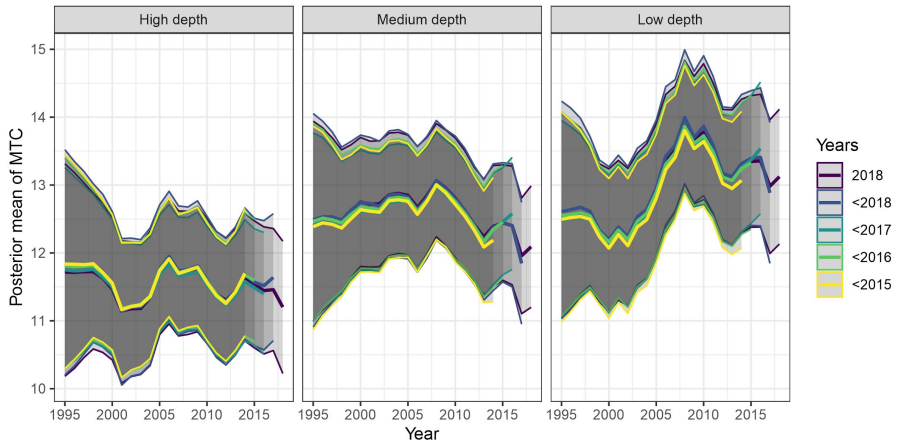


Figure 1.12: Retrospective comparison of the overall trend for 0 (complete dataset) to 4 years removed from the end of the studied period.

Figure 1.10 shows the posterior means of the predicted values of the MTC across the study area (upper and lower bounds of the corresponding credible intervals are represented in Figure 1.14 and 1.13, respectively), excluding point with depth $>$ 200m (hence the white pixels).

Also, to highlight both the direction and the rate of the change of the MTC level, the difference between model predictions at t and $t-1$ time values has been computed for each time point, across the study area (Figure 1.11). In particular, it can be noticed that the MTC increase that began in 2002 has become faster from 2005 to 2006. Than, a quite rapid decrease has happened between years 2010 and 2012. The MTC does not appear to increase linearly; instead, it displays a dynamic characterized by step-like changes (Figure 1.8 - low depth) because the distribution of marine species is not uniform in the ocean (Figure 1.10), and their spatial distribution may vary non-linearly with climate change. The warming of the oceans can lead to the redistribution of marine species, as they move away from their original distribution areas in search of cooler waters or adapt to new climatic conditions. This can lead to a step-like dynamic of the MTC, in which the average temperature of the catches remains relatively stable for a certain period of time and then increases sharply due to the redistribution of marine species.

The estimated overall trend of the MTC level, as is shown in Figure 1.8, suggests a clear increase starting in 2003 in shallow-water area, but not in the medium and deep strata. The observed MTC increase suggests an alteration in the relative

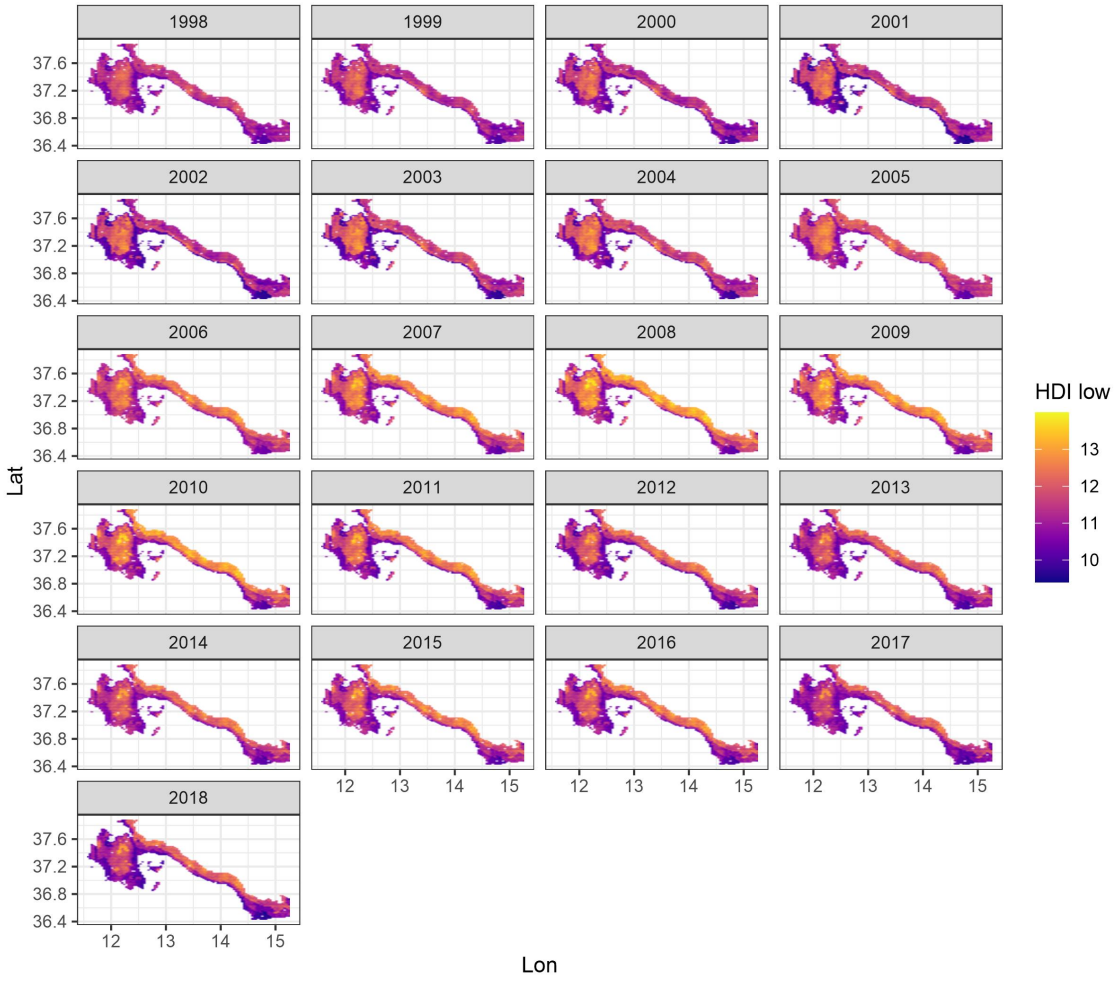


Figure 1.13: Lower limits of the credibility intervals of the predicted MTC in the study area.

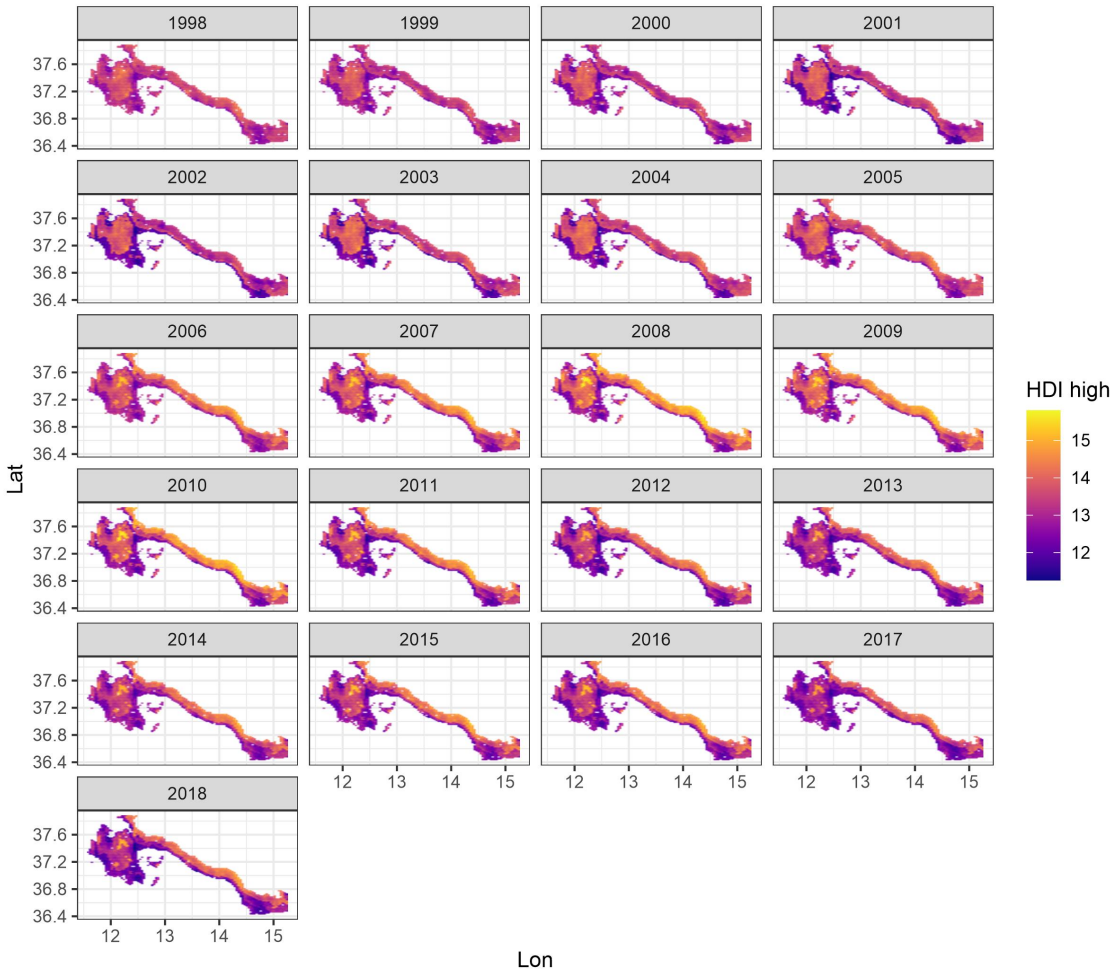


Figure 1.14: Upper limits of the credibility intervals of the predicted MTC in the study area.

catch proportions of species; the thermophilic species (those that prefer warmer temperatures) increased in proportion in the catches over the time series, while psychrophilous (those that prefer colder temperatures) decreased (until 2015). Such change could be due to the displacement of the thermophilous species to a higher latitude and the shift of the psychrophilous species in mean latitude or depth or both.

The MTC significantly decreases with depth, as expected (estimated posterior mode of the coefficients for medium and low depth are 1.07 and 1.55, respectively, with baseline category being high depth): regions characterized by shallow waters are, in general, inhabited by species with higher preferred temperature. Overall, the decrease in MTC with depth highlights the importance of considering depth as a factor in fisheries management and conservation. Understanding the factors that contribute to this pattern can inform the development of more effective management strategies that account for the unique characteristics of different fish communities at different depths.

1.6 Conclusions

In this Chapter, first we described the INLA-SPDE approach to the approximation of Gaussian random fields in spatial analysis. Then, we showed how we exploited the computational efficiency and the modeling flexibility of the INLA-SPDE approach to analyse the effects of environmental temperature on the MTC of demersal fish communities in the central Mediterranean Sea from 1995 to 2018. The proposed hierarchical model, which involves a GF and a stochastic smoother for the trend, allowed us to quantify the effect of ambient temperature (BST) and depth on the MTC, as well as to describe the spatio-temporal evolution of the MTC in the study area. The study emphasises the vulnerability of ectothermic organisms to even slight increases in temperature, as they operate near the upper limits of thermal tolerance. A one-degree increase in BST translates into an average 0.27°C increase in MTC, indicating a shift away from optimal temperature conditions for marine communities. The change in MTC patterns in the study area, particularly since 2003, reflects a shift towards thermophilic species and a decline in psychrophilic species, indicative of a redistribution due to warming waters. The potential of deep-water environments as refugia for species affected by climate change was also explored. Our results highlight the importance of considering depth in fisheries management and conservation. The significant decrease in MTC as depth increases underlines the need for customised strategies to protect different fish communities.

As a direction for future research, it would be beneficial to investigate the response of individual species to climate change in a multivariate statistical framework, in order to identify species that exhibit greater challenges in adapting to changing environmental conditions.

Chapter 2

Approaches to mediation analysis with nonlinear models

2.1 Introduction

In several applied research fields, it is common that the effect of a variable on a response of interest is not entirely direct, but is transmitted by one or more intermediate variables called *mediators*. Mediation analysis is nowadays a widely spread approach to address such settings. It was developed by social scientists Baron and Kenny (1986b), relying on the path-analytic framework developed by Wright (1934) and subsequently further extended to structural equation models (SEM, Bollen, 1989).

The SEM framework assumes linearity; that is, the mediator and the outcome models are assumed to be linear (identity link functions and no interactions). This allows us to estimate the indirect effect of the exposure on the outcome, i.e. the part of the effect conveyed by the mediator, as a product of regression coefficients (Baron and Kenny, 1986b; Bollen, 1989; MacKinnon, 2008a). This approach takes the name of *product method*, and the obtained indirect effect is interpreted as the change in the outcome associated to a one-unit change in the exposure via the mediator.

However, in many real-world applications, linearity may fail to hold because of interaction terms in the models or the presence of not normally distributed variables requiring link functions different from the identity. For instance, consider a situation where researchers are examining the effect of environmental pollution on respiratory diseases (coded as a binary variable indicating the presence or absence of the disease) mediated by the level of physical outdoor activity, which itself might be a categorical variable (such as low, medium, or high activity levels). Another scenario could involve studying the impact of dietary habits on the likelihood of developing type

2 diabetes, mediated by body mass index (BMI), a continuous but not normally distributed variable. In these instances, the assumption of normality cannot hold, with mediators being categorical or non-normally distributed continuous variables, and outcomes being binary.

Despite the widespread of such variables in applied research, mediation analysis with nonlinear models has primarily been addressed using non-parametric estimands of the indirect effect based on counterfactuals, typical of a causal framework (Albert and Nelson, 2011; Doretti et al., 2022; Gaynor et al., 2019). This non-parametric formalisation allows for a certain flexibility in the model specification when one fits parametric models for the mediator and the outcome.

However, in general the estimands in the counterfactual-based approach require the introduction of a different notation and several, sometimes untestable, assumptions for them being identified (i.e. expressed in terms of observed variables), which scholars may not be willing to do. On the other hand, to date, very few approaches have been proposed to deal with the issue of nonlinear mediation models using an associative, path-analytic approaches, typically used in traditional mediation (Rijnhart et al., 2021a,b), and they are based either on standardising coefficients (MacKinnon and Dwyer, 1993), or on the less employed difference method (Schluchter, 2008).

An exception is given by a generalisation of the product method based on partial derivatives, proposed in the '80s by Stolzenberg (1980), and recently revived by Hayes and Preacher (2010) and Geldhof et al. (2018). Although quite intuitive, this approach is not widely known and applied by practitioners, and for this reason it is not well developed, presenting theoretical shortcomings yet to be addressed.

It is worth mentioning that while counterfactual-based approaches to mediation analysis are inherently causal, traditional path-analytic methods can in some situations only be used to test the presence of a mediated effect (VanderWeele, 2015).

In the following Sections, examples of estimating nonlinear mediation models within both the counterfactual-based framework and the associative mediational setting are presented.

In Section 2.2, an application of a counterfactual-based method developed by Imai et al. (2010) is described. The study (Rubino et al., 2022a) is aimed at investigating the effect of socio-economic status on interregional mobility and mortality among cancer patients residing in Sicily (Italy). A novel two-step approach is employed, in which graphical models are initially used as an exploratory tool to investigate the joint structure of relationships between all variables in the system. Thereafter, mediation analysis focused on the estimation of causal effects.

In Section 2.3, we delve into the derivative-based approach: first, we describe the method; then we move to some of its aspects which have received little attention from practitioners, and we believe should be deepened, also carrying out a simulation study addressing some of these issues; finally we present an application, examining the relationship between high school background and academic success. The content of this Section is taken from Rubino et al. (2022b) and Di Maria et al. (2024).

2.2 Socio-economic inequality, interregional mobility and mortality among cancer patients: A counterfactual-based approach

2.2.1 Motivation of the study

Interregional mobility involves a patient travelling to another region of their county of residence for the purpose of receiving planned healthcare. The phenomenon of extra-regional hospitalizations in Italy is remarkable: extra-regional hospitalizations number almost 1 out of 10, involving approximately $730,000$ patients in 2018 (Ministero della Salute, 2020). The most significant pattern of patient mobility involves predominantly southern regions, which display the highest negative mobility balances. This implies that southern regions pay a higher amount for hospital care services used by residents outside their region of residence than the amount received as reimbursement from other regions (Balia et al., 2014). This, therefore, generates additional financial flows in favour of central-northern areas in Italy (Berta et al., 2021).

Moreover, the travel burden, associated with a long-distance trip, can exacerbate inequalities in access to healthcare services. Consequently, the topic of patient mobility regards considerations of health disparities. Indeed, the association between socio-economic status and health has been widely demonstrated, and, nowadays, the adjustment for socio-economic condition is routine in epidemiological analyses (Adler et al., 1994; Pickett and Pearl, 2001). Nonetheless, while geographical differences in cancer treatment are well documented (Stitzenberg et al., 2009), we contend that the relationship between socio-economic status, patient mobility and health outcome has not yet been suitably investigated.

From a methodological point of view, the majority of studies regarding patient mobility explore the issue according to a framework of *choice models*. Accordingly, standard conditional logit models (Aggarwal et al., 2020; Verevkina et al., 2019),

or their variations, such as the heteroscedastic conditional logit model (Balia et al., 2014) or multilevel logistic regression (Aggarwal et al., 2019), have been used in the majority of cases. Other approaches, also in relation to the response variable under consideration, have used: partial proportional odds models for ordinal response variables (Meleddu et al., 2020), quantile regression (Knisely et al., 2020; Propper et al., 2007), and Cox proportional hazard modeling, in the context of survival analysis (Moten et al., 2020; Turner et al., 2020). Within the general topic of survival from cancer analysis, a growing body of literature is beginning to apply mediation analysis to evaluate the direct and indirect effects of socio-economic status on health outcomes (Bedir et al., 2021; Li et al., 2013, 2016; Russell et al., 2020). Nonetheless, to the best of our knowledge, none of these studies has jointly considered the effects of socio-economic status and patient mobility on health outcomes within a framework of mediation analysis.

In the analysis outlined here, patient-level data regarding regional and extra-regional hospitalizations in cancer patients residing in Sicily (Italy) have been used to evaluate 3-year mortality. This research involves two different cancer sites by type, namely: i) the colon, and ii) the trachea, bronchus and lungs (hereafter abbreviated to TBL). By combining the two most common approaches to health disparities discussed in the literature, namely equal opportunities and equal outcomes (Abatemarco et al., 2020), this study will investigate the complex relationship among patients' characteristics (including age, comorbidities and socio-economic status), interregional mobility, and mortality among cancer patients. Moreover, this approach will initially make use of graphical models to discover the relationship structure among the considered variables. And, secondly, mediation analysis will be used to measure the direct and indirect effect of socio-economic status on mortality.

2.2.2 Data

Patient-level data from the hospital discharge records (SDO) of patients residing in Sicily were made available from the Epidemiology Department of the Sicily Region. The analysis included all those living in Sicily who had been diagnosed with one of two selected cancer sites, namely: the colon (ICD-9-CM: 153, 154, or 1590), and TBL (ICD-9-CM: 162). These two cancer sites are those which are responsible for the highest share of male and female cancer-related deaths in Sicily, and the mobility of these patients has been described by some scholar as "remarkable" (Cislaghi et al., 2013). Values of extra-regional hospitalizations of approximately 7% for colon cancer and 7.6% for TBL in 2020 have been observed. And, despite other cancer sites presenting higher fatality rates involving a higher share of mobility (e.g. brain

and nerve cancer), they were not included in the analysis due to their much lower incidence among the population being analysed.

These patients had been hospitalized within and beyond Sicily between 1 January 2010 - 31 December 2012. Further selection criteria stipulated that patients had not been hospitalized for cancer in the previous seven years, and that each patient had only one type of cancer. The SDO dataset was also used to determine whether the hospitalization occurred in- or extra-region and other patient-level characteristics, including sex, age, and the Charlson comorbidity index (Charlson et al., 1994). *Mobility* was defined as a dummy variable; it had value of '0' if the first hospitalization event occurred in Sicily and '1' if it occurred in another region of the country. Only the first hospitalization event was considered during the 3-year follow-up period.

Information relating to a 3-year mortality period was obtained from the Regional Register of Causes of Death (in Italian: ReNCaM). And, the availability of data regarding the census tract of residence for those patients residing in municipalities with more than 10,000 inhabitants permitted the inclusion of the deprivation index (Caranci et al., 2010) pertaining to that census tract of residence. In line with other proposals at the international level (Shavers, 2007), this index is a composite indicator which considers five dimensions of deprivation, namely: the share of population with at least primary level education, the proportion of unemployed people, the share of families with one parent and dependent offspring living together, the proportion of rented households, and the number of people by $100m^2$ at the census tract territorial level. A categorization of the deprivation index based on quantiles of the cohort distribution under analysis has been used in this study.

The cohort being analysed comprised $6,667$ incident episodes of hospitalization of patients, who had been diagnosed with colon ($n = 3,912$), and TBL ($n = 2,755$) cancer. Approximately 5.5% of the patients under investigation were hospitalized out-of-their region: 5.1% for colon cancer and 6% for TBL cancer. The 3-year case-fatality rate was approximately 51.7%, ranging from 34.4% for colon cancer to 76.3% for TBL cancer.

The distribution of factors potentially associated with patient mortality and the corresponding case-fatality rate for the cancer sites considered are reported in Table 2.1. The 3-year case-fatality rate increased with age and the Charlson index category for both cancer sites; lower values of the 3-year case-fatality rate were also found for patients who had been hospitalized extra-regionally compared to those who had not. The case-fatality rate seemed to be influenced by sex only for TBL patients (0.70 for females vs 0.78 for males); generally an increase in the deprivation level appeared to be associated with case-fatality rates.

Table 2.1: Distribution of patient-level characteristics and 3-year case-fatality rate by cancer site. The *Colon* column collates the descriptive statistics for the patient cohort with colon cancer; the *TBL* column collates the statistics of the patient cohort with trachea, bronchus and lung cancer

	Colon		**TBL**	
	n	Case-fatality rate	n	Case-fatality rate
Deprivation				
1 (Low)	978	0.33	689	0.73
2	978	0.32	689	0.78
3	980	0.35	688	0.77
4 (High)	976	0.38	689	0.78
Mobility				
0 (No)	3713	0.35	2589	0.77
1 (Yes)	199	0.21	166	0.61
Charlson				
$[0, 2]$	2081	0.21	794	0.65
3	594	0.30	468	0.65
$[4, 6]$	347	0.43	355	0.80
>6	890	0.66	1138	0.87
Sex				
Female	1917	0.34	631	0.70
Male	1995	0.34	2124	0.78
Age				
<60	794	0.23	569	0.67
(60,80]	2366	0.31	1791	0.77
>80	752	0.56	395	0.86

2.2.3 Methods

The methods employed in the analysis will be described in this Section. First, graphical models were used as an exploratory tool to investigate the joint structure of relationships between all variables in the system. Thereafter, mediation analysis focused on the estimation of causal effects, which were related to the subset of variables of main interest, namely: deprivation index, out-of-region mobility and 3-year mortality.

Graphical models

A graphical model is a statistical technique whose rationale and philosophy can be described within the framework of parametric statistical modelling (Lauritzen, 1996; Whittaker, 2009). Its main advantage is that the conditional independence among a set of random variables from a graph composed of nodes (representing the random variables) and a collection of links (representing the dependent association) can be ascertained.

Formally, a graph is an object $G = \{V, E\}$, where $V = \{V_1, \ldots, V_d\}$ is a finite set of nodes, and $E \subseteq \{\{V_i, V_j\} : (V_i, V_j) \in V^2\}$ is a finite set of links between the nodes. Undirected graphs assume unordered pairs, and thus the edges are visualized by lines, whereas directed graphs consider arrows to visualize ordered pairs and the edges. Undirected graphs where the set of random variables is discrete will be the focus in this study.

Let $Y = (Y_1, \ldots, Y_d) = (Y_v)_{v \in V}$ be a set of discrete random variables, observed on n statistical units, where the total number of units has been fixed. Write the number of levels of Y_v as card$|Y_v|$. Write a generic observation (or cell) of the contingency table as $i = (i_1, \ldots, i_d)$, and the set of possible cells as I. Express the raw case list in an aggregated list by adding a column representing the frequencies of each of the levels of the random variables and suitably aggregating the levels. Denote the observed frequency with $n(i)$ and its r.v. with $Y(i)$. Given a graph G and the joint probability of the observed contingency tables

$$\Pr(\{n(i)\}_{i \in I} | \{\pi(i)\}_{i \in I}) = \frac{N!}{\prod_{i \in I} n(i)!} \prod_{i \in I} \pi(i)^{n(i)}, \tag{2.1}$$

this model belongs to the class of discrete graphical models if the joint probability can be expressed as a product of maximum cliques (a subset of nodes forming a complete graph). This fact implies that pair, local and global Markov properties can be directly read from the graph.

The parameter $\pi(i)$ in (2.1) must be estimated. If the probabilities are not restricted in any way (except requiring that they are non-negative and sum to one), then it can be easily demonstrated that the maximum likelihood estimates are given by $\hat{\pi}(i) = n(i)/N$ for $i \in I$. However, such an unrestricted model would imply a complete graph (i.e. all nodes connected). Generally, it is of interest to identify a more parsimonious graph, which represents the complexity of the data. Decomposable graphical models are a subclass of graphical models, for which the Maximum Likelihood estimate for $\pi(i)$ in a closed form can be ascertained. Alternatively, iterative algorithms, such as iterative proportional scaling (Denteneer and Verbeek,

1985) and message passing (Wainwright, 2015), are required for non-decomposable models.

Using graphs to represent models shifts the emphasis of the analysis from estimating the parameters for a given model to estimating the model structure, that is, selecting an appropriate graph. Model selection is challenging mainly because the number of possible graphs is huge. Many different methods for selecting graphical models have been proposed and they typically have a threefold classification: low-order conditional independence tests; heuristic search methods which optimize a given criterion; and Bayesian methods, often involving Markov chain Monte Carlo methods (Højsgaard et al., 2012). Heuristic search methods, which optimize an information criterion, have been considered in this study. The Akaike Information Criterion (AIC) assigns the following score $-2\ell + kp$, to each model, where ℓ is the likelihood, p is the number of parameters to be estimated in the model, and $k = 2$ is a penalty parameter. A popular alternative to AIC is the Bayesian information criteria (BIC), which sets k to the logarithm of the number of observations. In general, a larger k penalizes complex models more severely and thus tends to select simpler graphs. The heuristic methods incrementally search by default from an initial graph, adding or deleting the edges which significantly decreases the AIC or BIC: if no edge satisfies this characteristic, the process terminates. The search is also directional: either forward (adding edges) or backward (deleting edges). The choice of k is usually made based on asymptotic considerations.

In order to be able to answer queries (e.g. $P(Y_d|Y_1, Y_2)$), a marginalization must be effected, according to the graph structure. This marginalization is made via the propagation algorithm by working twice through the set of cliques and passing the "messages" between neighbouring cliques: initially from the final clique in the RIP (running intersection property) ordering (Højsgaard et al., 2012) towards the first, i.e. inwards in the junction tree, and subsequently passing messages in the other direction (Dawid, 1992; Green, 2005).

One limitation of the graphical model is that there is no way of measuring indirect effects, whereas the graph assists our understanding of the association structure among variables. Having identified the model structure among the considered variables, a causal mediation analysis within a counterfactual framework can, therefore, be relied upon in a second step to quantify the indirect effect of an exposure X (e.g. *deprivation*) on a response Y (e.g. *mortality*) passing through a mediator M (e.g. *mobility*), controlling for various confounders (e.g. *age*, *sex*, the *Charlson comorbidity index*).

Causal mediation analysis

Mediation analysis can be used to investigate how the effect of an exposure on a certain outcome of interest spreads, whether that be directly or indirectly through a third variable called a *mediator*. Let X, M and Y denote the exposure of interest, the mediator and the outcome respectively. Classical methods of mediation analysis rely on the assumption of a normally-distributed outcome and mediator (Baron and Kenny, 1986a; MacKinnon, 2008b). Indeed, this assumption is not appropriate in our case since the mediator (*mobility*) and the outcome (*mortality*) are discrete. A more general approach to causal mediation analysis has, therefore, been described in this study; this approach was developed within a counterfactual framework, permitting researchers to deal with different kinds of variables and use a wide range of models (Imai et al., 2010).

For each subject i, $M_i(x)$ and $Y_i(x)$ are defined as the counterfactual values of the mediator and the outcome respectively, if the exposure were set to x. Similarly, $Y_i(x, m)$ is the counterfactual value of the outcome for the i-th subject if X were set to x and M to m. For two different values of the exposure, x and x', the (average) total causal effect of X on Y on the difference scale can be defined as $\mathbb{E}[Y_i(x) - Y_i(x')]$, i.e. the expected difference in the outcome under two interventions, one intervention setting X to a treatment value x, the other to a control value x'. In the absence of unobserved confounding and if this expectation differs from zero, then X has a causal effect on Y. If a third variable M mediates the exposure-outcome relationship, the total effect can be decomposed into the natural direct and indirect effects as (Pearl, 2001; VanderWeele, 2015a)

$$NDE = \mathbb{E}[Y_i(x, M_i(x')) - Y_i(x', M_i(x'))] \tag{2.2}$$

$$NIE = \mathbb{E}[Y_i(x, M_i(x)) - Y_i(x, M_i(x'))]. \tag{2.3}$$

The direct effect in Equation (2.2) is the change in response under two different values of the exposure. This change is not mediated by M, which is indeed kept fixed to the natural value that it would assume if X were x'. *Vice versa* the indirect effect is the effect of X on the response only through the mediator: it measures the extent to which the response would change if X were fixed to x, and the mediator to the natural value it would assume under two interventions, one setting X to x, the other to x' (notice that the arguments of the $M(\cdot)$ counterfactuals differ).

Evidently, the only quantities observable for each subject are those corresponding to the actual value of the subject's exposure. That is, if for subject i, $X_i = x$, then $M_i(x) = M_i$ and $Y_i(x) = Y_i$, while $M_i(x')$ and $Y_i(x')$, for any $x' \neq x$, are not

observable. Nonetheless, it can be proved that, under a set of assumptions known as sequential ignorability (Imai et al., 2010; VanderWeele, 2015a), the mediational effects can be identified, i.e. they can be expressed as functions of the observed data. As detailed in Imai et al. (2010) and VanderWeele (2015a), these assumptions entail the absence of unobserved confounders of the exposure-mediator, exposure-outcome and mediator-outcome relationships. Assuming that the set of covariates considered includes all the relevant confounders, identifiability can be deemed to hold.

In the setting described in this study, the focus of interest is the deprivation effect on mortality, which has been measured three years after the initial hospitalization, either direct or mediated by mobility. In order to estimate the mediational effects, the following two logistic models were considered:

$$\text{logit}(\pi_i^M) = \beta_0 + \beta_1 X_i + \boldsymbol{\beta}_2' \boldsymbol{Z}_i \tag{2.4}$$

$$\text{logit}(\pi_i^Y) = \gamma_0 + \gamma_1 X_i + \gamma_2 M_i + \boldsymbol{\gamma}_3' \boldsymbol{Z}_i \tag{2.5}$$

where i denotes individuals, both mobility (M) in Equation (2.4) and 3-year mortality (Y) in Equation (2.5) are assumed to follow a Bernoulli distribution with parameters $\pi_i^M = P(M_i = 1|X, \boldsymbol{Z})$ and $\pi_i^Y = P(Y_i = 1|X, \boldsymbol{Z})$ respectively; X is the deprivation index categorised according to the quartiles, $\boldsymbol{Z}$ is a vector of covariates (age, sex and the Charlson index), and $(\beta_0, \beta_1, \boldsymbol{\beta}_2)$ and $(\gamma_0, \gamma_1, \gamma_2, \boldsymbol{\gamma}_3)$ are regression coefficients.

First, the regression models expressed in Equations (2.5) and (2.4) were fitted and the model parameters estimated, using frequentist GLM approach. Thereafter they were deployed, by means of a quasi-Bayesian Monte Carlo approach, to simulate the counterfactuals of interest relating to two different values of X in order to estimate the effects in Equations (2.2)-(2.3). This algorithm was implemented in the *mediation* package, and a detailed description can be found in the Appendix of the work by Imai et al. (2010).

2.2.4 Results

The results of the graphical model (Section 2.2.3) and the causal mediation analysis (Section 2.2.3), as applied to the two data sets described in Section 2.2.2, will be elaborated in the following. The analyses were performed with the R statistical software, namely the *gRim* package (to estimate the graph), the *gRain* package (to propagate the probabilities) and *igraph* package (to visualize the graph). The *vcdExtra* package was used to visualize the propagated conditional probabilities, and, in order to perform the mediation analysis, the *mediation* package was used.

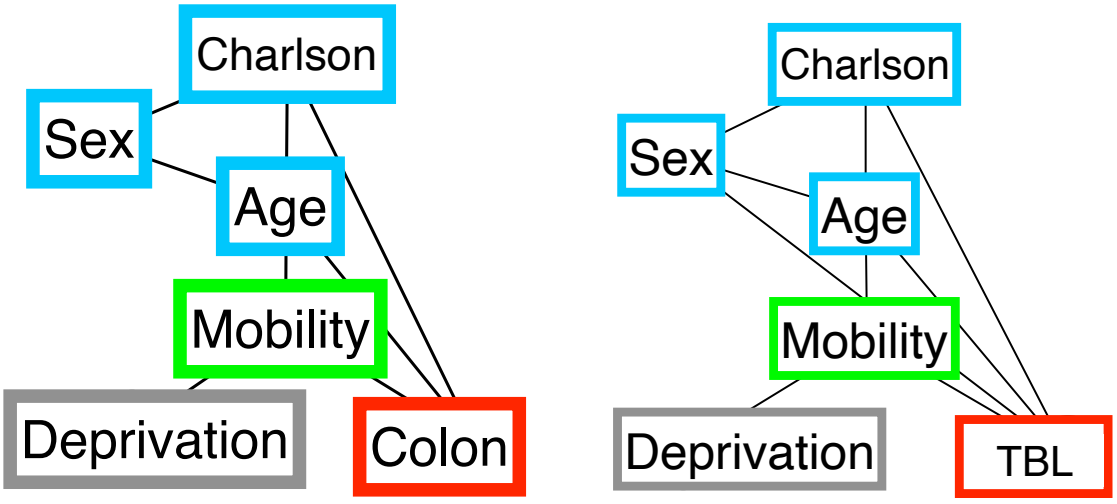
Decomposable Log-linear Graphical Models

A search was conducted for a conditional independence model, within the class of log-linear decomposable graphical models. The decomposable class of models is commonly employed to model conditional independence relationships in graphical models. In such models, closed-form expressions for maximum likelihood estimates exist, which allows to use triangulated graph based, efficient algorithms (Højsgaard et al., 2012). A model selection was made in order to select the optimal graph via a heuristic search algorithm, considering the AIC score. The analysis started with a fully connected graph, from which the links were removed, until the graph still summarized the dataset.

Given three set of variables, let's say A, B and C, if A and B are separated by C in the graph, then $A \perp\!\!\!\perp B \mid C$ in the model. Figures 2.1a and 2.1b show the graphs which were estimated for the patient cohorts who had been diagnosed with colon and with TBL cancer, respectively. These graphs display a similar conditional independence structure. Referring to both cancer sites under consideration, 3-year mortality was observed to be associated with mobility, age and the Charlson index. The only identified difference in the conditional probability structure was the link between sex and 3-year mortality, which only appeared in Figure 2.1b. That is, mortality seems to be influenced by sex only for TBL cancer patients, consistently with the results in Table 1.

Conditional probabilities estimated after applying the propagation algorithm, conditioning on $age = (60 - 80]$, are represented in Figure 2.2. In analysing these probabilities, the effects of various potential confounders must be taken into consideration. These include: *age*, the *Charlson index* and *sex*. Referring to both cancer sites, advanced age increases the chance of dying three years after the initial hospitalization; the higher the *Charlson index*, the higher the probability of dying, while *mobility* seems to decrease this chance (Figures 2.2a and 2.2b). Moreover, Figure 2.2b reveals no clear trend for the *sex* variable. Having analysed the aforementioned probabilities, the deprivation effect on *mobility* was the subsequent focus. It was less probable for patients characterized by a higher degree of deprivation (the worst economic conditions) to travel outwith their region of residency across all age categories. And *age* also affects *mobility* since younger people typically have more chances to move out of Sicily: the lower the degree of *deprivation*, the higher the probability to move outside the region (Figures 2.2c and 2.2d).

Finally, an analysis was performed on the effect of the *deprivation* and the *Charlson index* on the probabilities of 3-year mortality after initial hospitalization for colon cancer (Figure 2.3a); and after initial hospitalization for TBL cancer (Fig-



(a) (b)

Figure 2.1: (a) Undirected graph, estimated from a log-linear decomposable graphical model for the patient cohort, diagnosed with colon cancer (n=3,912). (b) Undirected graph, estimated from a log-linear decomposable graphical model for the patient cohort diagnosed with trachea, bronchus, lungs (TBL) cancer (n=2,755). The covariates are colored in blue, while the exposure, mediator, and outcome variables are colored in gray, green, and red, respectively.

ure 2.3b). The conditional probabilities in Figure 2.3a have been estimated after the propagation step, conditioning on age $= (60 - 80]$; and the conditional probabilities in Figure 2.3b have been estimated after the propagation step conditioning on *age*$= (60 - 80]$ and *sex* $=$ *male*. The effect of deprivation is visually negligible for both types of cancer, although that on TBL seems slightly more pronounced.

Causal Mediation Analysis

In this Section, the results of the mediator and the outcome regression models will be initially discussed individually. Thereafter, the focus will be on mediation analysis outputs. For any estimate, a significance level of 0.05 was considered. The results of the mediator and outcome models, which were described in Equations (2.4) and (2.5), regarding colon and TBL cancer types respectively, are shown in Table 2.2. This Table reports estimates relating to the exponentiated coefficients (odds ratios), in addition to 95% confidence intervals. Referring to the former (mediator models), the probability of interregional mobility is lower for older patients for both cancer sites, and it is unaffected by *sex*. Health conditions, as measured by the *Charlson comorbidity index*, show a significant (non-linear) effect only for TBL cancer patients. While improved health conditions seem to be positively correlated with interregional mobility, the two *Charlson index* categories that are associated with the lowest probability of *mobility* comprise the two intermediate categories. By

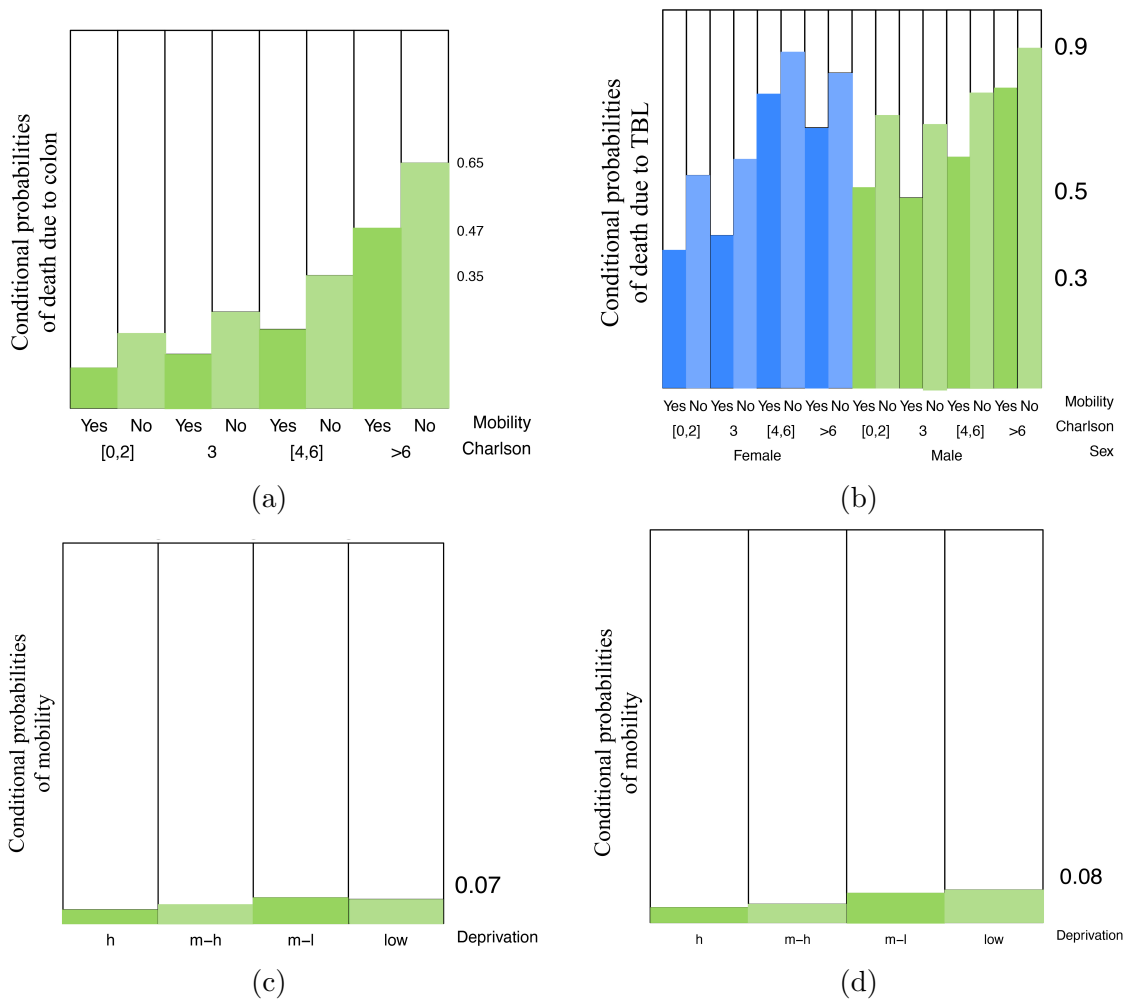


Figure 2.2: Conditional probabilities estimated after applying the propagation algorithm conditioning on $age = (60 - 80]$. Figure (a) shows conditional probabilities of death due to colon cancer, given changes in the levels of *mobility* and *Charlson index*. Figure (b) shows conditional probabilities of death due to TBL cancer, given changes in the levels of *mobility*, *Charlson index* and *sex*. Figures (c) and (d) show the conditional probability of *mobility*, by patient cohort (containing patients who had been diagnosed with colon cancer and TBL respectively) given changes in the levels of *deprivation*. The numbers on the right y-axes indicate the value of the conditional probabilities.

considering the effect of socio-economic conditions on interregional mobility, a significant association between the *deprivation index* and *interregional mobility* emerges for both cancer sites: those patients residing in areas characterised by higher levels of deprivation are less likely to move outwith the region. The results of the outcome models suggest that the relative risk of 3-year mortality for both cancer types is significantly lower for patients who experienced interregional mobility, compared with the risk of those who did not. The effect of *sex* is significant only for the TBL site, with males constituting the highest risk group. For both cancer sites under consideration, the risk of dying significantly increases with *age* and according to the

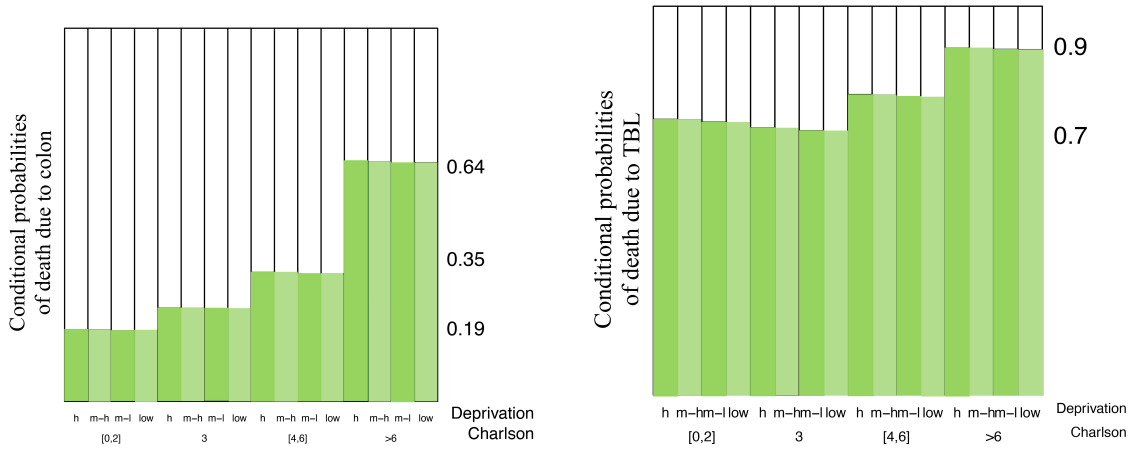


Figure 2.3: Conditional probabilities estimated after the propagation algorithm, conditioning on $age = (60 - 80]$. Figure (a) shows deprivation and the *Charlson index* effect on the probability of dying three years after the initial hospitalization for colon cancer. Figure (b) shows *deprivation* and the *Charlson index* effect on the probability of dying for TBL cancer. Both graphs are conditioned on $sex = male$.

Charlson index. The deprivation effect on mortality, accounting for mobility and the other covariates, proved to be statistically significant and positive only for TBL site.

Based on these findings, there is clearer evidence to suggest that *deprivation* may have a direct impact on the 3-year mortality for TBL cancer compared to colon cancer. This would indicate a differential impact of socio-economic status on cancer survival across the various cancer types (Woods et al., 2006; Yu et al., 2008). However, it is worth mentioning that the non-significance of the deprivation coefficient in the model for colon cancer mortality does not exclude the possibility that the deprivation effect is entirely mediated by mobility; this hypothesis cannot be tested by merely examining the model coefficients. This is the rationale for a mediation analysis, the results of which are reported in Table 2.3.

At this stage in the analysis, interest is focused on the extent to which a change in deprivation status (from less to the most deprived) can affect mortality, either directly and indirectly. Since the response variable (3-year mortality) is binary, the mediational effects can be interpreted as differences in the probability of dying three years after initial hospitalization, comparing two interventions as shown in Equations (2.2) and (2.3). The two interventions considered are a (hypothetical) intervention setting the deprivation status to 1 (less deprived patients) and an intervention setting the deprivation status to 4 (most deprived patients) for all patients.

It can be noticed that the indirect effect of deprivation on mortality through mo-

Table 2.2: Point estimates and 95% confidence intervals for mediator and outcome models effects on colon and TBL mortality.

	Colon		**TBL**	
	Mobility	3y-Mortality	Mobility	3y-Mortality
Deprivation 2 (Ref: 1)	1.087 (0.754, 1.569)	0.963 (0.779, 1.191)	0.839 (0.562, 1.249)	1.363 (1.053, 1.768)
Deprivation 3	0.673 (0.445, 1.010)	1.084 (0.879, 1.337)	0.534 (0.337, 0.832)	1.193 (0.923, 1.544)
Deprivation 4	0.544 (0.350, 0.832)	1.212 (0.984, 1.493)	0.424 (0.258, 0.678)	1.422 (1.095, 1.848)
Mobility Yes (Ref: No)	–	0.516 (0.348, 0.751)	–	0.533 (0.377, 0.758)
Age (60,80] (Ref: <60)	0.730 (0.528, 1.020)	1.674 (1.361, 2.065)	0.724 (0.509, 1.042)	1.661 (1.328, 2.075)
Age>80	0.307 (0.171, 0.526)	5.227 (4.090, 6.706)	0.208 (0.084, 0.439)	3.546 (2.500, 5.091)
Charlson 3 (Ref: [0,2])	0.649 (0.387, 1.034)	1.425 (1.150, 1.761)	0.458 (0.256, 0.777)	0.851 (0.663, 1.093)
Charlson [4,6]	0.767 (0.394, 1.367)	2.037 (1.583, 2.616)	0.450 (0.220, 0.839)	1.576 (1.160, 2.159)
Charlson>6	0.988 (0.693, 1.389)	8.037 (6.705, 9.658)	0.695 (0.489, 0.989)	3.761 (2.984, 4.757)
Sex M (Ref: F)	0.936 (0.701, 1.249)	1.099 (0.947, 1.276)	0.743 (0.524, 1.066)	1.528 (1.234, 1.889)
Intercept	0.104 (0.070, 0.151)	0.128 (0.099, 0.165)	0.228 (0.143, 0.356)	0.764 (0.565, 1.035)
Observations	3,912	3,912	2,755	2,755
Log Likelihood	−765.752	−2,116.610	−597.303	−1,375.842
Hosmer & Lemeshow test P-value	0.826	0.521	0.289	0.501

Note: reported estimates are exponentiated coefficients (odds ratios).

bility is positive and significant for both cancer sites, which is consistent with the results derived from graphical models and with that expected from the output of regression models. A reduced capability to receive medical treatment extra-regionally (due to belonging to the highest rather than the lowest level of deprivation) leads

Table 2.3: Point estimates, 95% confidence intervals and p-values for mediational effects on colon and TBL mortality.

	Colon		**TBL**	
	Effect	P-value	Effect	P-value
NIE	0.0029 (0.0005, 0.0062)	0.0136	0.0054 (0.0014, 0.0110)	0.0024
NDE	0.0347 (-0.0028, 0.0718)	0.0696	0.0586 (0.0153, 0.1009)	0.0080
Total Effect	0.0376 (-0.0001, 0.0747)	0.0508	0.0641 (0.0205, 0.1058)	0.0048

to an increased risk of death from colon and lung cancer of 0.29 and 0.54 percentage points respectively. The presence of a significant positive direct effect only for TBL, which was not detected by graphical model shown in Figure 2.1 (b), can also be noticed. The absence of a link in a graph is compatible with a non-null causal effect since the information conveyed by the two methodologies (graphical models and causal mediation analysis) is different, albeit complementary. As previously mentioned (at the beginning of Section 4), graphical models can be used in an exploratory way in order to estimate the joint density of variables in the system, and graphs display the conditional (in)dependencies among these variables. This approach can be said to be 'associational' and it can be applied to a sample which is naturally stratified by the covariates. Such an approach will take into account, for example, whether a group of subjects is more deprived, while other individuals display lower levels of deprivation. In contrast, the approach used to perform the mediation analysis is causal, and this implies the conceptualisation of an intervention throughout the entire sample, setting the deprivation status of *all* individuals to 4 and 1 (the deprivation categories chosen for the estimation of effects) in turn. The focus of causal mediation analysis is not, therefore, on conditional independencies but on the expected difference between the (simulated) potential outcomes corresponding to the two interventions.

The total effect of *deprivation* on *mortality* relating to TBL cancer is positive and significant. The p-value which is linked to the total effect of deprivation on colon cancer mortality is almost at the edge of the significance level. In such cases (where one of the mediational effects is significant but the total is not, or its significance is doubtful), following Tingley et al. (2014) (page 7), a joint interpretation of all

effects is advisable. Examining the results, the significance of the direct effect of deprivation on mortality due to colon cancer seems rather weak, while that of the indirect effect, mediated through mobility, appears more pronounced. In contrast, in the case of TBL cancer, the indirect effect of deprivation also appears to be significant; however, the evidence of a direct impact of deprivation on mortality is more clear.

This difference may not be surprising to the reader since it has been observed that socio-economic differences can act in a variety of ways on cancer survival. Some of these ways include: tumour characteristics, patient characteristics and healthcare factors (Woods et al., 2006). For example, differences in the disease stage at diagnosis are one of the ways in which socio-economic status can play a role (Yu et al., 2008). In this regard, Halpern et al. (2008) have demonstrated that cancer patients in the United States who were uninsured or inadequately insured were significantly more likely to be diagnosed at an advanced stage for several major cancers. This consideration may also provide an explanation to the differences observed regarding the significance of the direct effect between TBL and colon cancer. In addition, Booth et al. (2010) analysed the impact of socioeconomic status on stage of cancer at diagnosis and survival in Ontario (Canada): they reported that differences among social groups in stage of disease at diagnosis have an effect on survival disparity for breast cancer but not other disease sites, the latter including the colon.

2.3 The derivative-based approach to nonlinear mediation models: insights and applications

The traditional associational framework, in its most basic specification, involves three variables: an exposure X, a mediator M and an outcome Y. Variables M and Y are continuous and are modelled as linear

$$\mu_{_M} = \beta_0 + \beta_1 X \tag{2.6}$$

$$\mu_{_Y} = \gamma_0 + \gamma_1 X + \gamma_2 M, \tag{2.7}$$

where $\mu_{_M}$ and $\mu_{_Y}$ denote the conditional expectations $\mathbb{E}[M \mid X]$ and $\mathbb{E}[Y \mid X, M]$, respectively.

The coefficient γ_1 represents the direct effect of X on Y, while, as discussed in Baron and Kenny (1986b) and Bollen (1989), the indirect effect can be estimated as the product $\beta_1\gamma_2$, i.e. the product of regression coefficients lying on the $X \to M$ and $M \to Y$ paths. Alternatively, considering the marginal model for the outcome,

i.e. the model including only the exposure

$$\mu_{_Y} = \alpha_0 + \alpha_1 X, \tag{2.8}$$

the indirect effect can also be estimated as the difference between α_1, the total effect of X on Y, and the direct effect γ_1, that is $\alpha_1 - \gamma_1$. This approach is called *difference method*, and it is easy to prove that, in the linear case, the product and difference methods yield the same indirect effect estimate (MacKinnon, 2008a).

In the presence of nonlinearities, such as interaction terms or link functions different from identity, the indirect effect cannot be estimated as a simple product, and its value generally differs from that estimated through the difference method (MacKinnon and Dwyer, 1993). When the mediator and the outcome are not normal and are modelled with link functions different from identity, for example, using generalised linear models (GLMs), few approaches have been proposed to estimate the indirect effect. MacKinnon and Dwyer (1993) and MacKinnon (2008a) focus on the case of a binary outcome modelled via logistic or a probit model. The indirect effect is obtained as a product of *standardised* regression coefficients. Schluchter (2008) addresses the wider class of GLMs, proposing an extension of the difference method based on generalised estimating equations (GEE). Both these approaches suffer from limitations, the former because it works only for binary outcomes and the latter because it does not allow for exposure-mediator interaction or other forms of nonlinearity in the mediator and the outcome models. It is also worth mentioning the work by Tsai et al. (2006), which extends SEMs to the GLMs framework, but does not discuss how to formalise and estimate indirect effects.

The approach discussed here is based on the simple idea that the indirect effect can be interpreted as the variation in the outcome Y corresponding to a change in the exposure X through the variation in the mediator M (Stolzenberg, 1980). Such a definition can be formalised in terms of derivatives, that is:

$$\frac{\partial Y}{\partial M}\frac{\partial M}{\partial X}, \tag{2.9}$$

i.e. the product of the derivative of Y with respect to M and that of M with respect to X. Let us consider a typical GLM setting with the following models:

$$g_1(\mu_{_M}) = \beta_0 + \beta_1 X \quad \rightarrow \quad \mu_{_M} = h_1(\beta_0 + \beta_1 X) \tag{2.10}$$

$$g_2(\mu_{_Y}) = \gamma_0 + \gamma_1 X + \gamma_2 M \quad \rightarrow \quad \mu_{_Y} = h_2(\gamma_0 + \gamma_1 X + \gamma_2 M), \tag{2.11}$$

where 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 $h_k = g_k^{-1}$, $k \in \{1, 2\}$.

Notice that in the trivial case of identity link functions, the indirect effect in formula (2.9) reduces to the traditional expression obtained via the product method $\gamma_2\beta_1$.

In contrast, when at least one of the g functions differs from identity, the indirect effect assumes a more complex form. In this case, the expression of the indirect effect is not a single value but depends on X and/or M via the derivatives of h_1 and h_2. Assuming a continuous exposure, the researcher chooses some values of X of potential interest, say $x_1, \ldots, x_p$ and, if the expression of the indirect effect also involves the mediator, its values should be selected accordingly to those of X, as the predicted values corresponding to $x_1, \ldots, x_p$, obtained from the fitted model $\mu_{M_{|x_1}}, \ldots, \mu_{M_{|x_p}}$. For this reason, Geldhof et al. (2018) suggest calling the effect in Equation (2.9) *Conditional Indirect Effect* (CIE), since its values are conditional to those of X. To illustrate the concept of conditional indirect effect further, let us consider the case where both the mediator and the outcome are binary variables. Let us also assume that they are both modelled using logistic regression. In this case the indirect effect relative to x_p, using the model specification in Equations (2.10) and (2.11) is found as:

$$CIE_{|x_p} = \frac{\beta_1 \cdot \exp(\beta_0 + \beta_1 \cdot x_p)}{(1 + \exp(\beta_0 + \beta_1 \cdot x_p))^2} \cdot \frac{\gamma_2 \cdot \exp(\gamma_0 + \gamma_1 \cdot x_p + \gamma_2 \cdot \mu_{M_{|x_p}})}{(1 + \exp(\gamma_0 + \gamma_1 \cdot x_p + \gamma_2 \cdot \mu_{M_{|x_p}}))^2}.$$

As highlighted above, in the presence of these nonlinearities, the indirect effect becomes dependent on the values of x_p and $\mu_{M_{|x_p}}$, and this explains why the effect was named 'conditional' by Geldhof et al. (2018).

The main focus here is on mediation analysis with GLMs; nonetheless, it is worth remarking that the derivative-based approach can also be used in situations where the mediator or the outcome depends on nonlinear transformations of their regressors, such as X^2 or $\log(X)$, see Hayes and Preacher (2010) for some examples. Moreover, this approach is crucial even when the distributions of variables M and Y do not belong to the exponential family. In such cases, the key is establishing appropriate link functions that enable us to calculate the derivatives and obtain the CIE.

2.3.1 Potential issues

In this section, we discuss some relevant aspects of the derivative-based method which have not been addressed or satisfactorily deepened in the previous literature on the topic.

Binary, categorical and discrete exposures

The mathematical definition of derivatives is based on the concept of 'small increment' in the argument of the function to differentiate. Thus, the derivative of the mediator with respect to the exposure conceptually relies on a small increment of X, and analogously for the derivative of the outcome with respect to the mediator. This definition makes sense if the support $\mathcal{D}_{W}$ of the variable W with respect to which differentiation is made is continuous. The derivative is a continuous function defined over $\mathcal{D}_{W}$ or a subset. However, when W is discrete, the interpretation of the derivative becomes more challenging, as that of the indirect effect in Equation (2.9).

Let us start from a setting with a binary exposure. The mediator expectation is then a discrete function which can assume only two values. Consequently, the concept of infinitesimal increment is misspecified since the only increment meaningful to conceive is a unit increment. Derivatives cannot be applied to discrete functions; therefore an alternative definition to express change is required. We can use finite differences

$$D_{x,w}[f] = \frac{f(x+w) - f(x)}{w} \tag{2.12}$$

where f is the function of interest and w is the difference between two points in $\mathcal{D}_{X}$. Notice that, when $w \to 0, D_{x,w}(f) \equiv \frac{df}{dx}$. Going back to our mediational setting, and following the notation introduced in Equation (2.10), the derivative of the mediator in the case of binary exposure can then be written simply as the difference

$$D_{0,1}[h_1] = h_1(1) - h_1(0). \tag{2.13}$$

It is easy to prove (see Section 2.3.3) that the chain rule for composite functions holds also in the discrete case, and the indirect effect can then be written as

$$D_{\mu_{M}(x),wD_{x,w}[\mu_{M}(x)]}[\mu_{Y}(\mu_{M})] \cdot D_{x,w}[\mu_{M}(x)],$$

where we explicitly wrote the functional dependence of μ_{M} and μ_{Y}.

The case of binary exposure easily extends to that of categorical exposure. Suppose that X is a categorical variable with k categories. Without loss of generality, assume that the k-th category is the one chosen as baseline. The mediation model

can be rewritten as

$$h_1(\mu_{_M}) = \beta_0 + \sum_{j=1}^{k-1} \beta_j X_j, \tag{2.14}$$

where the X_j are binary variables assuming value 1 if X is in category j, 0 otherwise. In this case, it is necessary to specify the variable with respect to which one takes the difference or, in other words, the category with respect to which the indirect effect is estimated. An increment from 0 to 1 represents the passage from the baseline to this selected category.
The same line of reasoning holds for discrete exposures, for example, number of cigarettes smoked in a day or number of panic attacks in a month.

Binary mediator

As already mentioned, when the expression of the indirect effect involves both X and M, the values of M cannot be chosen arbitrarily, instead they should be fixed at the values the mediator takes in correspondence of the selected values of X, determined by the fitted model. This is generally straightforward unless the mediator is binary when some issues arise.

Consider a setting with a binary mediator, where g_1 in Equation (2.10) is the logit link and, and g_2 in the outcome model in Equation (2.11) is a generic function different from identity, say the logarithm to fix ideas. Therefore, applying the formula in Equation (2.9), the CIE is given by

$$\beta_1\gamma_2 \frac{\exp(\beta_0 + \beta_1 X)}{(1 + \exp(\beta_0 + \beta_1 X))^2} \exp(\gamma_0 + \gamma_1 X + \gamma_2 M),$$

which, as can be seen, depends on both X and M. However, when coming to the estimation of such an effect, the choice of the mediator value is not so immediate. Indeed, the mediator is binary, assuming only two values, while, for any selected value of X, its predicted values from model (2.10) with a logit link are probabilities, ranging in the continuum from 0 to 1. Which values to select, then? This issue is addressed neither by Hayes and Preacher (2010) nor by Geldhof et al. (2018). We believe that the most appropriate solution consistent with a data-generating mechanism of this type is to include binary values of the mediator obtained by the corresponding expected probabilities $\hat{\pi}_{M|X}$ by means of a cutoff c such that

$$\hat{M}|X = \begin{cases} 1 & \text{if } \hat{\pi}_{M|X} \geq c \\ 0 & \text{if } \hat{\pi}_{M|X} < c \end{cases}$$

For example, a possible criterion for the choice of c could be selecting the value for which the sensitivity and specificity of the classification are equal. The most trivial choice $c = 0.5$ may often not be appropriate, for example when classes are unbalanced. A possible alternative criterion for the choice of c could be selecting the value for which the sensitivity and specificity of the classification are equal based on the ROC curve, which can ensure a better performance. Clearly, other approaches are possible, for example maximizing the Youden's index or the F1-score (Berrar, 2019).

Inclusion of covariates

The models in Equations (2.10)-(2.11) are intentionally very simple, but real-word data generally require adjustment for covariates. Including covariates $\boldsymbol{Z}$ in the mediator and the outcome models affects the expression of the indirect effect, which may depend on the covariates' values in addition to those of X and M. For example, consider models as in Equations (2.10)-(2.11), where h_1 is the identity (i.e. the mediator model is linear), and h_2 is the exponential function, and include two (possibly overlapping) sets of covariates $\boldsymbol{Z}_M$ and $\boldsymbol{Z}_Y$ for the mediator and the outcome, respectively:

$$\mu_{_M} = \beta_0 + \beta_1 X + \sum_{k=1}^{p} \beta_{k+1} Z_{Mk}$$

$$\mu_{_Y} = \exp\left(\gamma_0 + \gamma_1 X + \gamma_2 M + \sum_{k=1}^{q} \gamma_{k+2} Z_{Yk}\right).$$

The indirect effect is

$$\beta_1 \gamma_2 \exp\left(\gamma_0 + \gamma_1 X + \gamma_2 M + \sum_{k=1}^{q} \gamma_{k+2} Z_{Yk}\right)$$

i.e., substituting $\mu_{_M}$ to M,

$$\beta_1 \gamma_2 \exp\left(\gamma_0 + \gamma_1 X + \gamma_2 (\beta_0 + \beta_1 X + \sum_{k=1}^{p} \beta_{k+1} Z_{Mk}) + \sum_{k=1}^{q} \gamma_{k+2} Z_{Yk}\right)$$

As already mentioned, the formula also includes covariates. Hayes and Preacher (2010) suggest estimating the indirect effects conditional on the values of X and M, setting the covariates to their mean values. The authors do not address the scenario where the covariates act as effect modifiers, i.e. when they interact with the exposure or the mediator. This simply makes the partial derivatives in Equation (2.9) more

complex but does not add any conceptual difficulty. An important interaction term, which is often included in the outcome model, is that between the exposure and the mediator. This is the case when the exposure moderates its own indirect effect on the outcome through the mediator by moderating the effect of M on Y. To see how the presence of such a term influences the indirect effect, it is sufficient to consider linear models for both the mediator and the outcome and include a term $\gamma_3 XM$ in the outcome model. The indirect effect is $\beta_1(\gamma_2+\gamma_3 X)$, which depends on X, in contrast to the indirect effect obtained from models excluding the presence of an exposure-mediator interaction, i.e. the simple product $\beta_1\gamma_2$. The expression $\beta_1(\gamma_2 + \gamma_3 X)$ is consistent with that obtained by VanderWeele (2015b) in a counterfactual-based framework.

Extension to multilevel models

Geldhof et al. (2018) claim that the derivative-based approach can be easily extended to the multilevel case, but we are not aware of any study addressing this issue. In the following, we discuss how the derivative-based method can be applied to generalized mixed-effect models (GLMMs).

Consider a setting with J clusters and $I = \sum_j n_j$ subjects, where n_j is the number of individuals belonging to each cluster. Typical examples of clustered data are children within classrooms, employees in an organization's departments, or patients in hospitals. Let us start from linear multilevel models where all variables are measured at the subject level (level 2), i.e. a $1 \to 1 \to 1$ design, using the notation introduced by Krull and MacKinnon (1999, 2001)[1]:

$$\mu_{M_{ij}} = (\beta_0 + b_{0j}) + (\beta_1 + b_{1j})X_{ij} \tag{2.15}$$

$$\mu_{Y_{ij}} = (\gamma_0 + g_{0j}) + (\gamma_1 + g_{1j})X_{ij} + (\gamma_2 + g_{2j})M_{ij} \tag{2.16}$$

where j denotes the cluster and i the subject, Greek letters denote fixed effects and $(b_{0j}, b_{1j}, g_{0j}, g_{1j}, g_{2j})'$ is the vector of random effects, which are assumed to be from a multivariate normal distribution with mean $\mathbf{0}$ and covariance matrix $\mathbf{\Sigma}$ possibly non-diagonal (i.e. random effects are allowed to have non-null covariances). Random effects capture the interdependence of units belonging to the same cluster and represent cluster-specific deviances from the average intercept or slope levels.

[1]Single observation and cluster level measurements are indicated by 1 and 2, respectively, which substitutes X, M and Y in the $X \to M \to Y$ setting. E.g., $1 \to 2 \to 2$ indicates that the exposure is measured at the single observation level, and the mediator and the outcome are both measured at the cluster level.

Using Equation (2.9), the indirect effect is the product:

$$(\beta_1 + b_{1j})(\gamma_2 + g_{2j}), \tag{2.17}$$

which depends on b_{1j} and g_{2j}, for $j = 1, \ldots, J$, or, in other words, the indirect effect is cluster-specific. To obtain a unique, average indirect effect, it is necessary to integrate random effects out. If b_1 and g_2 are uncorrelated, then the indirect effect is simply the product $\beta_1\gamma_2$; when, however, they are correlated, then the indirect effect is given by

$$\beta_1\gamma_2 + \sigma_{b_1 g_2}, \tag{2.18}$$

where $\sigma_{b_1 g_2}$ is the covariance between b_1 and g_2 (Kenny et al., 2003). The estimation of such a covariance term in the traditional multilevel setting is complex and requires *ad hoc* solutions, like those proposed in Kenny et al. (2003) and Bauer et al. (2006), or to address multilevel models from a structural perspective (Bauer et al., 2006; Curran, 2003; Preacher et al., 2011, 2010). Another option could be moving to a Bayesian framework (Di Maria et al., 2022; Yuan and MacKinnon, 2009a), which allows us to obtain the posterior distribution of $\sigma_{b_1 g_2}$ and of the indirect effect, also making the estimation of confidence intervals straightforward.

When at least one between h_1 and h_2 differs from identity, or, in other words, when the mediator and/or the outcome model is a generalized mixed model, estimation becomes more complex. For example, if the mediator is a count variable and we model it using a log link,

$$\mu_{M_{ij}} = \exp\{(\beta_0 + b_{0j}) + (\beta_1 + b_{1j})X_{ij}\}$$

while the outcome follows a linear model as in Equation 2.16, the indirect effect is

$$CIE = (\beta_1 + b_{1j})(\gamma_2 + g_{2j}) \exp\{(\beta_0 + b_{0j}) + (\beta_1 + b_{1j})X_{ij}\},$$

for which integrating out random effects and obtaining a closed form expression may be complex or even not feasible. In this case, numerical integration methods may be necessary.

So far, we have focused on the $1 \to 1 \to 1$ design, the one generally most complex due to the potential presence of a product between two random effects. Analogous considerations can, however, be extended to other multilevel mediational designs, like $2 \to 1 \to 1$ or $2 \to 2 \to 1$ [2]. When both the mediator and the outcome

[2]It is important to remark that, according to Krull and MacKinnon (1999, 2001), these are the only designs that can be addressed in the traditional multilevel framework. Other designs,

models are linear, the integration of random effects is straightforward, while when one or both models are nonlinear, some difficulties may arise, especially in the case of correlated random effects.

Confidence intervals for the indirect effect

Estimating confidence intervals (CIs) for the indirect effect poses challenges even in the linear case, since the distribution of the product $\beta_1\gamma_2$ does not follow a Normal distribution although the two coefficient estimators are assumed to be Normal (Lomnicki, 1967; Springer and Thompson, 1966). Indeed, the distribution of the product may be asymmetric and difficult to approximate with distributions traditionally used in statistics, see MacKinnon (2008a). This issue is potentially exacerbated in a nonlinear setting, where the indirect effect assumes complex forms, the distribution of which may be impractical (or simply impossible) to derive in closed form. Therefore, it seems convenient to rely on sampling-based approaches to retrieve an empirical distribution of the indirect effect, from which to compute statistics of interest. Geldhof et al. (2018) suggest using non-parametric bootstrap or Monte-Carlo confidence intervals (for a reference see, e.g., Efron and Tibshirani, 1994; Rubinstein and Kroese, 2016). The former creates B samples by resampling statistical units in the original sample, and for each of them, the parameter of interest is estimated, the indirect effect in our case.

The latter method does not require data resampling, but it generates samples of regression parameters in (2.10)-(2.11), assuming that they come from a multivariate normal distribution. Each of these samples is associated to an estimate of the indirect effect or, more generally, to the parameter of interest.

Supported by a growing body of literature which highlights its desirable properties (see, for example Biesanz et al., 2010; Koopman et al., 2015; Miočević et al., 2017; Yuan and MacKinnon, 2009a), we believe that another valuable option for the estimation of CIs could be the Bayesian approach (for an introduction see, e.g., Gelman et al., 2013). Each parameter is endowed with an *a priori* distribution, and an empirical (posterior) distribution of the indirect effect can be obtained via one of several methods available for this purpose, e.g. Monte Carlo Markov Chains (MCMC). Moving to the Bayesian framework can also provide additional advantages, like the possibility to embed prior information into the mediation model, if available, in order to improve estimates efficiency, ease of extension to the multilevel case, even assuming complex *a priori* correlation structures of fixed and random effects, and exact inference for small samples, for which asymptotic assumptions

including $1 \rightarrow 2$ components, i.e. *bottom-up* effects, cannot be dealt with

might not hold.

To the best of our knowledge, no simulation studies have been run so far to compare the performance of these three approaches for indirect effects estimated through the derivative-based method. This is the primary focus of Section 2.3.2.

2.3.2 Simulation study

In order to overcome the issues related to the closed-form estimation of confidence intervals for indirect effects in a non-linear context, sampling-based approaches may be a possible alternative. In particular, we focused on Bootstrap, Monte Carlo and Bayesian intervals, conducting a simulation study in order to compare their behaviour under different conditions. Namely, we considered three sample sizes ($n = 30, 100, 200$) and two combinations of assumed distribution and link function for both the mediator and the outcome models, that is Bernoulli distribution and logit link in one case and Poisson distribution and log link in the other, with expectations as in Equations (2.10)-(2.11). We generated the exposure variable X from a $\mathcal{N}(0, 5^2)$, and we arbitrarily chose three different exposure values ($x_1 = -5$, $x_2 = 0$ and $x_3 = +5$) on which to condition the indirect effect. The parameter values were arbitrarily set as follows: $\beta_0 = 0.8$, $\beta_1 = 0.2$, $\gamma_0 = 0.3$, $\gamma_1 = 0.40$, $\gamma_2 = -0.7$. Values of the indirect effect have been computed according to the derivative-based method as expressed in Equation (2.9).

For computing Bootstrap and Montecarlo intervals, we estimated GLM expressed in Equations (2.10) and (2.11) using frequentist approach. Bootstrap estimates are obtained by drawing 1,000 samples of size n, with replacement, from the original simulated dataset. Then, with each bootstrap sample, we fit models in Equations (2.10) and (2.11) and at each iteration we saved coefficients' estimates. To retrieve Monte Carlo samples, we sampled 1,000 regression coefficients values from a $\mathcal{N}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\theta}}})$, where

$$\hat{\boldsymbol{\theta}} = \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\boldsymbol{\gamma}} \end{bmatrix} \qquad \hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\theta}}} = \begin{bmatrix} \hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\beta}}} & \mathbf{0} \\ \mathbf{0} & \hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\gamma}}} \end{bmatrix} \tag{2.19}$$

$\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\gamma}}$ being the vectors of estimated coefficients of mediator and outcome models, respectively, and $\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\beta}}}$ and $\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\gamma}}}$ their asymptotic estimated covariance matrices. Matrix $\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\theta}}}$ is block diagonal because we assume that $Cov(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\gamma}}) = 0$. Bayesian posterior coefficients samples have been derived using diffuse priors ($\mathcal{N}(0, 10^3)$) for each parameter, by means of Monte Carlo Markov Chains, from two chains of length 10 000 with burn-in = 5 000. Graphical inspection of the chains showed that all the

chains converged. Simulations were carried out in `R`, using the package `rjags` for the bayesian part.

We repeated the process 500 times, each time computing quantile-based 95% intervals of the indirect effect from its empirical distribution for each approach and scenario. We compared the three approaches in terms of the average length of the intervals and the proportion of intervals which contain the "true" value of the indirect effect (i.e. coverage rate). Results are shown in Figures 2.4 and 2.5 and in Table 2.4.

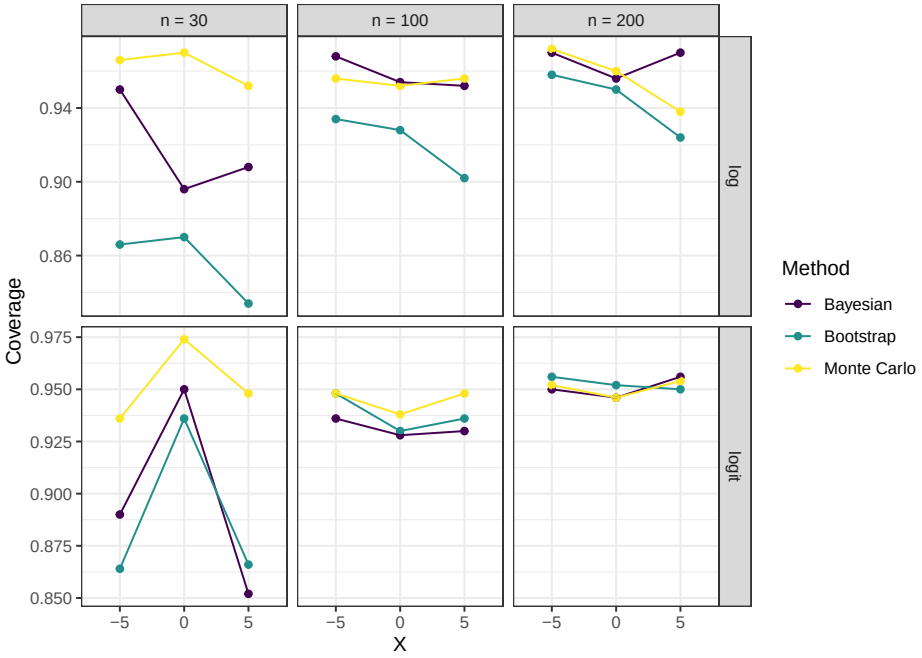


Figure 2.4: Results of the simulation study: coverage rates

As expected, the average lengths of the confidence intervals and the differences between the three methods decrease as the sample size increases. When the sample size is $n = 30$, it is observed that the coverage rate of Monte Carlo confidence intervals is slightly higher compared to the other two methods, while the Bayesian approach fall in between. Analysis of the average length of the credibility intervals in Figure 2.5 provides insights into the uncertainty of effect estimates obtained through the three methods; in scenarios with small samples, the Bayesian approach generally performs better than its counterparts. Generally, Bayesian intervals appears to perform well compared to the others and can be considered a reasonably good alternative.

Table 2.4: Results of the simulation study. For each method, the coverage rates and interval lengths are reported.

Sample size	**Link**	**X value**	**True eff.**	**Bayesian**		**Bootstrap**		**Monte Carlo**	
				Cov. rate	Avg. length	Cov. rate	Avg. length	Cov. rate	Avg. length
30	logit	-5	[-0.005]	0.89	0.045	0.864	0.052	0.936	0.044
		0	[-0.007]	0.95	0.058	0.936	0.121	0.974	0.05
		5	[-0.002]	0.852	0.018	0.866	0.027	0.948	0.017
	log	-5	[-0.01]	0.95	0.037	0.866	0.036	0.966	0.054
		0	[-0.066]	0.896	0.13	0.87	0.15	0.97	0.167
		5	[-0.097]	0.908	0.307	0.834	0.313	0.952	0.43
100	logit	-5	[-0.005]	0.936	0.021	0.948	0.021	0.948	0.02
		0	[-0.007]	0.928	0.025	0.93	0.027	0.938	0.024
		5	[-0.002]	0.93	0.008	0.936	0.009	0.948	0.009
	log	-5	[-0.012]	0.968	0.021	0.934	0.019	0.956	0.021
		0	[-0.097]	0.954	0.075	0.928	0.076	0.952	0.077
		5	[-0.133]	0.952	0.175	0.902	0.167	0.956	0.181
200	logit	-5	[-0.005]	0.95	0.014	0.956	0.014	0.952	0.014
		0	[-0.007]	0.946	0.017	0.952	0.018	0.946	0.017
		5	[-0.002]	0.956	0.006	0.95	0.006	0.954	0.006
	log	-5	[-0.012]	0.97	0.014	0.958	0.013	0.972	0.014
		0	[-0.095]	0.956	0.054	0.95	0.051	0.96	0.052
		5	[-0.124]	0.97	0.133	0.924	0.118	0.938	0.124

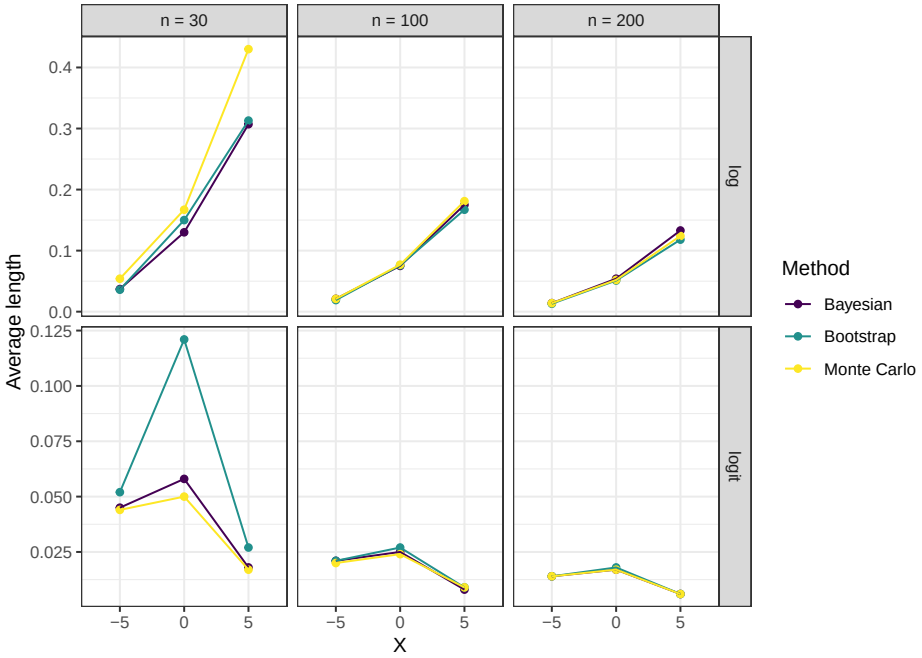


Figure 2.5: Results of the simulation study: average CI lengths

2.3.3 A real-data example: evaluating high school background's impact on academic success

In this section, we analyse data from the ANS (Anagrafe Nazionale Studenti), which serves as the database for Italian university students[3]. Each record in the database represents a statistical unit, specifically a freshman enrolled at an Italian university. These records contain variables about the student's high school background and university career. For this study, we have chosen to focus on the 2015 cohort, the most recent available cohort, which covers a sufficiently long time span to observe the completion of the degree. We have decided to limit the analysis to students enrolled at a non-online Sicilian university, comprising $N = 19\,770$ individuals. It is worth emphasising that the exclusion of students enrolled at online universities is driven by their unique behavioural patterns concerning degree completion (Priulla, 2023).

Our analysis employs associational nonlinear mediation analysis to examine the relationship between high school background and academic success. Specifically, the goal is to examine the impact of the high school final mark (HSFM) on the probability of achieving a bachelor's degree (BD) within four academic years while also exploring the mediating role of the number of University Credits (UC) earned

[3]Database MOBYSU.IT [Mobilità degli Studi Universitari in Italia], research protocol MUR - Universities of Cagliari, Palermo, Siena, Torino, Sassari, Firenze, Cattolica and Napoli Federico II, Scientific Coordinator Massimo Attanasio (UNIPA), Data Source ANS-MUR/CINECA

at the end of the first year.
In addition to the HSFM, CU and BD variables, which act as exposure, mediator and outcome, respectively, we included a set of covariates to control for possible confounders. The three variables of interest and the selected covariates are briefly described below:

- `HSFM`: final mark obtained by the student at the end of high school. In Italy, the final mark ranges from 60 ('Sufficient') to 100 *cum laude*, coded as 101. Decimal scores are not allowed.

- `UC`: number of university credits obtained by the student within the first year from his enrollment to their current degree course. Generally, the maximum number of credits a student can get during the first year is 60.

- `BD`: binary variable, taking value 1 if the student graduates within four years from their enrolment to the current degree course, 0 otherwise.

- `TUSS`: Type of Upper Secondary School diploma. In Italy, there are various types of upper secondary schools, each offering a different curriculum and training students for a particular career or academic path. In this study they have been categorized in Classical *lyceum*, Scientific *lyceum*, Technical institute, Industrial Technical institute, Vocational institute, Industrial Technical institute, Other *lyceum* (baseline), and Abroad/Other.

- `sex`: student's biological sex, male and female (baseline).

- `TDC`: area of the degree course at which the student is enrolled, categorised in "Agriculture, forestry, fisheries and veterinary" (baseline), "Arts and humanities", "Engineering, manufacturing and construction", "Health and welfare", "Business, administration and law", "Natural sciences, mathematics and statistics (NsMS)", "Services", "Social sciences, journalism and information", "Education", and "Information and Communication Technologies (ICTs)".

- `age`: student's age.

In this case study, the variables UC and BD acting as the mediator and outcome, respectively, do not follow a Gaussian distribution, suggesting the need for nonlinear mediation analysis. Specifically, the UC variable is bounded between 0 and 60, as mentioned before. To account for this, we first transformed UC into the proportion of UC (PUC) obtained in the first year, dividing UC by 60; then, to make these scores

strictly in the interval (0,1), we applied the transformation proposed by Smithson and Verkuilen (2006):

$$PUC' = \frac{PUC \cdot (N-1) + 0.5}{N},$$

and employed a Beta model with logit link for analysis:

$$\text{logit}\left(\mathbb{E}[PUC'|X,Z]\right) = \beta_0 + \beta_1 \text{HSFM} + \boldsymbol{\beta_2} \text{TUSS} + \beta_3 \text{sex} + \boldsymbol{\beta_4} \text{TDC} + \beta_5 \text{age}.$$

Regarding the outcome, BD is a binary variable that takes the value of 1 if the student successfully graduates on time. To examine the relationship between HSFM and PUC′, we employed a logistic regression model as follows:

$$\text{logit}\left(P[BD=1|M,X,Z]\right) = \gamma_0 + \gamma_1 \text{HSFM} + \gamma_2 \text{PUC}' + \boldsymbol{\gamma_3} \text{TUSS} + \gamma_4 \text{sex} + \boldsymbol{\gamma_5} \text{TDC} + \gamma_6 \text{age}.$$

An important point to highlight is that the exposure variable, HSFM, is a discrete variable. Consequently, estimating the indirect effect requires the use of finite differences methodology. Since the model also includes some other covariates, they need to be fixed to specific values. Specifically, we fixed TUSS, TDC, and age to their joint mode (i.e. the most recurrent profile of covariates in the data): TUSS = Scientific *lyceum*, TDC = Engineering, manufacturing and construction, and age = 19. In contrast, the variable sex has not been explicitly assigned. Indeed, we calculated the CIEs for both males and females, enabling a meaningful comparison. The indirect effect for the i-th value of HSFM is found as:

$$\text{CIE}_i = \text{logit}^{-1}(\eta_{i+1}) - \text{logit}^{-1}(\eta_i), \quad i = 1, \ldots, 41, \tag{2.20}$$

where

$$\eta_i = \gamma_0 + \gamma_1 \text{HSFM}_i + \gamma_2 \text{PUC}'_{|HSFM_i} + \gamma_3^{(TUSS=Sci)} + \gamma_4 \text{sex} + \gamma_5^{(TDC=Eng)} + \gamma_6 \cdot 19$$

and

$$PUC'_{|HSFM_i} = \text{logit}^{-1}(\beta_0 + \beta_1 \text{HSFM}_i + \beta_2{}^{(TUSS=Sci)} + \beta_3 \text{sex} + \beta_4{}^{(TDC=Eng)} + \beta_5 \cdot 19).$$

Note that $\text{HSFM}_1 = 60$ and $\text{HSFM}_{42} = 101$. In this terms, CIE_i quantifies how the probability of achieving BD changes when HSFM increases by one unit, from HSFM_i to HSFM_{i+1}, considering the mediating effect of PUC′.
The coefficients involved in the estimation of the CIEs, estimated using frequentist

Table 2.5: Estimates and pvalues of regression coefficients involved in the estimation of CIEs

	Mediator			***Outcome***		
Name	Coef	Estimate	pvalue	Coef	Estimate	pvalue
Intercept	β_0	-2.683	$<$ 2e-16	γ_0	-2.900	$<$ 2e-16
HSFM	β_1	0.039	$<$ 2e-16	γ_1	0.0004	0.774
PUC′	-	-	-	γ_2	4.583	$<$ 2e-16
TUSS = Sci	β_2	0.361	$<$ 2e-16	γ_3	-0.050	0.345
sex = male	β_3	-0.032	0.124	γ_4	-0.039	0.321
TDC = Engineering	β_4	-0.114	0.015	γ_5	-0.302	0.001
age	β_5	-0.034	$<$ 2e-16	γ_6	-0.001	0.903

approach, are reported in Table 2.5, while the whole set of coefficients is provided for reference in Table 2.6 in Section 2.3.3.

The results in Table 2.5 seem to suggest that the relationship between HSFM and BD is fully mediated by PUC′ since the effect of HSFM in the outcome model, γ_1, is not significant. HSFM is positively and significantly associated with PUC′, which is in turn positively and significantly associated with BD. The magnitude of the latter coefficient is remarkable (4.583). To formally test if the indirect effect is significant, we estimated the CIEs (as shown in Equation 2.20) and their confidence intervals using the three approaches discussed before. As in Section 2.3.2, Bootstrap and Monte Carlo intervals are based on frequentist models, while the Bayesian intervals are computed using MCMC. Figure 2.6 shows the results obtained using the Bayes approach, while those obtained with the other approaches are almost identical; actually, this is in agreement with what we observed in simulations in the scenario with a large sample size, as it is here.

Each point in the graph represents the difference in the probability of graduating within four years for a unitary increase in HSFM mediated by PUC′. It is worth noting that all the estimated CIEs are positive and significant (their CIs does not contain zero), meaning that getting a higher HSFM is associated with a higher probability of graduating on time through PUC′. However, the curve has a monotonic increasing trend until HSFM reaches 92, then it slightly starts to decrease. This may suggest that the mediating role of PUC′ becomes more and more important as HSFM increases until 92, at which point it becomes slightly less relevant. In addition, we can notice that the indirect effect for females is slightly larger than those of males; however, the confidence intervals overlap for all values of HSFM, implying that the observed differences in CIEs magnitude are not significant. This

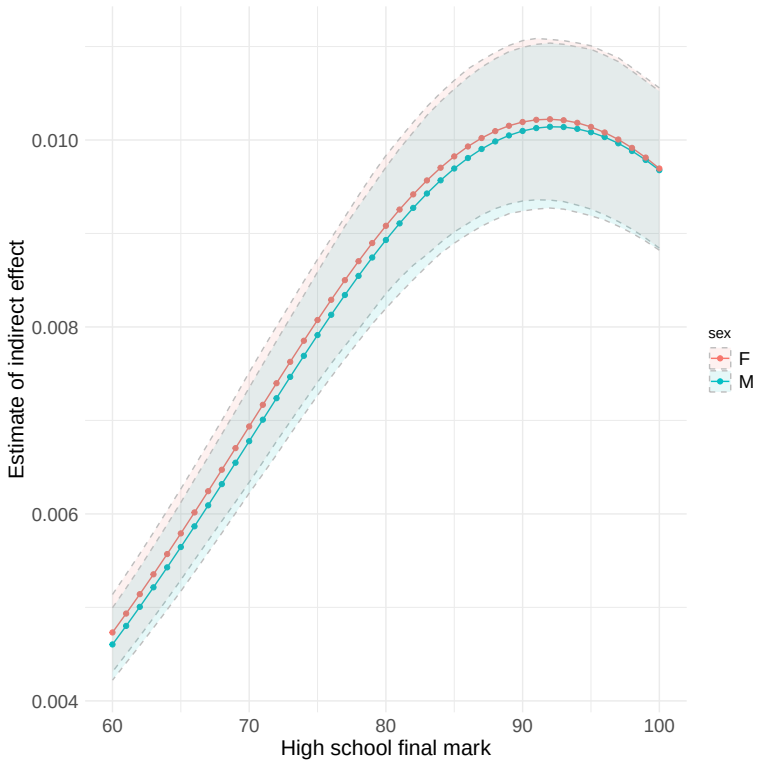


Figure 2.6: CIEs for male and females and their confidence intervals estimated with the Bayesian approach

is consistent with the regression model's results.

Supporting material

Chain rule for finite differences

Let $f(x)$ and $g(x)$ two discrete functions and $f \circ g \equiv f(g(x))$ the function obtained from their composition. We want to prove that

$$D_{x,w}[f(g)] = \frac{f(g(x+w)) - f(g(x))}{w}$$

can be written as a product of differences, deriving a chain rule analogous to that for derivatives of continuous functions. Indeed, $D_{x,w}[f(g)]$ can be written as

$$\begin{aligned} \frac{f(g(x+w)) - f(g(x))}{w} &= \frac{f\left(g(x) + w\,\frac{g(x+w)-g(x)}{w}\right) - f(g(x))}{w} \\ &= \frac{f\left(g(x) + wD_{x,w}[g]\right) - f(g(x))}{w} \end{aligned}$$

Noting that

$$D_{g(x),wD_{x,w}[g]}[f(g)] = \frac{f(g(x) + wD_{x,w}[g]) - f(g(x))}{wD_{x,w}[g]},$$

it is easy to derive that

$$D_{x,w}[f(g)] = D_{g(x),wD_{x,w}[g]}[f(g)] \cdot D_{x,w}[g].$$

Regression coefficients

Table 2.6: Estimates and pvalues of regression coefficients.

		Mediator			***Outcome***	
Name	Coef	Estimate	pvalue	Coef	Estimate	pvalue
Intercept	β_0	-2.683	< 2e-16	γ_0	-2.900	< 2e-16
HSFM	β_1	0.039	< 2e-16	γ_1	0.0004	0.774
PUC′	-	-	-	γ_2	4.583	< 2e-16
TUSS = Sci	β_2	0.361	< 2e-16	γ_3	-0.050	0.345
TUSS = Clas		0.209	< 2e-16		-0.301	< 2e-16
TUSS = Tech		-0.209	< 2e-16		-0.137	0.039
TUSS = Voc		-0.38	< 2e-16		-0.338	0.002
TUSS = Ind Tech		-0.101	0.047		-0.066	0.525
TUSS = Abroad		0.021	0.894		0.189	0.54
sex = male	β_3	-0.032	0.124	γ_4	-0.039	0.321
TDC = Engineering	β_4	-0.114	0.015	γ_5	-0.302	0.001
TDC = Arts		0.438	< 2e-16		0.092	0.331
TDC = Health		0.326	< 2e-16		-0.555	< 2e-16
TDC = Business		-0.017	0.708		-0.865	< 2e-16
TDC = NsMS		-0.425	< 2e-16		0.453	< 2e-16
TDC = Education		0.734	< 2e-16		-0.921	< 2e-16
TDC = ICTs		-0.299	<0.001		-0.141	0.419
age	β_5	-0.034	< 2e-16	γ_6	-0.001	0.903

Estimates of CIEs on real data

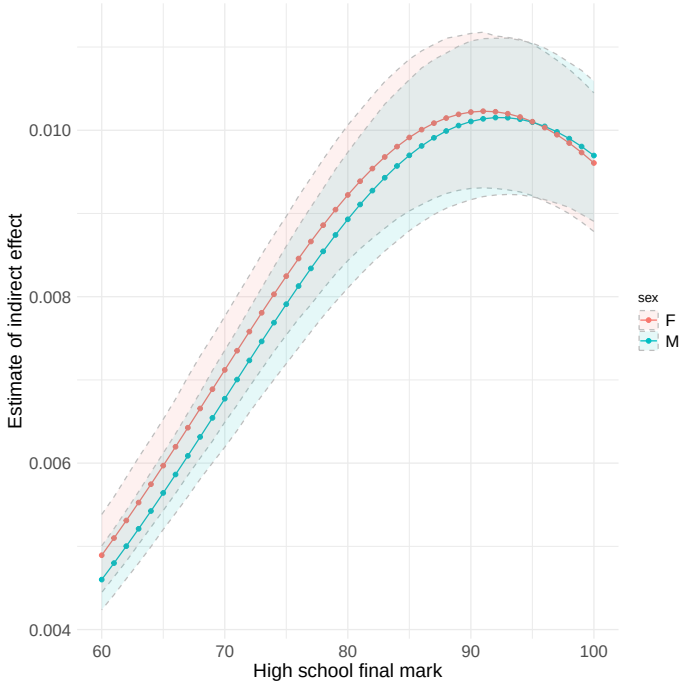


Figure 2.7: CIEs for male and females and their confidence intervals estimated with the Bootstrap approach

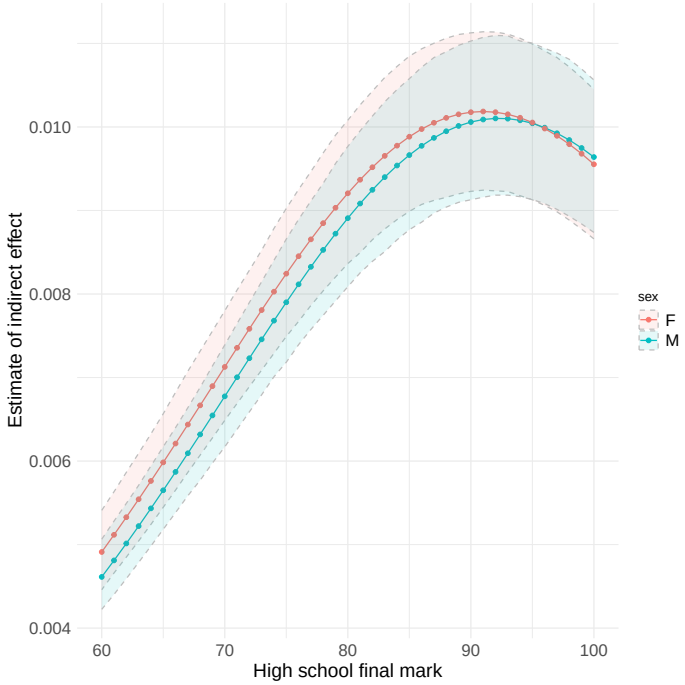


Figure 2.8: CIEs for male and females and their confidence intervals estimated with the Montecarlo approach

2.4 Conclusions

In this Chapter, we focused on the estimation of nonlinear mediation models. The existing literature primarily focuses on nonlinear mediation models within the counterfactual framework, which requires different assumptions and notation compared to the associational framework commonly used in mediation analysis.

Initially, we described a healthcare study conducted into the counterfactual framework. Namely, we investigated the effect of socio-economic status on cancer mortality as mediated by out-of-region mobility, focusing on lung and colon cancer patients in Sicily, Italy. We followed a novel two-step approach: first, we derived the structure of association among the variables from the data, using log-linear graphical models; than, we estimated mediational effects using a counterfactual-based approach.

The findings of the study highlight the existence of a complex structure of relationships among individual characteristics, patient mobility and health outcomes, herein described in a 3-year mortality in cancer patients. In line with much of previously published literature, older patients and those with a poorer profile in terms of comorbidities have demonstrated a lower probability of receiving treatments extra-regionally.

The results of the expounded graphical model have highlighted that most deprived patients have a diminished probability of travelling outwith their region of residence, and this holds across all age categories. Contemporaneously, mobility decreases the 3-year mortality for both cancer sites under consideration in this study.

Given the complex structure of the relationships among deprivation, mobility and the 3-year mortality, the mediation analysis outlined in this study has permitted the quantification of the indirect deprivation effect on mortality, as mediated by mobility. The results obtained have highlighted that, as regards TBL cancer, a state of deprivation acts on mortality indirectly (via considerations of mobility) and directly (via unmeasured factors, which are different in nature to mobility). In contrast, the effects of deprivation are mainly indirect for colon cancer. When considering the different effect of socio-economic status on cancer survival, further research is required to analyze the extent to which deprivation may be confounded by co-morbidity and lifestyle.

Our study acknowledges several limitations and identifies areas for future research. Key limitations include the absence of detailed medical information such as cancer stage, malignancy type, histological features, and tumor genetics in the hospital records, which are critical factors in determining 3-year mortality rates. While patient comorbidity was assessed using the Charlson index, more comprehensive

methods could be employed for a deeper analysis. Due to the study's cross-sectional nature, important dynamic factors over time were not captured, which could be addressed in future longitudinal studies using, for example, those approaches developed in Didelez (2019) and Aalen et al. (2020).

The second part of the Chapter is devoted to the derivative-based approach to the estimation of the indirect effect in an associational mediational context. Stolzenberg (1980) proposed to estimate the indirect effect as the derivative of a composite function, and Geldhof et al. (2018) discussed some applications of this approach in the GLM context. We offered a comprehensive discussion of the derivative-based method for nonlinear mediation analysis by deepening some aspects of the proposal of Geldhof et al. (2018), addressing some of its potential issues through novel solutions.

Specifically, we proposed how to address discrete exposures through a discrete version of the derivative method based on finite differences; we tackled binary mediators, proposing a way to select an appropriate value to include in the outcome model and discuss potential extensions to the multilevel framework and the application of the Bayesian framework for obtaining confidence intervals. Through a simulation study, we provided evidence that Bayesian intervals are a valid alternative compared to the Bootstrap-based and Montecarlo ones used by Geldhof et al. (2018).

Also, we present a real-data application which investigated the relationship between high-school background and academic success. By calculating finite differences, we were able to capture the effects of the mediator on the outcome variable while accounting for the discreteness of the high-school background variable. The results obtained through this approach were then interpreted accordingly, acknowledging the specific characteristics of the variables involved.

We believe that this work can serve as a guide for researchers who need to address a mediational setting with nonlinear models without switching to the counterfactual framework. This work can be extended in several ways, for example, by investigating the relationship between the total and indirect effects. Indeed, in the classical linear setting, the total effect is the sum of the direct and indirect effects; however, this property does not hold in the case of nonlinear models. Another possible venue for future research is clustered data, for which multilevel models are often employed. Settings with categorical mediators also need further investigation. This issue has traditionally received little attention by scholars, either in the counterfactual and the associational framework, and can be a promising research direction.

Chapter 3

Derivative-based spatial mediation with INLA-SPDE

3.1 Introduction

Mediation analysis, as introduced in Chapter 2, is commonly used in various disciplines like biology, social sciences and epidemiology to assess the indirect effect of an exposure on an outcome through a mediator variable. In those fields, encountering data with inherent spatial structure is a common occurrence. 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.

However, although recent explorations into spatial variability in causal inference indicate viable paths (Reich et al., 2021), the literature on mediation analysis predominantly lacks consideration for spatial heterogeneity, particularly in the context of spatial exposures, mediators, or outcomes. While not directly addressing spatial mediation, referring to the broader field of spatial causal inference, techniques such as case-control matching (Jarner et al., 2002), spatial smoothing adjustments (Schnell and Papadogeorgou, 2019), and propensity-score methods (Davis et al., 2019) have been developed to address bias due to missing spatial confounders. Approaches aimed at addressing the problem of interference, which occurs in cases where the treatment in one location affects outcomes in other locations, can be found in Zigler et al. (2012) and (Tchetgen Tchetgen et al., 2021). Also in Reich et al. (2021) are discussed spatial causal methods based on the potential outcomes framework (Rubin, 1974) and on Granger causality (Granger, 1969). All this approaches deal with areal data. Considering point-referenced data, in which the

treatment and response variables can be modeled as continuous random fields over an uncountable number of spatial locations (as described in Section 1.3), Reich et al. (2021) mention the possibility of using the INLA-SPDE approach to remove the effects of spatially-smooth confounding variables, viewing that as an extension of Simultaneous Autoregressive (SAR) models to the continuous spatial domain.

In Chapter 1 we have shown how the SPDE approximation to Gaussian spatial fields can be effectively applied to estimate Bayesian geostatistical models, exploiting all the model flexibility and the computational efficiency offered by INLA. Also, in Chapter 2 we described how the derivative-based approach can be used to estimate conditional indirect effects in nonlinear mediational settings. Following a growing body of literature (Yuan and MacKinnon, 2009b), we also explored the possibility to estimate Bayesian credibility intervals for the CIE. Although it is still a work-in-progress, in this Chapter we wanted to report some evidence on the opportunity to combine the INLA-SPDE approach and the derivative-based method, within the Bayesian inferential paradigm, in order to estimate nonlinear mediational effects while taking into account spatial correlation in the data.

3.2 The spatial conditional indirect effect

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 introduced in Chapter 1 and 2, and assuming the presence of spatial correlation at both the mediator and the outcome level, the mediation model expressed earlier in Equations (2.10) and (2.11) becames

$$g_1(\mu_M(\boldsymbol{s})) = \beta_0 + \beta_1 X(\boldsymbol{s}) + u_M(\boldsymbol{s}) \tag{3.1}$$

$$g_2(\mu_Y(\boldsymbol{s})) = \gamma_0 + \gamma_1 X(\boldsymbol{s}) + \gamma_2 M(\boldsymbol{s}) + u_Y(\boldsymbol{s}) \tag{3.2}$$

where X, M and Y are the exposure, the mediator and the outcome, respectively, observed at location $\boldsymbol{s}$ in a spatial domain $D \in \mathbb{R}^d$, $u_M(\boldsymbol{s})$ and $u_Y(\boldsymbol{s})$ 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 conditional indirect effect according to the derivative-based approach, it is necessary to compute the partial derivatives of the mediator and

the outcome model, as expressed in Equation 2.9. Table 3.1 shows 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 $\boldsymbol{s}$.

Table 3.1: Partial derivatives of spatial mediator and outcome models for identity, log and logit link functions.

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

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

$$\beta_1 \exp(\beta_0 + \beta_1 X + u_M(\boldsymbol{s})) \times \frac{\gamma_2 \exp(\gamma_0 + \gamma_1 X + \gamma_2 M + u_Y(\boldsymbol{s}))}{(1 + \exp(\gamma_0 + \gamma_1 X + \gamma_2 M + u_Y(\boldsymbol{s}))^2}. \tag{3.3}$$

3.3 Spatial mediation with INLA-SPDE: an example

Following the discussion reported in Section 1.4 of Chapter 1, an useful approach for estimating the spatial models expressed in Equations (3.1) and (3.2) is by using the SPDE approximation to the Gaussian field, as demonstrated in Lindgren et al. (2011). To ascertain the convenience of taking into account spatial correlation when estimating indirect effects with point-referenced data, we used simulated data to compare a traditional mediation model with a spatial mediation model. The spatial model includes a Gaussian spatial process at both the mediator and the outcome level.

3.3.1 Simulation of the data

We generated $n = 150$ point locations randomly in a unit square, simulating a higher density of points in the lower left corner, as shown in Figure 3.1.

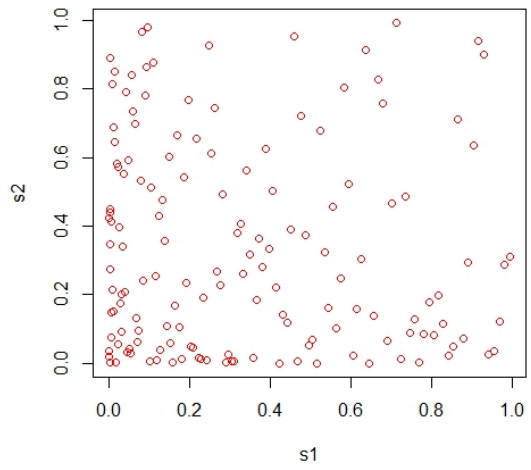


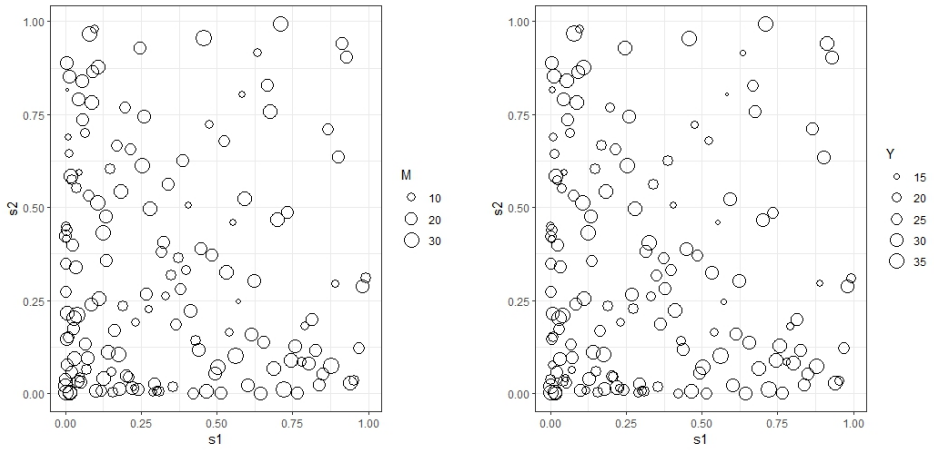
Figure 3.1: Simulated locations in the unit square.

Then, we arbitrarily set parameter values for the two spatial processes, that is $\sigma^2_{u_M} = \sigma^2_{u_Y} = 5$, $\kappa = 7$ and $\nu = 1$. After calculating the distances between each point, the Matérn covariance matrices of the two processes have been calculated as shown in Equation 1.8. Then, values from the spatial fields u_M and u_Y have been generated randomly using Cholesky decomposition.

To focus exclusively on the exclusion/inclusion of the spatial process, we have considered the non-generalised case (using the identity link function) assuming that both the mediator and the outcome have Normal errors ϵ_M and ϵ_Y, respectively, distributed as $\mathcal{N}(0, 2)$. Parameters of the mediator and outcome models have been set, respectively, as $\beta_0 = 4.5, \beta_1 = 1.5$ and $\gamma_0 = 10, \gamma_1 = 0.8, \gamma_2 = 0.4$. Hence, according to the formula reported in Table 3.1, the indirect effect is $\beta_1 \times \gamma_2 = 0.6$. Also, we set exposure $X \sim \mathcal{N}(10, 16)$. Figure 3.2 shows simulated mediator and response values in the unit square (the values in the legend are provided just to indicate the correspondence between the size of the points and the values of the variables).

3.3.2 Models estimation

We estimated posterior densities of parameters with INLA for both the spatial and non-spatial mediation models, using non-informative priors. In particular, for the spatial mediation model, we fitted an INLA-SPDE model as described in Section 1.4. For the hyperparameters of the SPDE component, we assumed a joint Normal prior for hyperparameters λ_1 and λ_2 with expected values $\mu_{\lambda_1} = -3.23$ and $\mu_{\lambda_1} = 1.96$, precision $\tau_{\lambda_1} = \tau_{\lambda_2} = 0.1$ and covariance zero.



(a) (b)

Figure 3.2: Simulated mediator (3.2a) and response (3.2b) values.

Figure 3.3 shows the mesh of the simulated locations for the INLA-SPDE model. The vertices of the triangles represent the nodes (bases) of Equation (1.13), the number of which is 390. To avoid boundary effects, we added an outer edge as suggested by Simpson et al. (2016). Considering the trade-off between accuracy and computational cost, we set a lower resolution for the outer region, which is of no practical interest, and allowed a higher number of nodes in the study region to compute an accurate approximation of the Gaussian field in that area. The maximum triangle side length is set to 0.2 in the interior and 0.3 in the outer region.

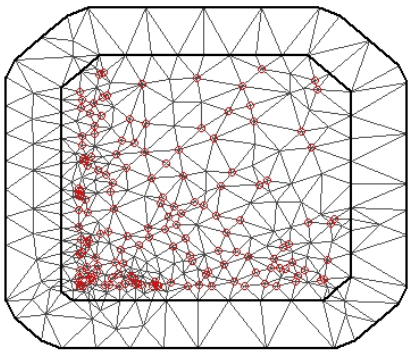
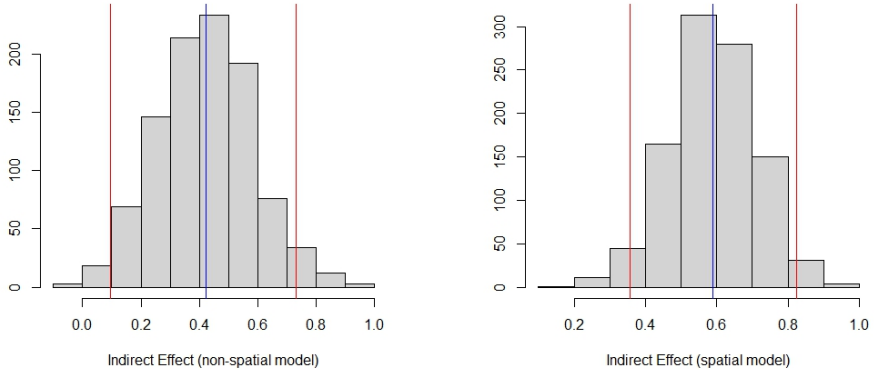


Figure 3.3: Mesh of the simulated locations (red dots) for the INLA-SPDE model.

Table 3.2 shows a comparison of the DIC of the non-spatial vs spatial mediation models fitted to the simulated dataset. It can be seen how accounting for spatial correlation by including the spatial processes at both the mediator and the outcome



(a) (b)

Figure 3.4: Histograms of the posterior distribution of the indirect effect computed using estimates from the non-spatial mediation model (3.4a) and from the spatial mediation model (3.4b). Red lines represent HDI intervals, and blue lines indicate the posterior mean. The true value of the indirect effect is equal to 0.6.

level improves the performance of the models.

Table 3.2: Comparison of the DIC of the non-spatial vs spatial mediation models fitted to the simulated dataset.

	DIC	
	Mediator model	Outcome model
Without spatial GF	664.35	756.07
With spatial GF	595.25	561.34

Recall that the indirect effect is a function of parameters from both the mediator and the outcome models. Therefore, to compare the estimates of the indirect effect obtained from the two approaches under comparison, from the joint posterior distribution we simulated $S = 1000$ values of parameters β_1 and γ_2, that is $\boldsymbol{\beta_1^p} = \{\beta_1^{(1)}, \ldots, \beta_1^{(S)}\}$ and $\boldsymbol{\gamma_2^p} = \{\gamma_2^{(1)}, \ldots, \gamma_2^{(S)}\}$, for both the spatial and non-spatial mediation models. We then retrieved the posterior sample of the indirect effect as $\boldsymbol{\beta_1^p} \odot \boldsymbol{\gamma_2^p}$ (where $\odot$ indicates the Hadamard product), from which we computed statistics of interest, specifically Highest Density Intervals and posterior means. These are illustrated in Figure 3.4 and reported in Table 3.3.

It turns out that the spatial mediation model has provided more accurate and precise estimates of the indirect effect compared to the non-spatial model.

Table 3.3: Posterior mean and HDI of the indirect effect estimated with the non-spatial and with the spatial mediation models.

	Posterior statistics of the IE	
	Mean	HDI
Without spatial GF	0.422	(0.10, 0.73)
With spatial GF	0.590	(0.36, 0.82)

3.4 Simulation study

In order to evaluate the impact on the estimation of the indirect effect of including/excluding the GFs, we conducted a simulation comparing 16 different scenarios. These scenarios varied in sample sizes and spatial correlation structures (i.e., combinations of parameter values r and σ_u^2). For each scenario, we repeated the process described in Section 3.3 300 times. At each iteration, we collected the highest posterior density intervals estimated from either model. We then computed their average length and coverage rate (i.e., the proportion of intervals that cover the 'true' indirect effect of 0.6), which are reported in Table 3.4. Given that the size of the finite-dimensional Gaussian approximation to the solution of the SPDE is solely dependent on the desired resolution and not on the number of observations, we kept the mesh construction parameters constant, as specified in Section 3.3.2.

An increase in sample size from 50 to 200 shows an improvement in the coverage rate for the spatial model compared to the non-spatial model, suggesting that the inclusion of the GF becomes more effective as the sample size increases. The average interval length decreases for both approaches with larger n, but this reduction is more pronounced for the spatial model.

A larger spatial range (increasing from 0.2 to 0.8) appears to enhance the coverage rate for the spatial model, especially in larger samples (n = 200). This indicates greater adaptability of the spatial model in scenarios with higher spatial variability. As σ_u^2 increases from 5 to 15, the coverage rate for the non-spatial model significantly drops, whereas that of the spatial model remains relatively stable. This demonstrates that the spatial model better accommodates the increase in variance compared to the non-spatial model.

In terms of coverage rate, the spatial model generally outperforms the non-spatial model across almost all scenarios, particularly in larger samples and with higher spatial variance (σ_u^2). Regarding average interval length, the GF produces shorter credibility intervals, indicating greater precision in estimating the indirect effect. In summary, the differences between the two models are more pronounced in scenarios

with larger sample sizes, wider ranges, and higher spatial variance. This suggests that ignoring spatial correlation in the data could be particularly detrimental under these conditions.

3.4.1 Correlated processes

Until now, we assumed that $\boldsymbol{u}_M \perp\!\!\!\perp \boldsymbol{u}_Y$. In this Section, we consider the scenario in which the two spatial processes $\boldsymbol{u}_M$ and $\boldsymbol{u}_Y$ are jointly Gaussian distributed, i.e.

$$\boldsymbol{u}_M, \boldsymbol{u}_Y \sim \mathcal{N}(\mathbf{0}_{\boldsymbol{u}_M,\boldsymbol{u}_Y}, \boldsymbol{\Sigma}_{\boldsymbol{u}_M,\boldsymbol{u}_Y}), \tag{3.4}$$

where $\mathbf{0}_{\boldsymbol{u}_M,\boldsymbol{u}_Y}$ 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:

$$\boldsymbol{\Sigma}_{\boldsymbol{u}_M,\boldsymbol{u}_Y} = \begin{bmatrix} \boldsymbol{\Sigma}_{\boldsymbol{u}_M} & \boldsymbol{\Sigma}_{\boldsymbol{u}_M,\boldsymbol{u}_Y} \\ \boldsymbol{\Sigma}^T_{\boldsymbol{u}_M,\boldsymbol{u}_Y} & \boldsymbol{\Sigma}_{\boldsymbol{u}_Y} \end{bmatrix}, \tag{3.5}$$

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

$$\rho \times \sqrt{\text{Cov}(u_M(\boldsymbol{s}_i), u_M(\boldsymbol{s}_j)) \times \text{Cov}(u_Y(\boldsymbol{s}_i), u_Y(\boldsymbol{s}_j))}, \tag{3.6}$$

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

Table 3.5 presents the results of the simulation study comparing the performance of the spatial and non-spatial mediation models, across different combinations of sample size n and the ρ parameter, which governs the correlation between the two spatial fields u_M and u_Y. We evaluated the coverage rate and the average length of credibility intervals for both models.

The table results suggest that, for both models (spatial and non-spatial), there is a decrease in the coverage of credibility intervals and an increase in the average length of intervals as ρ increases. This suggests that higher correlation between spatial models negatively impacts the precision of estimates. This is expected, as both models fail to take into account the correlation between the two spatial processes. However, despite the increase in ρ, the spatial model 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. Furthermore, increasing sample size seems to improve the performance of both models, as evidenced by higher coverage of credibility intervals and lower

average length of intervals for higher values of n.

3.5 An application on real data

3.5.1 Motivation of the study

One of the areas where spatial mediation can prove particularly useful is the analysis of ecological data, which often exhibit 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 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 (Massutí et al., 2000). 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).

3.5.2 The data

The data are sourced from the multivariate dataset of marine species described in Section 1.5.2 of Chapter 1. Specifically, we analysed $n = 56$ 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 18m (minimum observed depth in the dataset) and 75m (the thermocline, beyond which temperature no longer decreases with increasing depth), considering only the most recent available period from 2015 to 2018. Figure 3.5 shows the location of Pandora catches in the study area.

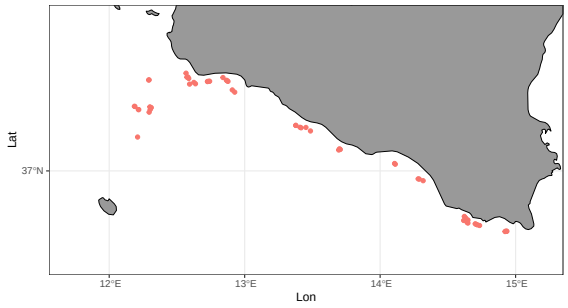


Figure 3.5: Location of Pandora fish catches in the study area.

3.5.3 An INLA-SPDE spatial mediation model for Pandora

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

$$\mu_M(\boldsymbol{s}) = \beta_0 + \beta_1 X(\boldsymbol{s}) + u_M(\boldsymbol{s}) \tag{3.7}$$

$$log(\mu_Y(\boldsymbol{s})) = \gamma_0 + \gamma_1 X(\boldsymbol{s}) + \gamma_2 M(\boldsymbol{s}) + u_Y(\boldsymbol{s}), \tag{3.8}$$

where X is the depth, M is the temperature, Y is the biomass of Pandora, $\boldsymbol{s}$ is the vector of spatial coordinates (latitude and longitude) of the catch, and $u_M(\boldsymbol{s})$ and $u_Y(\boldsymbol{s})$ 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 $M \sim \mathcal{N}(\mu_M, \sigma_M^2)$; also, considering the non-normal distribution of biomass, we assumed that $Y \sim$ Gamma with expected value μ_Y and variance σ_Y^2 and a log link function, which requires nonlinear mediation analysis.

The models expressed in Equations (3.7) and (3.8) were estimated using the INLA-SPDE approach. To choose the priors for the hyperparameters of the SPDE component, we followed the same criteria as described in Section 1.5.3, assuming a joint Normal prior for hyperparameters λ_1 and λ_2 with expected values $\mu_{\lambda_1} = -2.64$ and $\mu_{\lambda_1} = 1.38$, precision $\tau_{\lambda_1} = \tau_{\lambda_2} = 0.1$ and covariance zero.

Figure 3.6 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

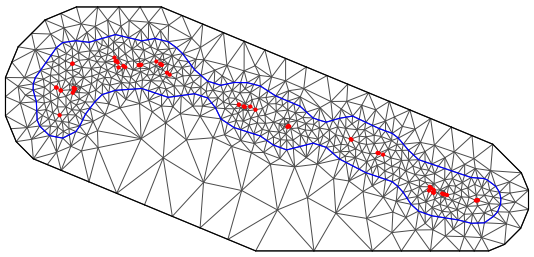


Figure 3.6: Triangulation of the study region and catch locations (red dots)

approach, as follows:

$$CIE^p(\boldsymbol{s}') = \boldsymbol{\beta}_1^p \odot \boldsymbol{\gamma}_2^p \odot exp(\boldsymbol{\gamma}_0^p + \boldsymbol{\gamma}_1^p x(\boldsymbol{s}') + \boldsymbol{\gamma}_2^p \odot \boldsymbol{m}^p(\boldsymbol{s}') + \sum_{j=1}^{m} A^*_{j\boldsymbol{s}'} \boldsymbol{w}_j^p), \qquad (3.9)$$

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 $A^*_{j\boldsymbol{s}'}$ 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 (as described in Section 1.4 of Chapter 1). The indirect effect was estimated both with and without the spatial component at the mediator and response levels.

3.5.4 Results

Table 3.6 reports posterior statistics for parameters and hyperparameters of the INLA-SPDE models expressed in Equations (3.7) and (3.8).

The inclusion of the spatial components lead 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 -0.41, respectively). The significance of γ_1 suggests that depth could affect Pandora abundance not only indirectly through a change in temperature, but also directly.

Figure 3.7 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 $P = \{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_L\}$ in the study region.

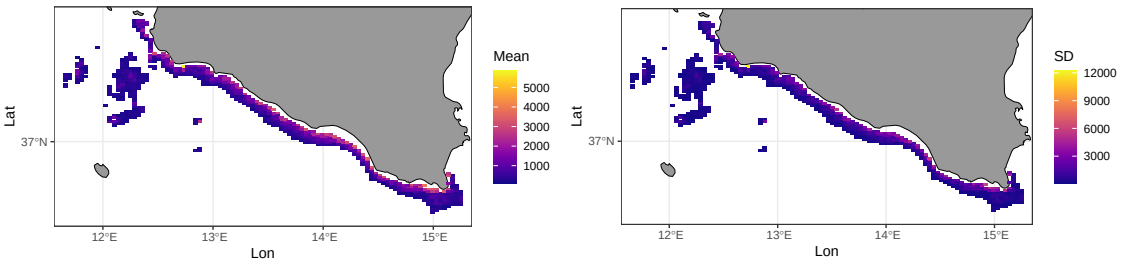


Figure 3.7: Mean and SD of the posterior distribution of the indirect effect in the study area.

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 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 Figure 3.8.

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 3.6.

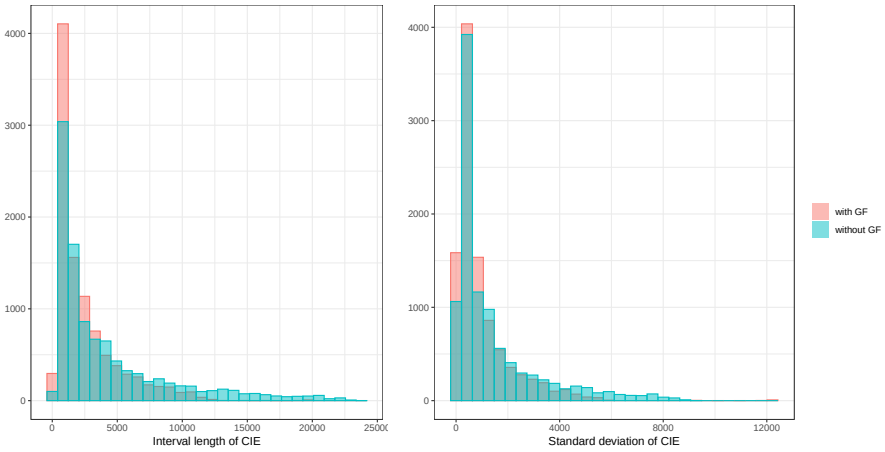


Figure 3.8: Comparison of the SDs and HDI interval lengths of the posterior distribution of the CIE, for the spatial and the non-spatial model.

3.6 Conclusions

In this chapter, we explored 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 the 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, high spatial variability, and a wide spatial range.

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.

Table 3.4: Comparison of performances in estimating the indirect effect including/excluding the GF in different scenarios.

n	r	σ_u^2	coverage (GF)	coverage (no GF)	avg. len. (GF)	avg. len. (no GF)	coverage diff.	avg. len. diff.
50	0.20	5	0.9167	0.87	0.4964	0.5198	0.0467	-0.0233
50	0.20	5	0.9167	0.9067	0.4987	0.5269	0.01	-0.0282
50	0.80	5	0.83	0.79	0.4622	0.5087	0.04	-0.0464
50	0.80	5	0.83	0.8133	0.4715	0.5085	0.0167	-0.037
200	0.20	5	0.9433	0.82	0.2327	0.2552	0.1233	-0.0225
200	0.20	5	0.9467	0.77	0.2332	0.2517	0.1767	-0.0185
200	0.80	5	0.9367	0.54	0.2277	0.2498	0.3967	-0.0221
200	0.80	5	0.92	0.5867	0.2286	0.2478	0.3333	-0.0192
50	0.20	15	0.8967	0.9233	0.5334	0.6255	-0.0267	-0.0921
50	0.20	15	0.8833	0.8867	0.5386	0.6264	-0.0033	-0.0878
50	0.80	15	0.8367	0.7567	0.4809	0.5956	0.08	-0.1147
50	0.80	15	0.8	0.75	0.4881	0.5889	0.05	-0.1008
200	0.20	15	0.9467	0.79	0.2435	0.3056	0.1567	-0.062
200	0.20	15	0.9367	0.74	0.2439	0.3065	0.1967	-0.0626
200	0.80	15	0.93	0.4567	0.2302	0.2813	0.4733	-0.0511
200	0.80	15	0.9133	0.5	0.2292	0.2904	0.4133	-0.0612

Table 3.5: Comparison of performances in estimating the indirect effect including/excluding the GF, varying sample size and correlation (ρ) between the spatial processes u_M and u_Y.

n	ρ	coverage (GF)	avg. len. (GF)	coverage (no GF)	avg. len. (no GF)	coverage diff.	avg. len. diff.
50	0	0.9270	0.2375	0.8719	0.2560	0.0552	-0.0185
50	0.2	0.8770	0.4800	0.7508	0.5132	0.1262	-0.0332
50	0.4	0.8385	0.4902	0.6935	0.5382	0.1450	-0.0480
50	0.6	0.7277	0.4964	0.6153	0.5449	0.1124	-0.0484
50	0.8	0.6016	0.4934	0.4387	0.5462	0.1629	-0.0528
200	0	0.7436	0.3605	0.4532	0.4024	0.2904	-0.0419
200	0.2	0.9204	0.2294	0.5246	0.2527	0.3958	-0.0233
200	0.4	0.8808	0.2309	0.4185	0.2605	0.4623	-0.0296
200	0.6	0.7893	0.2348	0.2468	0.2649	0.5425	-0.0301
200	0.8	0.6182	0.2390	0.1016	0.2682	0.5166	-0.0291

Table 3.6: Posterior statistics for parameters and hyperparameters of the mediation model.

	Mean	**Sd**	**Median**	**Mode**	**HPD low**	**HPD high**
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

Conclusions

The doctoral program that gave rise to this thesis was marked by two engaging collaborations, one with the Sicilian Epidemiological Observatory in the healthcare sector and the other with the Department of Marine Biology at the University of Palermo in the field of marine ecology. These partnerships offered a unique platform to delve into two distinct statistical domains: spatial statistics in ecology and mediation analysis in health research. They also facilitated the application of these statistical approaches to investigate real-world problems of actual interest.

We opted to perform our statistical analysis within the Bayesian inferential framework, renowned for its significant flexibility in handling hierarchical structures in spatial analysis. This approach is gaining momentum in mediation analysis as well, as evidenced by an expanding body of literature, as addressed by Yuan and MacKinnon (2009b). Presently, the computational challenges historically associated with Bayesian analysis are being effectively addressed through modern estimation methods like INLA. This advancements considerably enhances the range of options available to researchers in terms of modeling versatility, as we've experienced first-hand.

Specifically, we've applied the INLA-SPDE modelling approach investigating the effects of environmental temperature on the MTC of demersal fish communities in the central Mediterranean Sea. It has been a particularly gratifying aspect of this research to delve into a topic related to climate change, given the current prominence and urgency of this issue on a global scale. INLA's computational efficiency, as well as its modelling flexibility, were effectively exploited: by using the SPDE approximation to Gaussian fields, along with a stochastic smoother for the trend, we were allowed to quantify the effect of environmental temperature and depth on the MTC, as well as to describe the spatio-temporal evolution of the MTC in the study area. To the best of our knowledge, this is the first time that this methodological approach is applied to investigate the relation between climate change and thermal affinity of marine communities.

Our findings suggest the need for customised strategies to protect different fish

communities, indicating a shift away from optimal temperature conditions for marine communities, and a spatial redistribution of species due to warming waters. Our results also highlighted the importance of considering depth in fisheries management and conservation. While our study employed the MTC as a synthetic index for all collected species, future analyses in a multivariate framework could be directed towards investigating the species that are facing the most challenges.

In the second Chapter of the thesis the focus shifts on mediation analysis, particularly on the estimation of mediational effects in a non-linear setting. First, in the area of healtcare research, we investigated the effect of socio-economic status on cancer mortality as mediated by out-of-region mobility, focusing on lung and colon cancer patients in Sicily, Italy. In this context, as it is the case in many real-world applications, the presence of not normally distributed variables implies that it is not possible to estimate mediational effects using traditional approaches such as the *product method*. We found it helpful to use a two-step approach, first by deriving the structure of association among the variables from the data using log-linear graphical models as an exploratory tool, and than estimating mediational effects using a counterfactual-based approach. Our results provides evidence on the existence of potential equity threats in the provisioning of healthcare services, due to the economic barriers faced by lower socio-economic groups. To further investigate this issue further, important dynamic factors over time could be addressed in future longitudinal studies using, for example, approaches developed in Didelez (2019) and Aalen et al. (2020).

While mediation analysis with nonlinear models has primarily been addressed using non-parametric counterfactual-based approaches, as is the one we used in the above study, very few solutions have been proposed to deal with the issue of nonlinear mediation models in a path-analytic framework. An exception is given by a generalisation of the product method based on partial derivatives, proposed in the '80s by Stolzenberg (1980), and recently revived by Hayes and Preacher (2010) and Geldhof et al. (2018). Although quite intuitive, this approach is not widely known and applied by practitioners, and for this reason it is not well developed, presenting theoretical shortcomings yet to be addressed.

For that matter, in the last part of the second Chapter we offered a comprehensive discussion of the derivative-based method for nonlinear mediation analysis by deepening some aspects of the proposal of Geldhof et al. (2018). We addressed several potential issues and proposed possible solutions. Examples are the presence of binary mediators and the corresponding choice of predicted values for the estimation of CIEs, and dealing with binary/categorical/discrete exposure variables. We

also proposed the use of a Bayesian approach as a valuable option for the estimation of CIEs confidence intervals. Through a simulation study, we provided evidence that Bayesian intervals are a valid alternative compared to the Bootstrap-based and Montecarlo ones used by Geldhof et al. (2018).

Also, we presented a real-data application which investigated the relationship between high-school background and academic success, in which we dealt with the presence of a discrete exposure variable. This posed a limitation for applying the derivative-based approach, as the concept of derivative lacks meaning in the context of discrete variables. To overcome this challenge, we employed finite differences to estimate the conditional indirect effects. By calculating finite differences, we were able to capture the effects of the mediator on the outcome variable while accounting for the discreteness of the high-school background variable. The results obtained through this approach were then interpreted accordingly, acknowledging the specific characteristics of the variables involved.

We believe that this work can serve as a guide for researchers who need to address a mediational setting with nonlinear models without switching to the counterfactual framework. This work can be extended in several ways, for example by investigating the relationship between the total and indirect effects, or by moving the research into the venue of clustered data.

In the course of this study we eventually began to contemplate the integration of two specific methodologies we had investigated: spatial analysis using INLA and the derivative-based approach in mediation analysis. This consideration emerged particularly in practical scenarios where mediation analysis needs to account for spatial correlation. For instance, within the scope of our collaborative projects, we encountered issues such as examining the impact of socioeconomic status on cardiovascular disease mortality, as mediated by specific air pollutant levels. Although it is still a work-in-progress, we identified and highlighted a potential pathway for future research in this direction, and reported the evidences found so far.

For this matter, the last part of the thesis provides insights regarding the possibility to combine the INLA-SPDE method and the derivative-based approach to mediation analysis, within the Bayesian inferential paradigm, for estimating nonlinear mediational effects with spatially-correlated data. Although recent explorations into spatial variability in causal inference (Reich et al., 2021) indicate viable paths, the literature on mediation analysis predominantly lacks consideration for spatial heterogeneity. Barely mentioned in Reich et al. (2021) is the possibility of using the INLA-SPDE approach to estimate causal effects with spatially-correlated data, viewing that as an extension of Simultaneous Autoregressive (SAR) models to

the continuous spatial domain.

Contributing to fill the gap, using simulations we illustrated the advantages of considering spatial correlation in mediation analysis, quantifying how neglecting spatial correlation in the data impacts on the estimation of the indirect effect. We further demonstrate the practical utility of this approach by applying it to real ecological data. We acknowledge that many aspects deserve further investigation, for example by considering the possible correlation between spatial effects at the mediator and at the outcome level. Overall, this work underscores the need for continued exploration and development of methodologies for nonlinear mediation analysis in diverse real-world scenarios.

Bibliography

Aalen, O. O., Stensrud, M. J., Didelez, V., Daniel, R., Røysland, K., and Strohmaier, S. (2020). Time-dependent mediators in survival analysis: Modeling direct and indirect effects with the additive hazards model. *Biometrical Journal*, 62:532–549.

Abatemarco, A., Aria, M., Beraldo, S., and Stroffolini, F. (2020). Measuring disparities in access to health care: A proposal based on an ex-ante perspective. *Social Indicators Research*, pages 1–20.

Abrahamsen, P. (1997). A review of gaussian random fields and correlation functions. Technical Report 917, Norwegian Computing Center.

Adler, N. E., Boyce, T., Chesney, M. A., Cohen, S., Folkman, S., Kahn, R. L., and Syme, S. L. (1994). Socioeconomic status and health: the challenge of the gradient. *American psychologist*, 49(1):15.

Aggarwal, A., van der Geest, S. A., Lewis, D., van der Meulen, J., and Varkevisser, M. (2020). Simulating the impact of centralization of prostate cancer surgery services on travel burden and equity in the english national health service: A national population based model for health service re-design. *Cancer medicine*, 9(12):4175–4184.

Aggarwal, A. K., Sujenthiran, A., Lewis, D., Walker, K., Cathcart, P., Clarke, N., Sullivan, R., and van der Meulen, J. H. (2019). Impact of patient choice and hospital competition on patient outcomes after prostate cancer surgery: A national population-based study. *Cancer*, 125(11):1898–1907.

Albert, J. M. and Nelson, S. (2011). Generalized Causal Mediation Analysis. *Biometrika*, 67:1028–1039.

Bakka, H., Rue, H., Fuglstad, G.-A., Riebler, A., Bolin, D., Illian, J., Krainski, E., Simpson, D., and Lindgren, F. (2018). Spatial modeling with r-inla: A review. *Wiley Interdisciplinary Reviews: Computational Statistics*, 10(6):e1443.

Balia, S., Brau, R., and Marrocu, E. (2014). What drives patient mobility across italian regions? evidence from hospital discharge data. In *Health Care Provision and Patient Mobility*, pages 133–154. Springer.

Banerjee, S., Carlin, B. P., and Gelfand, A. E. (2003). *Hierarchical modeling and analysis for spatial data*. Chapman and Hall/CRC.

Baptista, V., Ullah, H., Teixeira, C. M., Range, P., Erzini, K., and Leitão, F. (2014). Influence of environmental variables and fishing pressure on bivalve fisheries in an inshore lagoon and adjacent nearshore coastal area. *Estuaries and Coasts*, 37(1):191–205.

Baron, R. M. and Kenny, D. A. (1986a). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6):1173–1182.

Baron, R. M. and Kenny, D. A. (1986b). The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic and Statistical Considerations. *Journal of Personality and Social Psychology*, 51(6):1173–1182.

Bauer, D. J., Preacher, K. J., and Gil, K. M. (2006). Conceptualizing and testing random indirect effects and moderated mediation in multilevel models: new procedures and recommendations. *Psychological Methods*, 11(2):142–163.

Bedir, A., Abera, S. F., Efremov, L., Hassan, L., Vordermark, D., and Medenwald, D. (2021). Socioeconomic disparities in head and neck cancer survival in germany: a causal mediation analysis using population-based cancer registry data. *Journal of cancer research and clinical oncology*, 147(5):1325–1334.

Berrar, D. P. (2019). Performance measures for binary classification. In *Encyclopedia of Bioinformatics and Computational Biology*.

Berta, P., Guerriero, C., and Levaggi, R. (2021). Hospitals' strategic behaviours and patient mobility: Evidence from italy. *Socio-Economic Planning Sciences*, page 101030.

Bertrand, J. A., Gil De Sola, L., Papaconstantinou, C., Relini, G., and Souplet, A. (2002). The general specifications of the medits surveys. *Scientia Marina*, 66(S2):9–17.

Biesanz, J. C., Falk, C. F., and Savalei, V. (2010). Assessing Mediational Models: Testing and Interval Estimation for Indirect Effects. *Multivariate Behavioral Research*, 45(4):661–701.

Bisanzio, D., Giacobini, M., Bertolotti, L., Mosca, A., Balbo, L., Kitron, U., and Vazquez-Prokopec, G. M. (2011). Spatio-temporal patterns of distribution of West Nile virus vectors in eastern Piedmont Region. *Parasit Vectors*, 4:230.

Bleuel, J., Pennino, M. G., and Longo, G. O. (2021). Coral distribution and bleaching vulnerability areas in southwestern atlantic under ocean warming. *Scientific reports*, 11(1):12833.

Bollen, K. A. (1989). *Structural Equations with Latent Variables*. Wiley, New York.

Booth, C. M., Li, G., Zhang-Salomons, J., and Mackillop, W. J. (2010). The impact of socioeconomic status on stage of cancer at diagnosis and survival: a population-based study in ontario, canada. *Cancer*, 116(17):4160–4167.

Cameletti, M., Lindgren, F., Simpson, D., and Rue, H. (2013). Spatio-temporal modeling of particulate matter concentration through the SPDE approach. *AStA Advances in Statistical Analysis*, 97:109–131.

Caranci, N., Biggeri, A., Grisotto, L., Pacelli, B., Spadea, T., and Costa, G. (2010). The italian deprivation index at census block level: definition, description and association with general mortality. *Epidemiologia e prevenzione*, 34(4):167–176.

Charlson, M., Szatrowski, T. P., Peterson, J., and Gold, J. (1994). Validation of a combined comorbidity index. *Journal of clinical epidemiology*, 47(11):1245–1251.

Cheung, W., Watson, R., and Pauly, D. (2013). Signature of ocean warming in global fisheries catch. *Nature*, 497(7449):365–368.

Cheung, W. W., Lam, V. W., Sarmiento, J. L., Kearney, K., Watson, R., Zeller, D., and Pauly, D. (2010). Large-scale redistribution of maximum fisheries catch potential in the global ocean under climate change. *Global Change Biology*, 16(1):24–35.

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

Cislaghi, C., Giuliani, F., Arena, V., and Olivadoti, S. (2013). *Impatto della patologia oncologica ed analisi delle prestazioni e della mobilità in Sicilia*. POAT Salute 2007-2013, Supporto allo svolgimento delle diverse fasi che accompagnano il ciclo della programmazione.

Colloca, F., Cardinale, M., Maynou, F., Giannoulaki, M., Scarcella, G., Jenko, K., Bellido, J. M., and Fiorentino, F. (2013). Rebuilding mediterranean fisheries: a new paradigm for ecological sustainability. *Fish and Fisheries*, 14(1):89–109.

Consoli, P., Esposito, V., Battaglia, P., Altobelli, C., Perzia, P., Romeo, T., Canese, S., and Andaloro, F. (2016). Fish distribution and habitat complexity on banks of the strait of sicily (central mediterranean sea) from remotely-operated vehicle (rov) explorations. *PLoS One*, 11(12):e0167809.

Cosandey-Godin, A., Krainski, E. T., Worm, B., and Flemming, J. M. (2014). Applying Bayesian spatiotemporal models to fisheries bycatch in the Canadian Arctic. *Canadian Journal of Fisheries and Aquatic Sciences*, 72:1–12.

Cressie, N. (1993). *Statistics for spatial data*. Wiley, Hoboken.

Cressie, N. and Wickle, C. (2011). *Statistics for Spatiotemporal Data*. Wiley, New York.

Curran, P. J. (2003). Have Multilevel Models Been Structural Equation Models All Along? *Multivariate Behavioral Res.*, 38(4):529–569.

Davis, M. L., Neelon, B., Nietert, P. J., Hunt, K. J., Burgette, L. F., Lawson, A. B., and Egede, L. E. (2019). Addressing geographic confounding through spatial propensity scores: a study of racial disparities in diabetes. *Statistical Methods in Medical Research*, 28(3):734–748.

Dawid, A. P. (1992). Applications of a general propagation algorithm for probabilistic expert systems. *Statistics and computing*, 2(1):25–36.

Denteneer, D. and Verbeek, A. (1985). A fast algorithm for iterative proportional fitting in log-linear models. *Computational Statistics & Data Analysis*, 3:251–264.

Di Lorenzo, M., Sinerchia, M., and Colloca, F. (2018). The north sector of the strait of sicily: a priority area for conservation in the mediterranean sea. *Hydrobiologia*, 821(1):235–253.

Di Maria, C., Abbruzzo, A., and Lovison, G. (2022). Bayesian causal mediation analysis through linear mixed-effect models. In *Book of Short Papers-SIS 2022*. Springer.

Di Maria, C., Rubino, C., and A., A. (2024). The derivative-based approach to nonlinear mediation models: insights and applications. *Quality & Quantity*. In publication.

Didelez, V. (2019). Defining causal meditation with a longitudinal mediator and a survival outcome. *Lifetime Data Analysis*, 25(4):593–610.

Doney, S. C., Ruckelshaus, M., Duffy, J. E., Barry, J. P., Chan, F., et al. (2012). Climate change impacts on marine ecosystems. *Annual Review of Marine Science*, 4:11–37.

Doretti, M., Raggi, M., and Stanghellini, E. (2022). Exact parametric causal mediation analysis for a binary outcome with a binary mediator. *Statistical Methods & Applications*, 31:87–108.

Dulčić, J. and Grbec, B. (2000). Climate change and its influence on the biology and management of the dentex (Pagellus erythrinus) in the adriatic sea. *Climatic Change*, 49(3):285–293.

Dulvy, N. K., Rogers, S. I., Jennings, S., Stelzenmüller, V., Dye, S. R., and Skjoldal, H. R. (2008). Climate change and deepening of the north sea fish assemblage: a biotic indicator of warming seas. *Journal of Applied Ecology*, 45(4):1029–1039.

Efron, B. and Tibshirani, R. J. (1994). *An introduction to the bootstrap*. CRC press.

Fahrmeir, L. and Lang, S. (2001). Bayesian inference for generalized additive mixed models based on markov random field priors. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 50(2):201–220.

Fong, Y., Rue, H., and Wakefield, J. (2010). Bayesian inference for generalized linear mixed models. *Biostatistics*, 11:397–412.

Fortibuoni, T., Aldighieri, F., Giovanardi, O., Pranovi, F., and Zucchetta, M. (2015). Climate impact on italian fisheries (mediterranean sea). *Regional Environmental Change*, 15:931–937.

Gaynor, S. M., Schwartz, J., and Lin, X. (2019). Mediation analysis for common binary outcomes. *Statistics in Medicine*, 38:512–529.

Geldhof, G. J., Anthony, K. P., Selig, J. P., and Mendez-Luck, C. A. (2018). Accommodating binary and count variables in mediation: A case for conditional indirect effects. *International Journal of Behavioral Development*, 42(2):300–308.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2013). *Bayesian data analysis*. CRC press.

Gómez-Rubio, V., Bivand, R., and Rue, H. (2015a). A new latent class to fit spatial econometrics models with integrated nested Laplace approximations. *Procedia Environmental Sciences*, 27:116–118.

Gómez-Rubio, V., Cameletti, M., and Finazzi, F. (2015b). Analysis of massive marked point patterns with stochastic partial differential equations. *Spatial Statistics*, 14:179–196.

Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: journal of the Econometric Society*, pages 424–438.

Green, P. J. (2005). Grappa: R functions for probability propagation. *URL http://www*, 287.

Halpern, M. T., Ward, E. M., Pavluck, A. L., Schrag, N. M., Bian, J., and Chen, A. Y. (2008). Association of insurance status and ethnicity with cancer stage at diagnosis for 12 cancer sites: a retrospective analysis. *The lancet oncology*, 9(3):222–231.

Harley, C. D. G., Randall Hughes, A., Hultgren, K. M., Miner, B. G., Sorte, C. J. B., et al. (2006). The impacts of climate change in coastal marine systems. *Ecology Letters*, 9(2):228–241.

Hastie, T., Tibshirani, R., Friedman, J. H., and Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction*, volume 2. Springer.

Hayes, A. F. and Preacher, K. J. (2010). Quantifying and Testing Indirect Effects in Simple Mediation Models When the Constituent Paths Are Nonlinear. *Multivariate Behavioral Research*, 45:627–660.

Helmuth, B., Kingsolver, J. G., and Carrington, E. (2005). Biophysics, physiological ecology, and climate change: does mechanism matter? *Annu. Rev. Physiol.*, 67:177–201.

Hoegh-Guldberg, O. and Bruno, J. F. (2010). The impact of climate change on the world's marine ecosystems. *Science*, 328(5985):1523–1528.

Højsgaard, S., Edwards, D., and Lauritzen, S. (2012). *Graphical models with R*. Springer Science & Business Media.

Illian, J. B. and Sørbye, Rue, H. et al., S. H. (2012). A toolbox for fitting complex spatial point process models using integrated nested Laplace approximation (INLA). *The Annals of Applied Statistics*, 6:1499–1530.

Imai, K., Tingley, D., and Keele, L. (2010). A general approach to causal mediation analysis. *Psychological methods*, 15(4):309–334.

Jarner, M. F., Diggle, P., and Chetwynd, A. G. (2002). Estimation of spatial variation in risk using matched case-control data. *Biometrical Journal: Journal of Mathematical Methods in Biosciences*, 44(8):936–945.

Kenny, D. A., Korchmaros, J. D., and Bolger, N. (2003). Lower Level Mediation in Multilevel Models. *Psychol. Methods*, 8(2):115–128.

Keskin, C. and Pauly, D. (2014). Changes in the 'mean temperature of the catch': application of a new concept to the north-eastern aegean sea. *Acta Adriatica*, 55(2):213–218.

Knisely, A., Huang, Y., Melamed, A., Tergas, A. I., Clair, C. M. S., Hou, J. Y., Khoury-Collado, F., Ananth, C. V., Neugut, A. I., Hershman, D. L., et al. (2020). Effect of regionalization of endometrial cancer care on site of care and patient travel. *American journal of obstetrics and gynecology*, 222(1):58–e1.

Koopman, J., Howe, M., Hollenbeck, J. R., and Sin, H.-P. (2015). Small sample mediation testing: Misplaced confidence in bootstrapped confidence intervals. *Journal of Applied Psychology*, 100(1):194–202.

Krainski, E. T., Gómez-Rubio, V., Bakka, H., Lenzi, A., Castro-Camilo, D., Simpson, D., Lindgren, F., and Rue, H. (2019). *Advanced Spatial Modeling with Stochastic Partial Differential Equations Using R and INLA*. CRC Press/Taylor and Francis Group, Boca Raton.

Krull, J. L. and MacKinnon, D. P. (1999). Multilevel mediation modeling in group-based intervention studies. *Evaluation Review*, 23(4):418–444.

Krull, J. L. and MacKinnon, D. P. (2001). Multilevel Modeling of Individual and Group Level Mediated Effects. *Multivariate Behavioral Research*, 36(2):249–277.

Lauritzen, S. L. (1996). *Graphical models*, volume 17. Clarendon Press.

Legendre, P. and Legendre, L. (1998). *Numerical ecology*. Elsevier.

Leitão, F., Alms, V., and Erzini, K. (2014). A multi-model approach to evaluate the role of environmental variability and fishing pressure in sardine fisheries. *Journal of Marine Systems*, 139:128–138.

Levitus, S., Antonov, J. I., Boyer, T. P., and Stephens, C. (2000). Warming of the world ocean. *Science*, 287(5461):2225–2229.

Lezama-Ochoa, N., Pennino, M. G., Hall, M. A., Lopez, J., and Murua, H. (2020). Using a bayesian modelling approach (inla-spde) to predict the occurrence of the spinetail devil ray (mobular mobular). *Scientific reports*, 10(1):18822.

Li, R., Daniel, R., and Rachet, B. (2013). How much of socioeconomic differences in survival in patients with breast cancer can be explained by differences in stage of diagnosis and treatment? application of causal mediation analysis to routine data. *The Lancet*, 382:S61.

Li, R., Daniel, R., and Rachet, B. (2016). How much do tumor stage and treatment explain socioeconomic inequalities in breast cancer survival? applying causal mediation analysis to population-based data. *European journal of epidemiology*, 31(6):603–611.

Lindgren, F. and Rue, H. (2015). Bayesian spatial modelling with r-inla. *Journal of statistical software*, 63(19).

Lindgren, F., Rue, H., and Lindstrom, J. (2011). An explicit link between Gaussian fields and Gaussian Markov random fields: The stochastic partial differential equation approach. *Journal of Royal Statistical Society, Series B*, 73:423–498.

Lomnicki, Z. A. (1967). On the Distribution of Products of Random Variables. *Journal of the Royal Statistical Society. Series B*, 29(3):513–524.

MacKinnon, D. P. (2008a). *Introduction to Statistical Mediation Analysis*. Taylor and Francis Group, New York.

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

MacKinnon, D. P. and Dwyer, J. H. (1993). Estimating mediated effects in prevention studies. *Evaluation Review*, 17:144–158.

Maharaj, R. R., Lam, V. W., Pauly, D., and Cheung, W. W. (2018). Regional variability in the sensitivity of caribbean reef fish assemblages to ocean warming. *Marine Ecology Progress Series*, 590:201–209.

Martino, S., Akerkar, R., and Rue, H. (2011). Approximate Bayesian inference for survival models. *Scandinavian Journal of Statistics*, 38:514–528.

Martins, T. G., Simpson, D., Lindgren, F., and Rue, H. (2013). Bayesian computing with inla: new features. *Computational Statistics & Data Analysis*, 67:68–83.

Massutí, E., Morales-Nin, B., and Lloris, D. (2000). Fish community structure and depth-related trends on the continental shelf of the balearic islands (alboran sea, western mediterranean). *Marine Ecology Progress Series*, 211:157–167.

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

Meleddu, M., Pulina, M., and Scuderi, R. (2020). Public and private healthcare services: What drives the choice? *Socio-Economic Planning Sciences*, 70:100739.

Ministero della Salute (2020). Rapporto annuale sull'attività di ricovero ospedaliero. Dati SDO 2018.

Miočević, M., MacKinnon, D. P., and Levy, R. (2017). Power in Bayesian Mediation Analysis for Small Sample Research. *Structural Equation Modeling*, 24(5):666–683.

Moten, A. S., von Mehren, M., Reddy, S., Howell, K., Handorf, E., and Farma, J. M. (2020). Treatment patterns and distance to treatment facility for soft tissue sarcoma of the extremity. *Journal of Surgical Research*, 256:492–501.

Munoz, F., Pennino, M. G., Conesa, D., Lopez-Quilez, A., and Bellido, J. M. (2013). Estimation and prediction of the spatial occurrence of fish species using bayesian latent gaussian models. *Stochastic Environmental Research and Risk Assessment*, 27:1171–1180.

Olden, J. D. and Neff, B. D. (2001). Cross correlation bias in lag analysis of aquatic time series. *Marine Biology*, 138:1063–1070.

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

Pennino, M. G., Munoz, F., Conesa, D., López-Quílez, A., and Bellido, J. M. (2014). Bayesian spatio-temporal discard model in a demersal trawl fishery. *Journal of sea research*, 90:44–53.

Perry, A. L., Low, P. J., Ellis, J. R., and Reynolds, J. D. (2005). Climate change and distribution shifts in marine fishes. *Science*, 308(5730):1912–1915.

Pickett, K. E. and Pearl, M. (2001). Multilevel analyses of neighbourhood socioeconomic context and health outcomes: a critical review. *Journal of Epidemiology & Community Health*, 55(2):111–122.

Poloczanska, E., Brown, C., Sydeman, W., et al. (2013). Global imprint of climate change on marine life. *Nature Climate Change*, 3:919–925.

Pörtner, H.-O. and Farrell, A. P. (2008). Physiology and climate change. *Science*, 322(5902):690–692.

Preacher, K. J., Zhang, Z., and Zyphur, M. J. (2011). Alternative Methods for Assessing Mediation in Multilevel Data: The Advantages of Multilevel SEM. *Structural Equation Modeling*, 18(2):161–182.

Preacher, K. J., Zyphur, M. J., and Zhang, Z. (2010). A General Multilevel SEM Framework for Assessing Multilevel Mediation. *Psychological Methods*, 15(3):209–233.

Priulla, A. (2023). *Inequalities in student performances in the Italian universities*. PhD Thesis, University of Palermo. Available at `https://iris.unipa.it/handle/10447/582705`.

Propper, C., Damiani, M., Leckie, G., and Dixon, J. (2007). Impact of patients' socioeconomic status on the distance travelled for hospital admission in the english national health service. *Journal of Health Services Research & Policy*, 12(3):153–159.

Reich, B. J., Yang, S., Guan, Y., Giffin, A. B., Miller, M. J., and Rappold, A. (2021). A review of spatial causal inference methods for environmental and epidemiological applications. *International Statistical Review*, 89(3):605–634.

Rijnhart, J. J. M., Valente, M. J., MacKinnon, D. P., Twisk, J. W. R., and Heymans, M. W. (a2021a). The use of traditional and causal estimators for mediation models with a binary outcome and exposure-mediator interaction. *Struct. Equ. Model.*, 28(3):345–355.

Rijnhart, J. J. M., Valente, M. J., Smyth, H. L., and MacKinnon, D. P. (b2021b). Statistical mediation analysis for models with a binary mediator and a binary outcome: the differences between causal and traditional mediation analysis. *Prev. Sci.*, pages 1–11.

Robert, C. P., Casella, G., and Casella, G. (1999). *Monte Carlo statistical methods*, volume 2. Springer.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology*, 66(5):688.

Rubino, C., Adelfio, G., Abbruzzo, A., Bosch-Belmar, M., Lorenzo, M. D., Fiorentino, F., Gancitano, V., Colloca, F., and Milisenda, G. (2024). Exploring the effects of temperature on demersal fish communities in the central mediterranean sea using INLA-SPDE modeling approach. *Environmental and Ecological Statistics*. Minor revisions (2nd round).

Rubino, C., Di Maria, C., Abbruzzo, A., and Ferrante, M. (2022a). Socio-economic inequality, interregional mobility and mortality among cancer patients: A mediation analysis approach. *Socio-Economic Planning Sciences*, 82:101247.

Rubino, C., Di Maria, C., et al. (2022b). Insights into the derivative-based method for nonlinear mediation models. In *Book of Short Papers-SIS 2022*. Pearson.

Rubino, C., Milisenda, G., Abbruzzo, A., Adelfio, G., Bosch-Belmar, M., Colloca, F., Lorenzo, M. D., and Gancitano, V. (2021). Sea surface temperature effects on the mediterranean marine ecosystem: a semiparametric model approach. In *Book of Short Papers - SIS 2021*, page 895–900. Pearson.

Rubinstein, R. Y. and Kroese, D. P. (2016). *Simulation and the Monte Carlo method*. John Wiley & Sons.

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

Rue, H., Martino, S., and Chopin, N. (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations (with discussion). *Journal of Royal Statistical Society, Series B*, 71:319–392.

Rue, H., Riebler, A., Sorbye, S. H., Illian, J. B., Simpson, D. P., and Lindgren, F. (2017). Bayesian computing with INLA: A review. *Annual Review of Statistics and Its Application*, 4:395–421.

Russell, B., Hemelrijck, M. V., Gårdmark, T., Holmberg, L., Kumar, P., Bellavia, A., and Häggström, C. (2020). A mediation analysis to explain socio-economic differences in bladder cancer survival. *Cancer medicine*, 9(20):7477–7487.

Schluchter, M. D. (2008). Flexible Approaches to Computing Mediated Effects in Generalized Linear Models: Generalized Estimating Equations and Bootstrapping. *Multivariate Behavioral Research*, 43(2):268–288.

Schnell, P. and Papadogeorgou, G. (2019). Mitigating unobserved spatial confounding bias with mixed models. *arXiv preprint arXiv:1907.12150*.

Schrodle, B. and Held, L. (2011). Spatio-temporal disease mapping using INLA. *Environmetrics*, 22:725–734.

Schrodle, B., Held, L., and Rue, H. (2012). Assessing the impact of a movement network on the spatiotemporal spread of infectious diseases. *Biometrics*, 68:736–744.

Shavers, V. L. (2007). Measurement of socioeconomic status in health disparities research. *Journal of the national medical association*, 99(9):1013.

Simpson, D., Illian, J. B., Lindgren, F., Sørbye, S. H., and Rue, H. (2016). Going off grid: Computationally efficient inference for log-gaussian cox processes. *Biometrika*, 103:49–70.

Smithson, M. and Verkuilen, J. (2006). A better lemon squeezer? Maximum-Likelihood Regression with Beta-Distributed Dependent Variables. *Psychol. Methods*, 11(1):54.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 64(4):583–639.

Springer, M. D. and Thompson, W. E. (1966). The Distribution of Products of Independent Random Variables. *SIAM Journal on Applied Mathematics*, 14(3):511–526.

Stitzenberg, K. B., Sigurdson, E. R., Egleston, B. L., Starkey, R. B., and Meropol, N. J. (2009). Centralization of cancer surgery: implications for patient access to optimal care. *Journal of Clinical Oncology*, 27(28):4671.

Stolzenberg, R. M. (1980). The measurement and decomposition of causal effects in nonlinear and nonadditive models. *Sociological Methodology*, 11:459–488.

Sunday, J. M., Pecl, G. T., Frusher, S., Hobday, A. J., Hill, N., Holbrook, N. J., Edgar, G. J., Stuart-Smith, R., Barrett, N., Wernberg, T., et al. (2015). Species traits and climate velocity explain geographic range shifts in an ocean-warming hotspot. *Ecology letters*, 18(9):944–953.

Tchetgen Tchetgen, E. J., Fulcher, I. R., and Shpitser, I. (2021). Auto-g-computation of causal effects on a network. *Journal of the American Statistical Association*, 116(534):833–844.

Tierney, L. and Kadane, J. B. (1986). Accurate approximations for posterior moments and marginal densities. *Journal of the american statistical association*, 81(393):82–86.

Tingley, D., Yamamoto, T., Hirose, K., Keele, L., and Imai, K. (2014). mediation: R package for causal mediation analysis. *Journal of Statistical Software*, 59:Issue 5.

Tsai, T.-L., Shau, W., and Hu, F. (2006). Generalized Path Analysis and Generalized Simultaneous Equations Model for Recursive Systems with Responses of Mixed Types. *Structural Equation Modeling*, 13(2):229–251.

Tsikliras, A. C., Peristeraki, P., Tserpes, G., and Stergiou, K. I. (2015). Mean temperature of the catch (mtc) in the greek seas based on landings and survey data. *Frontiers in Marine Science*, 2:23.

Tsikliras, A. C. and Stergiou, K. I. (2014). Size at maturity of mediterranean marine fishes. *Reviews in fish biology and fisheries*, 24:219–268.

Turner, M., Jawitz, O., Adam, M., Srinivasan, E., Niedzwiecki, D., Migaly, J., Fisher, D., and Mantyh, C. (2020). Comparison of survival of stage i–iii colon cancer by travel distance and hospital volume. *Techniques in coloproctology*, 24(7):703–710.

VanderWeele, T. (2015). *Explanation in causal inference: methods for mediation and interaction*. Oxford University Press.

VanderWeele, T. J. (2015a). *Explanation in Causal Inference*. Oxford University Press, New York.

VanderWeele, T. J. (2015b). *Explanation in Causal Inference*. Oxford University Press, New York.

Verevkina, N. I., Short, P. F., Yang, T.-C., Matthews, S. A., Camacho, F., and Anderson, R. (2019). Analyzing hospital choices of colon cancer patients in four states in appalachia. *Journal of health care for the poor and underserved*, 30(2):587–608.

Wainwright, M. J. (2015). Graphical models and message-passing algorithms: Some introductory lectures. In *Mathematical Foundations of Complex Networked Information Systems*, pages 51–108. Springer.

Wang, X., Yue, Y. R., and Faraway, J. J. (2018). *Bayesian regression modeling with INLA*. CRC Press.

Whittaker, J. (2009). *Graphical models in applied multivariate statistics*. Wiley Publishing.

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

Woods, L., Rachet, B., and Coleman, M. (2006). Origins of socio-economic inequalities in cancer survival: a review. *Annals of oncology*, 17(1):5–19.

Wright, S. (1934). The Method of Path Coefficients. *The Annals of Mathematical Statistics*, 5(3):161–215.

Yu, X. Q., O'Connell, D. L., Gibberd, R. W., and Armstrong, B. K. (2008). Assessing the impact of socio-economic status on cancer survival in new south wales, australia 1996–2001. *Cancer Causes & Control*, 19(10):1383–1390.

Yuan, Y. and MacKinnon, D. P. (2009a). Bayesian Mediation Analysis. *Psychological Methods*, 14(4):301–322.

Yuan, Y. and MacKinnon, D. P. (2009b). Bayesian mediation analysis. *Psychol. Methods*, 14(4):301.

Zigler, C. M., Dominici, F., and Wang, Y. (2012). Estimating causal effects of air quality regulations using principal stratification for spatially correlated multivariate intermediate outcomes. *Biostatistics*, 13(2):289–302.