Received: 29 March 2023 | Revised: 21 August 2023 | Accepted: 30 August 2023

DOI: 10.1002/bimj.202300096

Biometrical Journal

RESEARCH ARTICLE

A scalable approach for short-term disease forecasting in high spatial resolution areal data

Erick Orozco-Acosta[1,2] | **Andrea Riebler**[3] | **Aritz Adin**[1,2] | **Maria D. Ugarte**[1,2]

[1]Department of Statistics, Computer Science and Mathematics, Public University of Navarre, Pamplona, Spain

[2]Institute for Advanced Materials and Mathematics, InaMat[2], Public University of Navarre, Pamplona, Spain

[3]Department of Mathematical Sciences, Norwegian University of Science and Technology, Trondheim, Norway

Correspondence
Maria D. Ugarte, Department of Statistics, Computer Science and Mathematics, Public University of Navarre, 31006 Pamplona, Spain.
Email: lola@unavarra.es

Funding information
MCIN/AEI/ 10.13039/501100011033, Grant/Award Number: PID2020-113125RB

This article has earned an open data badge “**Reproducible Research**” for making publicly available the code necessary to reproduce the reported results. The results reported in this article were reproduced partially due to confidentiality issues.

Abstract

Short-term disease forecasting at specific discrete spatial resolutions has become a high-impact decision-support tool in health planning. However, when the number of areas is very large obtaining predictions can be computationally intensive or even unfeasible using standard spatiotemporal models. The purpose of this paper is to provide a method for short-term predictions in high-dimensional areal data based on a newly proposed “divide-and-conquer” approach. We assess the predictive performance of this method and other classical spatiotemporal models in a validation study that uses cancer mortality data for the 7907 municipalities of continental Spain. The new proposal outperforms traditional models in terms of mean absolute error, root mean square error, and interval score when forecasting cancer mortality 1, 2, and 3 years ahead. Models are implemented in a fully Bayesian framework using the well-known integrated nested Laplace estimation technique.

KEYWORDS

cancer projections, disease mapping, high-dimensional data, integrated nested Laplace approximation

1 | INTRODUCTION

Bayesian hierarchical models have been developed extensively to model area-level incidence or mortality data and to estimate their underlying spatial, temporal, and spatiotemporal patterns. Traditionally, generalized linear mixed models including spatially and temporally structured random effects have been proposed for smoothing disease risks or rates by borrowing information from neighboring areas and time periods. In addition, extensions of these hierarchical models have also been considered for forecasting of rare and noncommunicable diseases in areal data. For example, Assuncao et al. (2001) provide an extension of the parametric model of Bernardinelli et al. (1995) to predict human visceral Leishmaniasis incidence rates in 117 health zones of a Brazilian municipality. Etxeberria et al. (2014) extend the nonparametric models proposed by Knorr-Held (2000) that include conditional autoregressive (CAR) priors for space and random walk (RW) priors for time for short-term cancer mortality risk predictions in the 50 provinces of Spain. Similar projections of cancer mortality risks using spatiotemporal P-spline models were also considered by Ugarte et al. (2012) and Etxeberria et al. (2015). In Corpas-Burgos and Martinez-Beneito (2021), an enhancement of a previous autoregressive (AR) spatiotemporal model proposed by Martínez-Beneito et al. (2008) was considered for 5-year ahead forecasting of different cancer site mortality data in the 540 municipalities of the Valencian autonomous region of Spain. Etxeberria et al. (2023) predict incidence rates for rare and lethal cancers by borrowing strength from mortality data using spatiotemporal models with shared spatial and age components. Different extensions of age-period-cohort models including spatial random effects have also been proposed for the prediction of cancer mortality and incidence data (see Lagazio et al., 2003; Papoila et al., 2014; Schmid & Held, 2004; or Etxeberria et al., 2017 among others). Very recently, several extensions of different CAR, AR, and RW models have also been proposed for representing the geographical variation of risk processes that underlay the dynamic outbreaks of COVID-19 infection and related outcomes (see, e.g., MacNab, 2023, and the references therein).

All these models perform well when the spatial domain has a limited number of areas. If the number of areas is very large, the model fitting becomes computationally expensive or even unfeasible (Van Niekerk et al., 2021), mainly due to the huge dimension of the spatiotemporal covariance/structure matrices and the high number of identifiability constraints (Goicoa et al., 2018; Schrödle & Held, 2011). However, forecasting short-term disease risks or rates in high-resolution areal data is very important to take high-impact decisions in health planning and addressing health inequalities (Sartorius et al., 2021; Utazi et al., 2019). Cancer mortality projections also play an important role in epidemiology, as they support the decision-making process for population intervention plans and health resource planning. According to the Spanish Statistical Institute, cancer was the second leading cause of mortality in Spain in 2021 after cardiovascular diseases (24.3% and 22.8%, respectively), being the first leading cause of death among the male population. The estimated direct cost of cancer in Spain for the year 2019 was more than 7000 million euros, which represents about 10% of Spanish health costs (Diaz-Rubio, 2019).

The aim of the current paper is to evaluate if the scalable Bayesian spatiotemporal modeling approach proposed by Orozco-Acosta et al. (2023) for estimating risks is also appropriate for short-term forecasting in high spatial resolution areal data. This methodology is based on a “divide-and-conquer” approach and has been shown to provide reliable risk estimates with a substantial reduction in computational time in comparison with classical spatiotemporal CAR models. Specifically, Orozco-Acosta et al. (2023) propose to divide the spatial domain into smaller subregions where independent models can be fitted simultaneously by using parallel or distributed computation strategies. To reduce the border effect in the risk estimates caused by the spatial partitions, neighboring areas are added to each subdomain when fitting the models. Finally, the risk estimates from different submodels are properly combined to obtain unique posterior marginal estimates of the risks for each areal-time unit. This approach has also been extended to high-dimensional multivariate spatial models to jointly analyze several disease outcomes (Vicente et al., 2023). However, it has not been checked yet in a forecasting framework.

The rest of the paper is structured as follows. Sections 2 and 3 outline the methodology and briefly describe the different spatiotemporal models considered for short-term forecasting of cancer mortality. A validation study is presented in Section 4 to assess and compare the predictive performance of the models. In Section 5, the proposed methodology is applied to project male lung cancer and overall cancer (all sites) mortality data by considering 3-year ahead predictions in the 7907 municipalities of continental Spain. The paper concludes with a discussion.

2 | BAYESIAN SPATIOTEMPORAL MODELS

The great variability inherent to classical risk estimation measures such as crude rates when analyzing very small domains or low-populated areas, requires the use of statistical models to smooth risks by borrowing information from spatial and temporal neighbors (Wakefield, 2007). Let O_{it} and N_{it} denote the observed number of cancer deaths and the corresponding number of populations at risk in region $i = 1, \dots, n$ and time period (year) $t = 1, \dots, T$, respectively. Here, we assume that all regions are connected and that years are consecutive. We further assume that the observations are conditionally independent and model them as

$$\begin{aligned} O_{it}|\lambda_{it} &\sim \text{Poisson}(\mu_{it} = N_{it} \cdot \lambda_{it}) \text{ for } i = 1, \dots, n;\ t = 1, \dots, T, \\ \log \mu_{it} &\sim \log N_{it} + \log \lambda_{it}, \end{aligned}$$

where $\log N_{it}$ is an offset and λ_{it} is the mortality rate in region i at time t. Depending on the specification of $\log \lambda_{it}$, we define different models that are all placed in a hierarchical Bayesian inference scheme.

2.1 | Classical Bayesian disease-mapping models

Here, we assume a linear predictor of the form

$$\log \lambda_{it} = \beta_0 + \xi_i + \gamma_t + \delta_{it}, \tag{1}$$

where β_0 is an intercept representing the overall log-rate, ξ_i is a spatial random effect that follows the so-called BYM2 prior distribution (Riebler et al., 2016), γ_t is a temporally structured random effect that follows a first-order random walk (RW1), and δ_{it} is a spatiotemporal random effect allowing for space-time interactions (Knorr-Held, 2000). All the components of this model can be formulated as Gaussian Markov random fields (Rue & Held, 2005), and prior densities can be written according to some structure matrices.

The BYM2 model for the spatial random effect $\boldsymbol{\xi} = (\xi_1, \dots, \xi_n)'$ is expressed as

$$\boldsymbol{\xi} = \frac{1}{\sqrt{\tau_\xi}}\left(\sqrt{\phi}\mathbf{u}_* + \sqrt{1-\phi}\mathbf{v}\right),$$

where τ_ξ is a precision parameter, $\mathbf{u}_* \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_*^-)$ is the scaled intrinsic CAR model with $\mathbf{R}_*^-$, representing the generalized inverse of the standardized neighborhood structure matrix $\mathbf{R}_*$ (see Sørbye & Rue, 2014), $\mathbf{v} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_n)$ is a vector of unstructured random effects, and $\phi \in [0, 1]$ is a parameter that weights the regional variability between the unstructured and spatially structured component. Therefore, the covariance matrix of $\boldsymbol{\xi}$ is

$$\text{Var}(\boldsymbol{\xi}|\tau_\xi) = \frac{1}{\tau_\xi}(\phi\mathbf{R}_*^- + (1-\phi)\mathbf{I}_n),$$

expressed as a weighted average of the covariance matrices of the spatially structured and unstructured components, $\mathbf{R}_*^-$ and $\mathbf{I}_n$, respectively. Values of ϕ larger than 0.5 indicate that more than 50% of the spatial variation is explained by the structured component, indicating the benefits of having a joint model for all regions. For further details, we refer to Riebler et al. (2016).

A RW1 prior distribution is assumed for the temporal random effects $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_T)'$, that is,

$$\boldsymbol{\gamma} \sim N(\mathbf{0}, [\tau_\gamma \mathbf{R}_\gamma]^-),$$

where τ_γ is a precision parameter and $\mathbf{R}_\gamma$ is the $T \times T$ structure matrix defined as

$$\mathbf{R}_\gamma = \begin{pmatrix} 1 & -1 & 0 & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & 0 & -1 & 1 \end{pmatrix}.$$

Finally, for the space-time interaction random effect $\boldsymbol{\delta} = (\delta_{11}, \dots, \delta_{n1}, \dots, \delta_{1T}, \dots, \delta_{nT})'$, the following prior distribution is assumed

$$\boldsymbol{\delta} \sim N(\mathbf{0}, [\tau_\delta \mathbf{R}_\delta]^-),$$

where $\mathbf{R}_\delta$ is the $nT \times nT$ matrix which represents one of the four types of interaction models originally proposed by Knorr-Held (2000). Type I interaction corresponds to a simple adjustment for an overdispersion setting with $\mathbf{R}_\delta = \mathbf{I}_T \otimes \mathbf{I}_n$ and implies the introduction of independent and identically distributed (iid) normally distributed random effects with zero mean and precision τ_δ for each observation. Type II interaction ($\mathbf{R}_\delta = \mathbf{R}_\gamma \otimes \mathbf{I}_n$) assumes structure in time but not in

TABLE 1 Specification of the different types of space-time interactions (Knorr-Held, 2000) and the corresponding sum-to-zero constraints (Goicoa et al., 2018).

Interaction	$\mathbf{R}_\delta$	Constraints
Type I	$\mathbf{I}_T \otimes \mathbf{I}_n$	$\sum_{i=1}^{n} \xi_i = 0, \ \sum_{t=1}^{T} \gamma_t = 0, \ \text{and} \ \sum_{i=1}^{n}\sum_{t=1}^{T} \delta_{it} = 0$
Type II	$\mathbf{R}_\gamma \otimes \mathbf{I}_n$	$\sum_{i=1}^{n} \xi_i = 0, \ \sum_{t=1}^{T} \gamma_t = 0, \ \text{and} \ \sum_{t=1}^{T} \delta_{it} = 0, \text{ for } i = 1, \dots, n$
Type III	$\mathbf{I}_T \otimes \mathbf{R}_\xi$	$\sum_{i=1}^{n} \xi_i = 0, \ \sum_{t=1}^{T} \gamma_t = 0, \ \text{and} \ \sum_{i=1}^{n} \delta_{it} = 0, \text{ for } t = 1, \dots, T$
Type IV	$\mathbf{R}_\gamma \otimes \mathbf{R}_\xi$	$\sum_{i=1}^{n} \xi_i = 0, \ \sum_{t=1}^{T} \gamma_t = 0, \ \text{and}$ $\sum_{t=1}^{T} \delta_{it} = 0, \text{ for } i = 1, \dots, n$ $\sum_{i=1}^{n} \delta_{it} = 0, \text{ for } t = 1, \dots, T$

space, that is, each $\delta_{i\cdot} = (\delta_{i1}, \dots, \delta_{iT})'$ for $i = 1, \dots, n$ follows an independent RW1 prior distribution. Similarly, Type III interaction ($\mathbf{R}_\delta = \mathbf{I}_T \otimes \mathbf{R}_\xi$) assumes structure in space but not in time, that is, each $\delta_{\cdot t} = (\delta_{1t}, \dots, \delta_{nt})'$ for $t = 1, \dots, T$ follows an intrinsic CAR prior distribution with structure matrix $\mathbf{R}_\xi$. Finally, for the Type IV interaction, a completely structured precision matrix $\mathbf{R}_\delta = \mathbf{R}_\gamma \otimes \mathbf{R}_\xi$ is assumed. As for the spatially structured random effect, scaled structure matrices have been considered for the temporal and interaction random effects.

Although these models are flexible enough to describe real situations and their interpretation is fairly straightforward, appropriate sum-to-zero constraints must be imposed on random effects to warrant the identifiability of the intercept, the main spatial and temporal effects, and the space-time interaction effect (Goicoa et al., 2018; Schrödle & Held, 2011). Table 1 shows the constraints chosen for each interaction type in this work. For further details, see Goicoa et al. (2018).

2.2 | Model fitting with the INLA method

In this paper, we use the integrated nested Laplace approximation (INLA) method which provides approximate Bayesian inference in latent Gaussian models (Bakka et al., 2018; Martino & Riebler, 2019; Rue et al., 2009, 2017) due to its faster computational speed compared to simulation techniques based on Markov chain Monte Carlo methods. INLA relies on numerical approximations and integration methods to estimate the posterior marginal distributions of model parameters. This technique can be easily used in the free software `R` through the `R-INLA` package (http://www.r-inla.org/). In what follows, we briefly describe the INLA method.

According to the notation used by Rue et al. (2017), the model class abstraction is obtained using a three-stage hierarchical model formulation, in which observations $\mathbf{y} = (y_1, \dots, y_N)^{\mathrm{T}}$ can be assumed to be conditionally independent, given a latent Gaussian random field $\mathbf{x}$ and hyperparameters θ_1,

$$\mathbf{y} \mid \mathbf{x}, \theta_1 \sim \prod_{i=1}^{N} \pi(y_i \mid x_i, \theta_1).$$

The versatility of the model class relates to the specification of the latent Gaussian field (second stage):

$$\mathbf{x} \mid \theta_2 \sim N(\mathbf{0}, \mathbf{Q}^{-1}(\theta_2)),$$

which includes all random terms in a statistical model, describing the underlying dependence structure of the data. The hyperparameters $\theta = (\theta_1, \theta_2)$ control the latent Gaussian field and/or the likelihood for the data (third stage). The joint posterior distribution of $\mathbf{x}$ and θ given $\mathbf{y}$ is

$$\begin{aligned} \pi(\mathbf{x}, \theta \mid \mathbf{y}) &\propto \pi(\theta) \times \pi(\mathbf{x} \mid \theta) \times \pi(\mathbf{y} \mid \mathbf{x}, \theta) \\ &\propto \pi(\theta) |\mathbf{Q}(\theta)|^{1/2} \exp\left(-\frac{1}{2} \mathbf{x}^T \mathbf{Q}(\theta) \mathbf{x} + \sum_{i=1}^{N} \log\left(\pi(y_i \mid x_i, \theta)\right) \right). \end{aligned}$$

The components of the latent Gaussian field $\mathbf{x}$ are supposed to be conditionally independent with the consequence that $\mathbf{Q}(\theta)$ is a sparse precision matrix (Blangiardo & Cameletti, 2015, Chapter 4, p. 109). Note that if the components x_i and x_j are conditionally independent given all the other components $\mathbf{x}_{-ij}$, that is, if the joint conditional distribution can be factorized as $\pi(x_i, x_j \mid \mathbf{x}_{-ij}) = \pi(x_i \mid \mathbf{x}_{-ij})\pi(x_j \mid \mathbf{x}_{-ij})$, then $\mathbf{Q}_{ij}(\theta) = 0$ and vice versa (Rue & Held, 2005, Chapter 2, Theorem 2.2). This specification is known as *latent Gaussian Markov random field* (GMRF). Therefore, numerical methods for sparse matrices can be used when making inferences with GMRFs, which are much quicker than general algorithms for dense matrices.

The posterior distribution is usually a high-dimensional density that is hard to interpret. However, the interest often lies in the univariate posterior marginals $\pi(\mathbf{x}_i|\mathbf{y})$ and $\pi(\theta_j|\mathbf{y})$ and INLA provides an approximation to such posterior marginal densities. For details, we refer to Rue et al. (2009) and Martino and Riebler (2019).

2.3 | Prediction with INLA

Our main objective is to obtain short-term predictions of mortality rates, which allows us to compute the predictive distribution of mortality cases in nonobserved time point. Suppose that O_{it}^* represents the number of cancer deaths in region i at a future time point. According to Blangiardo and Cameletti (2015, Chapter 5, p. 162) and Gómez-Rubio (2020, Chapter 12, p. 260), its predictive distribution can be derived as

$$
\begin{aligned}
\pi\left(O_{it}^* \mid \mathbf{o}_{-it}\right) &= \frac{\pi\left(O_{it}^*, \mathbf{o}_{-it}\right)}{\pi(\mathbf{o}_{-it})} \\
&= \frac{\displaystyle\int \pi\left(O_{it}^* \mid \theta\right)\pi(\mathbf{o}_{-it} \mid \theta)\pi(\theta)\mathrm{d}\theta}{\pi(\mathbf{o}_{-it})} \\
&= \frac{\displaystyle\int \pi\left(O_{it}^* \mid \theta\right)\pi(\theta \mid \mathbf{o}_{-it})\pi(\mathbf{o}_{-it})\mathrm{d}\theta}{\pi(\mathbf{o}_{-it})} \\
&= \int \pi\left(O_{it}^* \mid \theta\right)\pi(\theta \mid \mathbf{o}_{-it})\mathrm{d}\theta,
\end{aligned}
$$

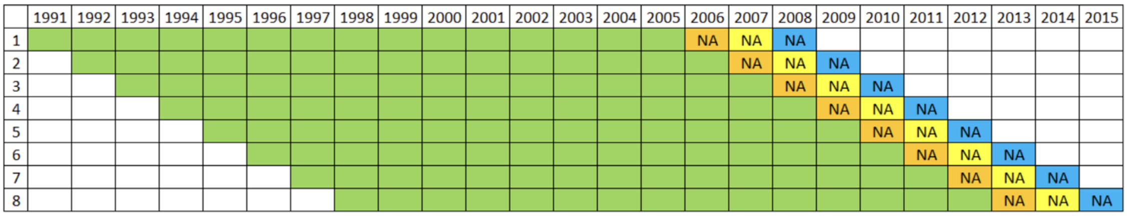
where $\mathbf{o}_{-it}$ is the vector of responses without the observation O_{it}^*. INLA allows for missing values in the response variable and computes posterior marginals for the corresponding linear predictor. If O_{it} is set as NA, this means that O_{it}^* is not observed and hence gives no contribution to the likelihood (see Appendix A.1 for details about how to compute this predictive distribution).

3 | SCALABLE APPROACH FOR HANDLING LARGE SPATIAL DOMAINS

When using Model (1) with Type II or IV interaction effects, a total of $n-1$ and $n+T-1$ sum-to-zero restrictions on interaction effects are required to avoid identifiability problems, respectively (see Table 1). Consequently, if the number of areas n is very large, the model fitting in INLA becomes computationally challenging since inference is affected by the number of constraints added to the random effects. Specifically, INLA uses the kriging technique to correct for constraints (Rue & Held, 2005) with a computational complexity of $\mathcal{O}(nk^2)$ that grows quadratically with the number of constraints k. For a high number of constraints, the cost of this technique dominates the overall cost for approximate inference using sparse matrices.

Recently, Fattah and Rue (2022) have proposed a new implementation for fitting this type of the spatiotemporal model using INLA based on a dense matrix formulation that automatically imposes the necessary set of identifiability constraints. However, this new approach depends on the accessibility to a high-performance computing architecture to speed up inference. In addition, a data set with almost 8000 regions still seems challenging.

Our interest relies on evaluating the scalable Bayesian modeling approach proposed by Orozco-Acosta et al. (2021, 2023) for high-dimensional areal count data and extends it for short-term forecasting in time. Under this approach, two different partition models are defined: the *disjoint model* and *k-order neighborhood models*. In the *disjoint model*, we divide the spatial domain of interest $\mathscr{D}$ into D subdomains, so that $\mathscr{D} = \bigcup_{d=1}^{D} \mathscr{D}_d$, where $\mathscr{D}_i \cap \mathscr{D}_j = \emptyset$ for all $i \neq j$. The partition



	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
1	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟧 NA	🟨 NA	🟦 NA							
2		🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟧 NA	🟨 NA	🟦 NA						
3			🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟧 NA	🟨 NA	🟦 NA					
4				🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟧 NA	🟨 NA	🟦 NA				
5					🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟧 NA	🟨 NA	🟦 NA			
6						🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟧 NA	🟨 NA	🟦 NA		
7							🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟧 NA	🟨 NA	🟦 NA	
8								🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟧 NA	🟨 NA	🟦 NA

FIGURE 1 Validation setup: Each row corresponds to one of the eight validation configurations. Green cells represent years with data that are used in the model. Orange, yellow, and blue cells indicate years for which 1, 2, and 3-year ahead predictions are computed, respectively.

can be chosen based on administrative boundaries, such as states, provinces, autonomous communities, or local health zones, or randomly defined based on a regular grid that is placed over the reference cartography. For our data analyses, partitions are made according to Spanish province boundaries leading to $D = 47$ subdomains. Then, separate Bayesian spatiotemporal models are fitted for each subdomain and the results are just the union of the posterior marginal estimates of the log-rates obtained from each submodel.

To reduce border effects caused by the partition of the spatial domain, Orozco-Acosta et al. (2021) propose to include k-order neighbors for the regions that lie at the boundary of the spatial subdomains. This causes an overlapping of the spatial subdomains, and, consequently, multiple posterior estimates are obtained for regions lying at the subdomain boundaries. Two different merging strategies were compared in Orozco-Acosta et al. (2023) to properly combine the posterior marginal estimates obtained from different submodels: (i) to weight the estimated posterior probability density functions using mixture distributions and (ii) to use the posterior marginal distribution given by the model for which the areal unit of interest originally belongs to. Based on the results obtained from a simulation study, they show that the latter strategy gives better results in terms of risk estimation accuracy and true positive/negative rates. We denote these models as *k-order neighborhood models*. Previous simulation studies on spatial (Orozco-Acosta et al., 2021) and spatiotemporal (Orozco-Acosta et al., 2023) areal data have shown that using first-order neighborhood models ($k = 1$) is often deemed suitable.

One of the main advantages of this scalable approach is that submodels can be simultaneously fitted using both parallel or distributed computation strategies. The `R` package `bigDM` (Adin et al., 2023) implements this methodology and allows it to fit several scalable univariate and multivariate disease mapping models for high-dimensional data in a fully Bayesian setting using INLA.

4 | PREDICTIVE VALIDATION STUDY

To evaluate the predictive ability of all the models, predictions of lung cancer mortality counts in all 7907 continental municipalities of Spain have been made for three consecutive periods (years). We note that mortality registries often provide data with a delay of up to 3 years. A total of $K = 8$ configurations have been considered, where each configuration uses $T = 15$ years of data to fit the model and predict at time points $T + 1$, $T + 2$, and $T + 3$. The first configuration uses data from 1991 to 2005, the second configuration data from 1992 to 2006, and so on. This results in predictions for the years 2006–2015, whereby years 2006 and 2015 are only predicted in one configuration, years 2007 and 2014 in two configurations, and all the other years in three configurations. Figure 1 illustrates the validation setup, which is similar to the one used by Ghosh and Tiwari (2007) and Etxeberria et al. (2014).

4.1 | Assessment criteria

To assess the predictive performance of the models, we compute the mean absolute error (MAE) and the root mean square error (RMSE) of predicted mortality counts for each municipality $i = 1, \dots, 7907$ differing between $k = 1, 2, 3$ year ahead predictions as

$$\text{MAE}_i^{(k)} = \frac{1}{8}\sum_{t=2005+k}^{2012+k} \left|O_{it} - \widehat{O}_{it}\right|, \qquad \text{and} \qquad \text{RMSE}_i^{(k)} = \sqrt{\frac{1}{8}\sum_{t=2005+k}^{2012+k} \left(O_{it} - \widehat{O}_{it}\right)^2},$$

where O_{it} is the number of observed cases and $\widehat{O}_{it}$ is the expected value for the posterior predictive counts for areal unit i and time period t, respectively (see Appendix A.1 for details about its computation with INLA). Looking at Figure 1 this means that average scores over the orange, yellow, and blue cells are built.

To assess not only point predictions but the entire predictive distribution, we compute the 95% interval score (IS), which is a proper scoring rule that combines calibration and sharpness of predictions (Gneiting & Raftery, 2007). This measure transforms interval width and empirical coverage into a single score and has recently become popular (see, for instance, Hofer & Held, 2022; Paige et al., 2022). Let O_{it} be the number of cases and $[l, u]$ be the respective $(1-\alpha) \cdot 100\%$ posterior predictive credible interval at credible level $\alpha \in (0,1)$, then

$$IS_{\alpha}\left(O_{it}\right) = (u - l) + \frac{2}{\alpha}\left(l - O_{it}\right) I\left[O_{it} < l\right] + \frac{2}{\alpha}\left(O_{it} - u\right) I\left[O_{it} > u\right].$$

Here, $I[\cdot]$ denotes an indicator function that penalizes the length of the credible interval if the number of observed cases is not contained within that interval.

4.2 | Implementation details

To fit the classical disease mapping models described in Model (1), calculations are made on a computer with Intel Xeon E5-2620 v4 processors and 256 GB RAM (CentOS Linux release 7.3.1611 operative system), using the simplified Laplace approximation strategy in `R-INLA` (stable version INLA 22.05.07) of `R-4.2.0`. Partition models, disjoint, and first-order neighborhood models based on 47 provinces, are distributed over five machines (with the same specifications described above) simultaneously running eight models in parallel on each machine using the `bigDM` package.

4.3 | Hyperprior distributions

In a Bayesian framework, prior distributions need to be assigned to all parameters. Here, we use a uniform improper prior on the positive real line for all standard deviations (square root inverse of precision parameters) in the model (Gelman, 2006; Gómez-Rubio, 2020, Chapter 5.3) and a Uniform [0,1] distribution for the mixing parameter ϕ of the BYM2 prior. We also use Penalized Complexity (PC) priors (Simpson et al., 2017) for the same parameters using the default PC prior values given in R-INLA, that is, $P(\sigma > 1) = 0.01$ and $P(\phi > 0.5) = 0.5$ and the results obtained remain very similar. Finally, a vague zero mean normal distribution with a precision close to zero (0.001) is assigned to the intercept β_0. Additional details on the implementation of these hyperprior distributions in INLA can be found in Ugarte et al. (2016).

4.4 | Results

Table 2 compares average values of 95% IS, MAE, and RMSE over all the municipalities for 1, 2, and 3-year ahead predictions when using the different models described in the previous section. For all these criteria, lower values are preferred. As expected, higher scores are obtained as we increase the time of prediction in future years. Note that we were not able to fit with INLA the classical disease mapping models using Types II and IV interactions due to the high-number of identifiability constraints. However, we were able to fit the partition models proposed by Orozco-Acosta et al. (2023) with all types of space-time interactions. In addition, a substantial reduction in computational time for fitting Types I and III models was obtained. We observe that partition models outperform the classical models when these latter models can be fitted (only Types I and III interactions). The best predictive measures are obtained using the Type IV interaction model. Compared to a disjoint model that estimates all 47 provinces separately, we find that the first-order neighborhood models obtain slightly better IS values when predicting 2- and 3-year ahead. The running time is slightly more than twice as long.

Figure 2 investigates predictions in more detail for three selected provincial capitals: the municipalities of Madrid, Palencia, and Ávila. Here, expected values for the posterior predictive counts (deaths per 100,000 inhabitants) together with 95% predictive intervals are plotted using the disjoint and first-order neighborhood models with Type IV interaction. Different colors are used for 1, 2, and 3-year ahead predictions. Very similar forecasts are obtained for the municipality of Madrid (a region with a large number of population at risk) under the different models. As expected, wider predictive

TABLE 2 Validation study: average values of evaluation scores (interval score [IS], mean absolute error [MAE], and root mean square error [RMSE]) and computational time (in min) for classical, disjoint, and first-order neighborhood models fitted using the simplified Laplace approximation strategy of INLA.

		1-year ahead			2-year ahead			3-year ahead			
Model	**Space-time interaction**	**$\text{IS}_{0.05}$**	**MAE**	**RMSE**	**$\text{IS}_{0.05}$**	**MAE**	**RMSE**	**$\text{IS}_{0.05}$**	**MAE**	**RMSE**	**Time**
Classical	Type I	3.91	0.68	0.84	3.97	0.69	0.85	3.97	0.70	0.85	343
	Type II	–	–	–	–	–	–	–	–	–	–
	Type III	3.94	0.68	0.84	3.96	0.69	0.85	4.03	0.70	0.86	1767
	Type IV	–	–	–	–	–	–	–	–	–	–
Disjoint	Type I	3.91	0.68	0.83	3.95	0.69	0.84	3.97	0.70	0.85	14
by provinces	Type II	3.87	0.67	0.82	3.92	0.68	0.83	3.98	0.69	0.84	152
	Type III	3.88	0.68	0.83	3.92	0.69	0.84	3.95	0.69	0.85	22
	Type IV	3.84	0.67	0.82	3.90	0.68	0.83	3.96	0.69	0.84	179
First-order	Type I	3.90	0.68	0.83	3.94	0.69	0.84	3.96	0.69	0.85	18
neighborhood	Type II	3.86	0.67	0.82	3.91	0.68	0.83	3.97	0.69	0.84	405
by provinces	Type III	3.88	0.68	0.83	3.91	0.69	0.84	3.93	0.69	0.85	40
	Type IV	3.84	0.67	0.82	3.88	0.68	0.83	3.93	0.69	0.84	433

intervals are obtained for those areas with lower values of observed cases and population at risk, as is the case of the municipalities of Palencia and Ávila. For these municipalities, slightly wider predictive intervals are observed when using the disjoint model.

To perform a more in-depth analysis of the predictive performance of the partition models, we compute average values of prediction evaluation scores for two subsets of the data: (i) municipalities for which the proportion of zero observed cases during the study period is less or equal to 0.2 and (ii) municipalities lying at the boundary between two or more provinces with at least two observed cases per 100,000 inhabitants during the whole study period. The results are shown in Tables A1 and A2 (in the Appendix), respectively. In both scenarios, the first-order neighborhood model outperforms the disjoint model in terms of prediction accuracy and interval score.

5 | ILLUSTRATION: PROJECTIONS OF CANCER MORTALITY IN SPAIN

The aim of this section is to illustrate our proposal for forecasting cancer burden for up to 3 years, as cancer registries often suffer from this delay in data provision. Here, we use both male lung cancer and overall cancer (all sites) mortality data in the 7907 municipalities of continental Spain in the period 1991–2012 to forecast cancer data for the years 2013, 2014, and 2015. This allows us to check whether actual rates are close to predicted rates. The same models described in the validation study of Section 4 have been considered here.

To compare the predictive performance of the different models, we use cross-validation techniques to compute scoring rules based on the estimated predictive distribution of the mortality counts. Typically, cross-validation techniques are based on the idea of splitting the observed data into a training set (sample data used for model's parameter estimation) and a testing set (set of points used to compute the prediction error based on the training model) multiple times to estimate the predictive accuracy of the model (Gelman et al., 1995; Hastie et al., 2009). A particular interesting feature of INLA is that it provides leave-one-out cross-validatory (LOOCV) model checks without rerunning the model for each observation in turn (Held et al., 2010; Rue et al., 2009), something that would be computationally unfeasible when fitting complex models on large data sets. Specifically, INLA provides approximations of the conditional predictive ordinates (CPO; Pettit, 1990), $\text{CPO}_{it} = \pi(O_{it} = o_{it}|\mathbf{o}_{-it})$, which is defined as the cross-validates predictive probability mass at the observed count o_{it}. However, it is well-known that LOOCV techniques may not be appropriate to measure the predictive performance of a model that includes spatially and/or temporally structured random effects to deal with correlated data (Rabinowicz & Rosset, 2022; Roberts et al., 2017). To solve this problem, Liu and Rue (2022) propose an automatic group construction procedure for leave-group-out cross-validation (LGOCV) to estimate the predictive performance of structured models for latent Gaussian models with INLA.

Table 3 shows the sum of the log-predictive densities computed over each area-time point using both the LOOCV (usually named as logarithmic score; Gneiting & Raftery, 2007) and the LGOCV techniques, as well as model selection criteria

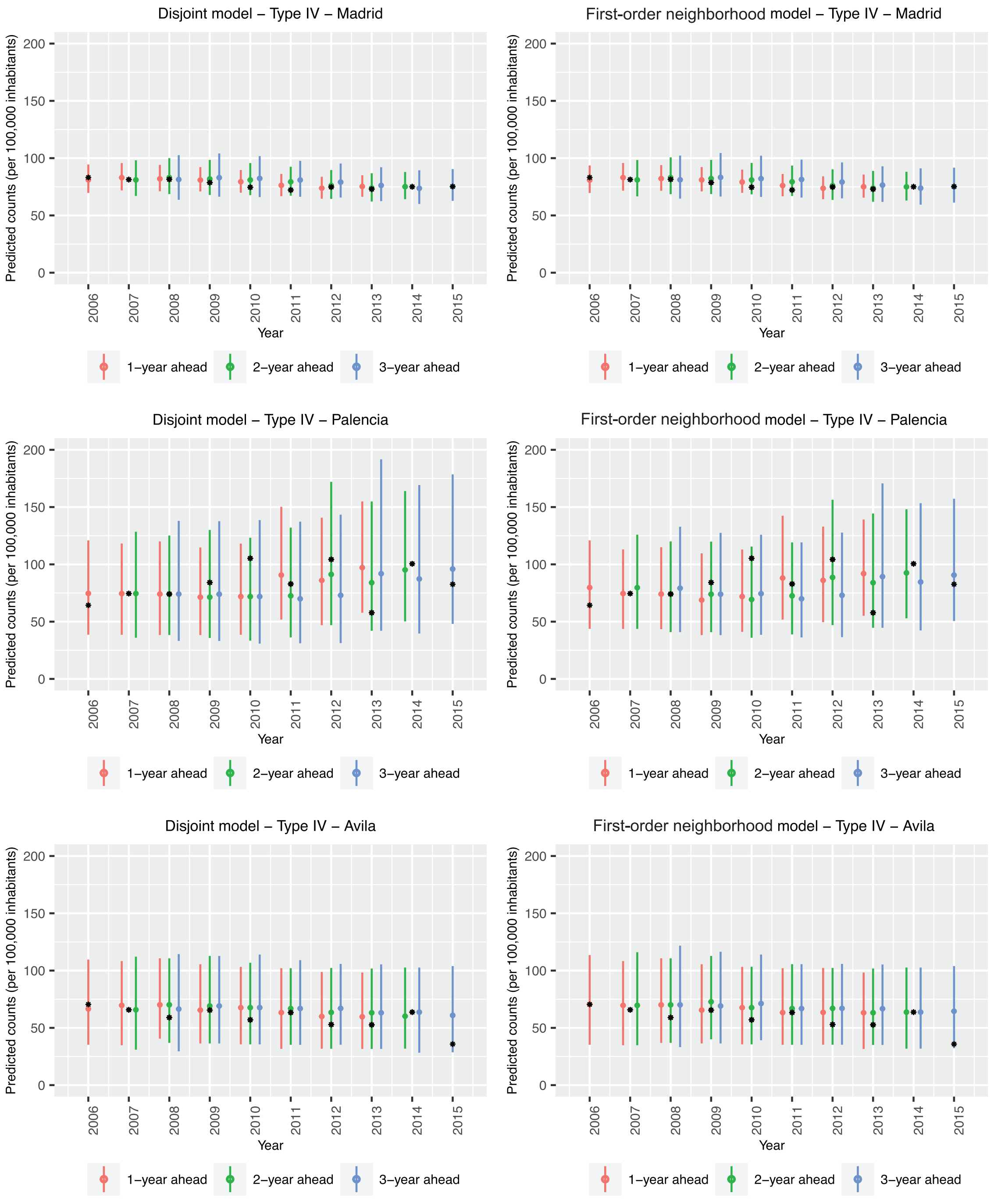


FIGURE 2 One, 2, and 3-year ahead predictions for the municipalities of Madrid (top), Palencia (middle), and Ávila (bottom) using the disjoint model (left column) and first-order neighborhood model (right column) with Type IV interactions. Expected values of posterior predictive counts per 100,000 inhabitants (dots) and 95% predictive intervals (color lines) are plotted. The number of the real observed number of cases is also included as black stars.

such as the deviance information criterion (DIC; Spiegelhalter et al., 2002) and the Watanabe–Akaike information criterion (WAIC; Watanabe, 2010). For comparison purposes, the reference value has been set to zero by subtracting the minimum value for each column when computing the cross-validation measures and DIC/WAIC values. As in the validation study, we were not able to fit the classical spatiotemporal models with Types II and IV interactions. For both lung and overall cancer mortality data analyses, partition models show better predictive performance and better values for model selection criteria (see Table 3). In particular, the first-order neighborhood models with Type IV interactions. As expected, the differences are more pronounced when comparing the predictive performance of the models for overall cancer data, as a higher number of deaths are observed for each areal time unit compared to lung cancer data.

TABLE 3 Logarithmic score using both LOOCV and LGOCV techniques, model selection criteria, and computational time (in min) for lung cancer mortality and overall cancer mortality data with models fitted using the simplified Laplace approximation strategy of INLA (stable version INLA_22.12.16).

		Lung cancer					**Overall cancer**				
		LOOCV[a]	**LGOCV**[a]	**DIC**[a]	**WAIC**[a]	**Time**	**LOOCV**[a]	**LGOCV**[a]	**DIC**[a]	**WAIC**[a]	**Time**
Classical	Type I	440	1925	698	703	622	1542	5763	2591	2662	598
	Type II	–	–	–	–	–	–	–	–	–	–
	Type III	365	2048	557	596	3914	1110	5072	1780	1905	4674
	Type IV	–	–	–	–	–	–	–	–	–	–
Disjoint	Type I	244	363	437	373	20	925	1743	1516	1442	20
by provinces	Type II	52	78	150	130	296	46	101	123	67	275
	Type III	209	330	392	369	36	836	1533	1385	1411	36
	Type IV	0	0	48	47	379	18	8	76	74	323
First-order	Type I	234	373	376	333	28	912	2078	1451	1387	28
neighborhood	Type II	47	73	93	78	661	57	100	114	61	633
by provinces	Type III	209	376	351	351	55	799	1716	1301	1342	55
	Type IV	0	18	0	0	800	0	0	0	0	754

Abbreviations: DIC, deviance information criterion; LGOCV, leave-group-out cross-validation; LOOCV, leave-one-out cross-validation; WAIC, Watanabe–Akaike information criterion.

aReference value has been set to zero by subtracting the minimum value for each column.

5.1 | Lung cancer mortality

In this section, we provide lung cancer mortality projections in the municipalities of Spain for the period 2013–2015 using the first-order neighborhood model with Type IV space-time interactions. Table 4 shows posterior median estimates of the predicted mortality rates per 100,000 males, and its corresponding 95% credible intervals for years 2013 and 2015 for the 47 municipalities that are provincial capitals. As expected, wider credible intervals are obtained when computing 3-year ahead predictions (year 2015).

Figure 3a shows the maps with the temporal evolution of posterior median estimates of lung cancer mortality rates for some selected years between 1991 and 2015. An increasing trend is observed in the northwest and central-west regions of Spain, in particular during the period 1991–2001. The average mortality rate for Spain is about 71.7 deaths per 100,000 males in the year 1991, rising to 78.3 and 83.2 deaths per 100,000 males in the years 1996 and 2001, respectively. For the second half of the period, the estimated rates are fairly constant without major spatial changes (average mortality rates close to 83.5 deaths per 100,000 males during the years 2005–2010, with a slight increase to 84.5 deaths per 100,000 males in 2012), something that is also observed for the predicted years 2013–2015.

In Figure 4a, we show the temporal evolution of mortality rate forecasts for the provincial capitals of Girona, Madrid, and Bilbao (selected to show areas with different estimated temporal trends). In general, the 95% credible intervals contain the crude rates over all the study period. As expected, wider credible intervals are obtained for those areas with lower populations at risk.

5.2 | Overall cancer mortality

Similar to the previous section, here we provide overall cancer mortality projections in the municipalities of Spain for the period 2013–2015 using the first-order neighborhood model with Type IV space-time interactions. Table 5 provides posterior median estimates of predicted mortality rates per 100,000 males and its corresponding 95% credible intervals for the provincial capitals. The municipalities of Santander, Bilbao, and Salamanca have the highest overall cancer mortality rates, while Murcia, Toledo, and Guadalajara show the lowest values. In general, the temporal trend in forecasts is quite stable with slight variation in some areas, as is the case of Soria with an increase of five cases per 100,000 males.

The maps with the temporal evolution of the overall cancer mortality rates for the 7907 municipalities of continental Spain are shown in Figure 3b. We observed a remarkable increase in the estimated cancer rates in the regions located in the west (Extremadura) and northwest (Castilla y León, Galicia, Asturias, and part of Cantabria) during the period

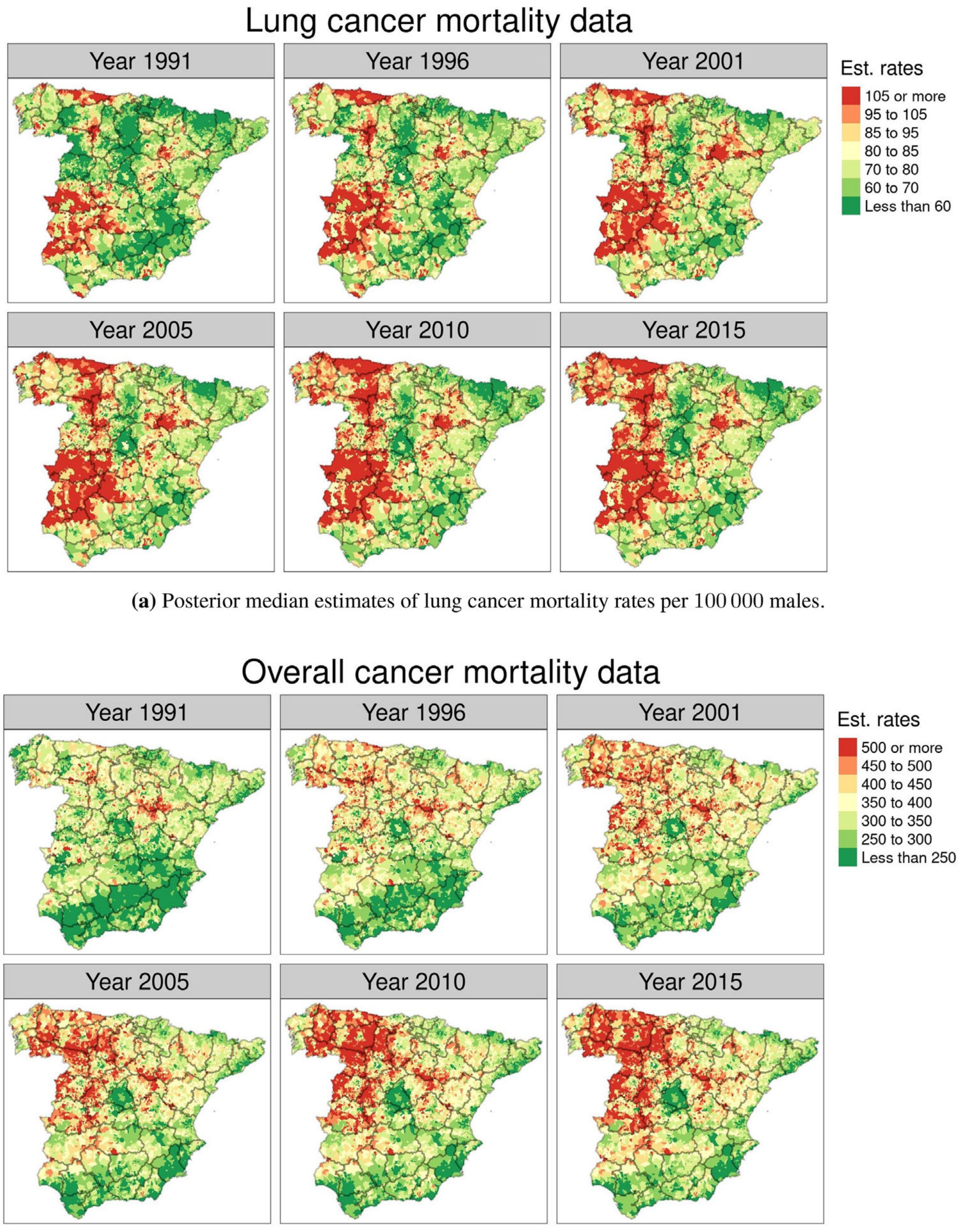
TABLE 4 Posterior median estimates of predicted lung cancer mortality rates ($\hat{\lambda}_{it}$) per 100,000 males, its corresponding 95% credible intervals (CI), and width of the CIs for years 2013 and 2015 for the 47 municipalities that form the provincial capitals (sorted by increasing order).

	2013			2015		
Municipality	$\hat{\lambda}_{it^\star} \times 10^5$	**95% CI**	**Width**	$\hat{\lambda}_{it^\star} \times 10^5$	**95% CI**	**Width**
Guadalajara	56.1	(33.6,85.2)	51.6	57.2	(34.3,87.0)	52.7
Jaén	58.8	(37.4,82.0)	59.4	59.4	(36.0,84.6)	48.6
Girona	59.9	(38.5,85.6)	47.1	59.9	(36.4,89.8)	53.4
Albacete	60.4	(42.9,80.1)	37.2	59.6	(40.9,81.7)	40.8
Segovia	61.6	(30.8,100.1)	69.3	64.0	(32.0,108.0)	76.0
Ávila	63.2	(35.1,101.8)	66.7	64.5	(32.3,107.5)	75.2
Murcia	63.3	(51.2,75.8)	24.6	63.2	(50.2,78.1)	27.9
Soria	63.6	(26.5,106.0)	79.5	59.3	(21.6,113.2)	91.6
Granada	64.2	(47.0,85.0)	38.0	64.0	(44.8,87.8)	43.0
Tarragona	65.9	(46.9,89.3)	42.4	65.9	(44.9,91.3)	46.4
Burgos	66.5	(46.6,87.5)	40.9	66.3	(45.0,90.0)	45.0
Toledo	67.5	(42.5,100.0)	57.5	67.7	(40.1,102.9)	62.8
Castellón	67.9	(48.6,89.4)	40.8	67.1	(48.0,91.1)	43.1
Vitoria	69.2	(52.3,88.6)	36.3	68.7	(50.3,92.2)	41.9
Almería	69.3	(49.1,91.8)	42.7	68.8	(47.6,95.3)	47.7
Logroño	69.5	(49.1,95.4)	46.3	69.4	(47.2,98.5)	51.3
Lérida	72.2	(50.5,98.1)	47.6	71.6	(48.2,102.3)	54.1
Córdoba	74.7	(60.1,91.2)	31.1	75.0	(59.8,92.2)	32.4
Nadrid	74.8	(65.4,85.6)	20.2	75.0	(61.3,91.6)	30.3
Teruel	74.9	(40.3,126.7)	86.4	76.2	(35.2,129.0)	93.8
Málaga	75.0	(62.5,88.1)	25.6	74.9	(61.0,91.3)	30.3
Cuenca	77.9	(40.8,118.7)	77.9	75.5	(41.5,120.7)	79.2
Sevilla	78.4	(67.3,90.8)	23.5	78.6	(65.5,93.1)	27.6
Alicante	78.6	(65.1,94.6)	29.5	79.0	(64.0,94.7)	30.7
Pontevedra	78.8	(50.8,111.8)	61.0	76.8	(46.1,115.2)	69.1
Ciudad Real	79.0	(48.0,115.6)	67.6	79.5	(48.3,116.5)	68.2
Huesca	83.2	(47.6,126.8)	79.2	84.1	(44.0,132.1)	88.1
Valladolid	84.0	(66.2,105.2)	39.0	84.5	(63.5,109.6)	46.1
Huelva	84.0	(60.2,110.6)	50.4	83.9	(59.8,115.2)	55.4
Badajoz	84.3	(62.3,109.0)	46.7	83.7	(61.4,112.4)	51.0
Pamplona	85.1	(64.9,107.5)	42.6	84.8	(64.4,109.5)	45.1
San Sebastián	86.8	(64.0,113.1)	49.1	86.8	(59.4,117.7)	58.3
Valencia	87.4	(75.3,100.9)	25.6	87.4	(73.0,103.4)	30.4
Barcelona	88.0	(77.8,99.0)	21.2	88.3	(74.9,103.2)	28.3
Zamoraona	88.4	(52.4,130.9)	78.5	86.9	(46.8,140.4)	93.6
Cáceres	88.9	(58.5,121.4)	62.9	89.2	(58.8,126.2)	67.4
Palencia	89.3	(55.1,133.9)	78.8	89.3	(50.6,146.6)	96.0
Ourense	91.2	(60.8,127.6)	66.8	90.5	(57.6,135.8)	78.2
Zaragoza	91.2	(76.1,107.8)	31.7	91.2	(71.9,114.3)	42.4
Lugo	93.3	(62.9,128.0)	65.1	94.2	(61.4,133.7)	72.3
Bilbao	95.8	(75.8,117.6)	41.8	95.3	(72.0,123.7)	51.7
Santander	97.6	(72.0,128.1)	56.1	98.4	(69.7,135.7)	66.0
Salamanca	99.0	(69.9,135.4)	65.5	98.7	(62.8,145.1)	82.3

(Continues)

TABLE 4 (Continued)

	2013			2015		
Municipality	$\hat{\lambda}_{it^*} \times 10^5$	**95% CI**	**Width**	$\hat{\lambda}_{it^*} \times 10^5$	**95% CI**	**Width**
A Coruña	99.8	(76.1,125.2)	49.1	98.9	(72.4,133.4)	61.0
León	102.2	(72.0,139.0)	67.0	103.3	(68.8,144.6)	75.8
Oviedo	104.9	(81.0,131.6)	50.6	104.8	(76.7,138.8)	62.1
Cádiz	115.1	(80.7,156.3)	75.6	112.3	(73.7,168.4)	94.7



(a) Posterior median estimates of lung cancer mortality rates per 100 000 males.

(b) Posterior median estimates of overall cancer mortality rates per 100 000 males.

FIGURE 3 Posterior median estimates of mortality rates per 100,000 males for the 7907 municipalities of continental during the period 1991–2015. Years 2013–2015 were predicted.

TABLE 5 Posterior median estimates of predicted overall cancer mortality rates ($\hat{\lambda}_{it}$) per 100,000 males, its corresponding 95% credible intervals (CI) and width of the CIs for years 2013 and 2015 for the 47 municipalities that form the provincial capitals (sorted by increasing order).

	2013			2015		
Municipality	$\hat{\lambda}_{it^\star} \times 10^5$	**95% CI**	**Width**	$\hat{\lambda}_{it^\star} \times 10^5$	**95% CI**	**Width**
Murcia	229.4	(203.3,255.9)	52.6	229.6	(199.8,262.1)	62.3
Toledo	239.9	(184.9,304.9)	120.0	243.4	(183.1,316.1)	133.0
Guadalajara	239.9	(186.1,298.2)	112.1	240.3	(180.8,309.0)	128.2
Albacete	242.6	(204.3,284.4)	80.1	242.9	(198.5,293.1)	94.6
Girona	243.9	(194.7,297.3)	102.6	243.8	(192.5,305.8)	113.3
Castellón	248.9	(204.7,297.5)	92.8	249.4	(196.6,310.5)	113.9
Jaén	253.1	(203.2,310.1)	106.9	253.8	(192.6,325.8)	133.2
Almería	253.9	(215.5,296.6)	81.1	254.0	(206.4,307.0)	100.6
Tarragona	256.2	(212.3,303.1)	90.8	255.9	(208.0,308.3)	100.3
Málaga	266.2	(237.7,297.3)	59.6	267.0	(229.4,310.1)	80.7
Lérida	277.1	(228.0,330.5)	102.5	277.7	(220.7,343.5)	122.8
Córdoba	277.3	(245.0,314.0)	69.0	277.8	(238.4,323.6)	85.2
Alicante	280.3	(243.7,321.1)	77.4	281.1	(232.1,335.6)	103.5
Granada	281.1	(237.7,328.1)	90.4	281.6	(229.5,344.7)	115.2
Ciudad Real	284.9	(220.0,352.6)	132.6	284.1	(215.9,363.7)	147.8
Badajoz	288.0	(241.3,336.0)	94.7	286.3	(234.0,346.4)	112.4
Cuenca	289.4	(218.9,371.1)	152.2	290.5	(211.3,388.6)	177.3
Sevilla	291.8	(264.8,319.5)	54.7	291.0	(257.0,329.5)	72.5
Cáceres	294.9	(236.4,355.6)	119.2	293.8	(230.7,367.9)	137.2
Ávila	294.9	(221.2,379.2)	158.0	293.9	(211.5,397.8)	186.3
Madrid	296.1	(264.8,329.9)	65.1	296.3	(247.1,352.9)	105.8
Segovia	300.4	(223.4,392.8)	169.4	304.0	(208.0,411.9)	203.9
Huelva	302.5	(247.9,362.7)	114.8	303.1	(237.6,379.9)	142.3
Logroño	306.6	(256.2,365.2)	109.0	306.6	(247.0,381.5)	134.5
Vitoria	311.5	(269.3,356.2)	86.9	310.8	(261.4,367.8)	106.4
Valencia	316.9	(288.1,346.4)	58.3	317.6	(280.8,359.5)	78.7
Pamplona	323.4	(279.8,370.2)	90.4	323.2	(273.8,375.8)	102.0
Zaragoza	323.4	(289.3,358.5)	69.2	324.1	(278.3,374.5)	96.2
Soria	328.6	(249.1,424.0)	174.9	328.7	(242.5,431.1)	188.6
Pontevedra	330.4	(269.4,396.5)	127.1	330.1	(261.0,404.4)	143.4
Barcelona	335.9	(304.4,368.9)	64.5	336.0	(288.2,385.8)	97.6
Teruel	357.2	(259.2,472.4)	213.2	357.7	(252.1,492.6)	240.5
San Sebastián	368.3	(319.8,421.5)	101.7	367.9	(308.5,435.3)	126.8
A Coruña	368.4	(313.3,432.3)	119.0	370.2	(295.1,461.2)	166.1
Burgos	368.5	(314.9,425.7)	110.8	369.3	(306.6,442.7)	136.1
Valladolid	370.8	(320.9,426.1)	105.2	370.8	(306.6,451.1)	144.5
Huesca	372.6	(285.4,471.7)	186.3	372.3	(276.2,488.3)	212.1
Zamora	376.4	(288.1,477.9)	189.8	377.8	(270.8,511.6)	240.8
Lugo	384.0	(314.6,459.9)	145.3	385.7	(302.4,477.8)	175.4
Palencia	396.5	(315.1,483.1)	168.0	394.5	(303.9,498.5)	194.6
Cádiz	398.4	(309.1,499.7)	190.6	396.5	(280.7,543.9)	263.2
Oviedo	403.2	(353.7,457.6)	103.9	403.6	(342.5,471.6)	129.1
León	405.3	(329.9,492.4)	162.5	406.2	(315.0,523.2)	208.2

(Continues)

TABLE 5 (Continued)

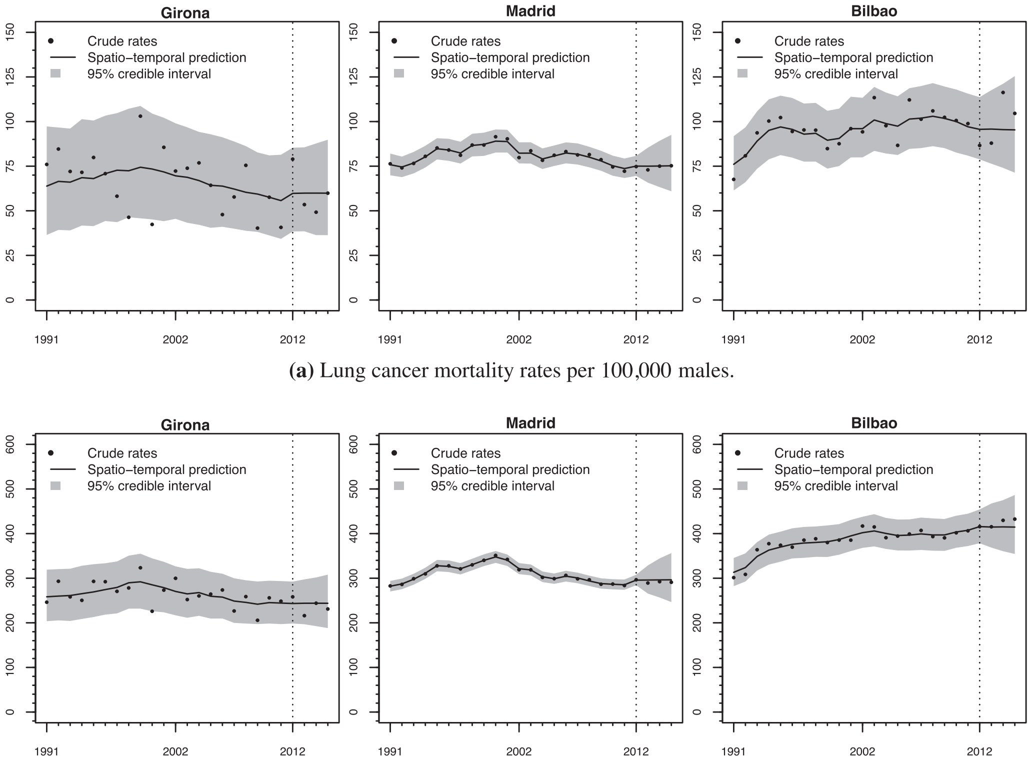
	2013			2015		
Municipality	$\hat{\lambda}_{it^\star} \times 10^5$	95% CI	Width	$\hat{\lambda}_{it^\star} \times 10^5$	95% CI	Width
Ourense	407.2	(336.3,484.2)	147.9	407.5	(325.1,506.3)	181.2
Salamanca	413.5	(345.0,490.6)	145.6	412.8	(327.5,520.5)	193.0
Bilbao	415.3	(368.0,466.8)	98.8	414.6	(356.2,484.1)	127.9
Santander	419.8	(339.3,517.4)	178.1	420.8	(303.8,574.0)	270.2

1991–2015. The overall mortality rate for the whole of Spain shows a steady rise during the study period, with average values of 325.5, 355.5, 373.4, 381.6, and 386.3 deaths per 100,000 males in the years 1991, 1996, 2001, 2010, and 2015, respectively.

In Figure 4b, we show the temporal evolution of estimated overall cancer mortality rates for the municipalities of Girona, Madrid, and Bilbao. Wider credibility intervals are observed in Bilbao and Girona compared to Madrid, as expected due to the number of inhabitants. Crude rates are always included in the credible intervals.

6 | DISCUSSION

Short-term disease forecasting is of great interest in epidemiology and public health as it supports health decision-making processes. However, this might be a very complex task when predicting counts for high-dimensional areal data. As far as we know, our paper is the first attempt to extend classical spatiotemporal disease mapping models for predicting short-term cancer burden when the number of areas is very large. Under this high-resolution spatial setting, the main limitation of classical models is their computational complexity due to the huge dimension of the spatiotemporal covariance matrices and the high number of constraints to make the models identifiable. The "divide-and-conquer" approach for high-dimensional count data proposed by Orozco-Acosta et al. (2023) is a strategy that involves partitioning the spatial



(a) Lung cancer mortality rates per 100,000 males.

(b) Overall cancer mortality rates per 100,000 males.

FIGURE 4 Posterior predictive median estimates of mortality rates and its corresponding 95% credible interval per 100,000 males during the period 1991-2015 for the municipalities of Girona, Madrid, and Bilbao. Crude rates are shown as a filled circle. The vertical dotted line indicates where the prediction starts.

domain into smaller regions and then fitting a separate model for each partition. This approach has been shown to be effective for estimating disease risks as it can provide better accuracy and computational efficiency compared to classical spatiotemporal disease mapping models. In this paper, we extend this approach to a forecasting setting, where the goal is to forecast short-term mortality rates.

To evaluate the predictive performance of both classical and scalable modeling approaches, we perform a validation study imitating the real scenario of forecasting short-term cancer mortality rates in almost 8000 municipalities in Spain. For the partition models, we use the 47 provinces of Spain (NUTS3 level from the European nomenclature of territorial units for statistics) although other partitions can be used depending on the researcher's interest. Then, we compute mean absolute errors, root mean square errors, and interval scores when making 1, 2, and 3-year ahead predictions. In general, partition models outperform the classical models in terms of predicted counts, being the Type IV interaction model the one showing better results. When computing these measures by stratifying the data based on the proportion of areas with zero observed cases during the study period or selecting only municipalities lying at the boundary of two or more provinces, we observe that the first-order neighborhood model outperforms the disjoint model in terms of prediction accuracy and interval score. Of note, including second-order neighborhood models in our validation study does not improve the results further (results are available upon request).

We also illustrate our proposed methodology by forecasting 3 years ahead of lung cancer and overall cancer mortality data in the municipalities of continental Spain using data from the period 1991–2012. To compare the different models in terms of their predictive performance, we compute the logarithmic score based on both LOOCV and LGOCV techniques. For the latter, we use the automatic group construction for latent Gaussian models proposed by Liu and Rue (2022), which computes the set of testing points for each data point based on posterior correlations of the linear predictor. Under this cross-validation setting, two alternatives exist for comparing the predictive performance of the models: (i) calculating the set of testing points for each model or (ii) utilizing the groups derived from a specific model as a reference for the remaining models. To simplify the practical calculation of these measures when analyzing real data, we adopt the first strategy. According to the results, a first-order neighborhood model with the Type IV interaction model is chosen as the best model in terms of predictive performance. The conclusions align with those obtained when computing the logarithmic score under the second strategy. The differences between the models are more pronounced when analyzing the data for overall cancer mortality because the number of observed cases is higher. Well-known model selection criteria such as the DIC and WAIC provide similar conclusions.

In summary, the results of this paper show that the "divide-and-conquer" approach performs well in terms of accuracy and computational time and outperforms classical methods in all the scenarios. These findings suggest that this approach is a promising strategy for forecasting high-dimensional count data and can provide valuable insights for decision-making in various fields, such as public health and environmental monitoring. We remark that the scalable methodology presented in this paper could also be used to forecast rates using other hierarchical Bayesian disease mapping models including, for example, AR terms or second-order random walks for time. Moreover, the fact that the proposed methodology is general and can be applied to other health indicators such as cancer incidence or other health indicators is also valuable. This suggests that the methodology can be used for a broader range of applications beyond the scope of the paper.

ACKNOWLEDGMENTS

We would like to thank the Spanish Statistical Office (INE) for providing the data. This research has been supported by the project PID2020-113125RB- I00/MCIN/AEI/10.13039/501100011033 (principal investigator: M.Dolores Ugarte).

Open access funding provided by Universidad Pública de Navarra.

CONFLICT OF INTEREST STATEMENT

The authors declare no potential conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from Spanish Statistical Office. Restrictions apply to the availability of these data, which were used under license for this study. Data are available from the author(s) with the permission of Spanish Statistical Office.

OPEN RESEARCH BADGES

This article has earned an Open Data badge for making publicly available the digitally-shareable data necessary to reproduce the reported results. The data is available in the Supporting Information section.

This article has earned an open data badge "**Reproducible Research**" for making publicly available the code necessary to reproduce the reported results. The results reported in this article were reproduced partially due to confidentiality issues.

ORCID

Maria D. Ugarte https://orcid.org/0000-0002-3505-8400

REFERENCES

Adin, A., Orozco-Acosta, E., & Ugarte, M. D. (2023). *bigDM: Scalable Bayesian disease mapping models for high-dimensional data*. R package version 0.5.1.

Assuncao, R. M., Reis, I. A., & Oliveira, C. D. L. (2001). Diffusion and prediction of Leishmaniasis in a large metropolitan area in Brazil with a Bayesian space–time model. *Statistics in Medicine*, *20*(15), 2319–2335.

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

Bernardinelli, L., Clayton, D., Pascutto, C., Montomoli, C., Ghislandi, M., & Songini, M. (1995). Bayesian analysis of space-time variation in disease risk. *Statistics in Medicine*, *14*(21–22), 2433–2443.

Blangiardo, M., & Cameletti, M. (2015). *Spatial and spatio-temporal Bayesian models with R-INLA*. John Wiley & Sons.

Corpas-Burgos, F., & Martinez-Beneito, M. A. (2021). An autoregressive disease mapping model for spatio-temporal forecasting. *Mathematics*, *9*(4), 384.

Diaz-Rubio, E. (2019). La carga del cáncer en España: situación en 2019. *Real Academia Nacional de Medicina de España*, *136*(1), 25–33.

Etxeberria, J., Goicoa, T., López-Abente, G., Riebler, A., & Ugarte, M. D. (2017). Spatial gender-age-period-cohort analysis of pancreatic cancer mortality in Spain (1990–2013). *PLoS One*, *12*(2), e0169751.

Etxeberria, J., Goicoa, T., & Ugarte, M. D. (2023). Using mortality to predict incidence for rare and lethal cancers in very small areas. *Biometrical Journal*, *65*(3), 2200017.

Etxeberria, J., Goicoa, T., Ugarte, M. D., & Militino, A. F. (2014). Evaluating space-time models for short-term cancer mortality risk predictions in small areas. *Biometrical Journal*, *56*(3), 383–402.

Etxeberria, J., Ugarte, M. D., Goicoa, T., & Militino, A. F. (2015). On predicting cancer mortality using ANOVA-type P-spline models. *REVSTAT-Statistical Journal*, *13*(1), 21–40.

Fattah, E. A., & Rue, H. (2022). *Approximate Bayesian inference for the interaction types 1, 2, 3 and 4 with application in disease mapping*. arXiv. https://doi.org/10.48550/arXiv.2206.09287

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian Analysis*, *1*(3), 515–534.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (1995). *Bayesian data analysis*. Chapman and Hall/CRC.

Ghosh, K., & Tiwari, R. C. (2007). Prediction of US cancer mortality counts using semiparametric Bayesian techniques. *Journal of the American Statistical Association*, *102*(477), 7–15.

Gneiting, T., & Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, *102*(477), 359–378.

Goicoa, T., Adin, A., Ugarte, M., & Hodges, J. (2018). In spatio-temporal disease mapping models, identifiability constraints affect PQL and INLA results. *Stochastic Environmental Research and Risk Assessment*, *32*(3), 749–770.

Gómez-Rubio, V. (2020). *Bayesian inference with INLA*. CRC Press.

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

Held, L., Schrödle, B., & Rue, H. (2010). Posterior and cross-validatory predictive checks: A comparison of MCMC and INLA. In *Statistical modeling and regression structures* (pp. 91–110). Springer.

Hofer, L. J., & Held, L. (2022). *Comparing confidence intervals for a binomial proportion with the interval score*. arXiv. https://doi.org/10.48550/arXiv.2207.03199

Knorr-Held, L. (2000). Bayesian modeling of inseparable space-time variation in disease risk. *Statistics in Medicine*, *19*(17–18), 2555–2567.

Lagazio, C., Biggeri, A., & Dreassi, E. (2003). Age–period–cohort models and disease mapping. *Environmetrics*, *14*(5), 475–490.

Liu, Z., & Rue, H. (2022). *Leave-group-out cross-validation for latent Gaussian models*. arXiv. https://doi.org/10.48550/arXiv.2210.04482

MacNab, Y. C. (2023). Adaptive Gaussian Markov random field spatiotemporal models for infectious disease mapping and forecasting. *Spatial Statistics*, *53*, 100726.

Martínez-Beneito, M. A., López-Quilez, A., & Botella-Rocamora, P. (2008). An autoregressive approach to spatio-temporal disease mapping. *Statistics in Medicine*, *27*(15), 2874–2889.

Martino, S., & Riebler, A. (2019). *Integrated nested Laplace approximations (INLA)* (pp. 1–19). Wiley StatsRef: Statistics Reference Online.

Orozco-Acosta, E., Adin, A., & Ugarte, M. D. (2021). Scalable Bayesian modelling for smoothing disease risks in large spatial data sets using INLA. *Spatial Statistics*, *41*, 100496.

Orozco-Acosta, E., Adin, A., & Ugarte, M. D. (2023). Big problems in spatio-temporal disease mapping: methods and software. *Computer Methods and Programs in Biomedicine*, *231*, 107403.

Paige, J., Fuglstad, G.-A., Riebler, A., & Wakefield, J. (2022). Spatial aggregation with respect to a population distribution: Impact on inference. *Spatial Statistics*, *52*, 100714.

Papoila, A. L., Riebler, A., Amaral-Turkman, A., São-João, R., Ribeiro, C., Geraldes, C., & Miranda, A. (2014). Stomach cancer incidence in Southern Portugal 1998–2006: A spatio-temporal analysis. *Biometrical Journal*, *56*(3), 403–415.

Pettit, L. I. (1990). The conditional predictive ordinate for the normal distribution. *Journal of the Royal Statistical Society: Series B (Methodological)*, *52*(1), 175–184.

Rabinowicz, A., & Rosset, S. (2022). Cross-validation for correlated data. *Journal of the American Statistical Association*, *117*(538), 718–731.

Riebler, A., Sørbye, S. H., Simpson, D., & Rue, H. (2016). An intuitive Bayesian spatial model for disease mapping that accounts for scaling. *Statistical Methods in Medical Research*, *25*(4), 1145–1165.

Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schröder, B., Thuiller, W., Warton, D. I., Wintle, B. A., Hartig, F., & Dormann, C. F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, *40*(8), 913–929.

Rue, H., & Held, L. (2005). *Gaussian Markov random fields: Theory and applications*. Monographs on Statistics and Applied Probability, Vol. 104. Chapman and Hall/CRC.

Rue, H., Martino, S., & Chopin, N. (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *71*(2), 319–392.

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

Sartorius, B., Lawson, A., & Pullan, R. (2021). Modelling and predicting the spatio-temporal spread of COVID-19, associated deaths and impact of key risk factors in England. *Scientific Reports*, *11*(1), 1–11.

Schmid, V., & Held, L. (2004). Bayesian extrapolation of space–time trends in cancer registry data. *Biometrics*, *60*(4), 1034–1042.

Schrödle, B., & Held, L. (2011). Spatio-temporal disease mapping using INLA. *Environmetrics*, *22*(6), 725–734.

Simpson, D., Rue, H., Riebler, A., Martins, T. G., & Sørbye, S. H. (2017). Penalising model component complexity: A principled, practical approach to constructing priors. *Statistical Science*, *32*(1), 1–28.

Sørbye, S. H., & Rue, H. (2014). Scaling intrinsic Gaussian Markov random field priors in spatial modelling. *Spatial Statistics*, *8*, 39–51.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & 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.

Ugarte, M. D., Adin, A., & Goicoa, T. (2016). Two-level spatially structured models in spatio-temporal disease mapping. *Statistical Methods in Medical Research*, *25*(4), 1080–1100.

Ugarte, M. D., Goicoa, T., Etxeberria, J., & Militino, A. F. (2012). Projections of cancer mortality risks using spatio-temporal P-spline models. *Statistical Methods in Medical Research*, *21*(5), 545–560.

Utazi, C., Thorley, J., Alegana, V., Ferrari, M., Nilsen, K., Takahashi, S., Metcalf, C. J. E., Lessler, J., & Tatem, A. (2019). A spatial regression model for the disaggregation of areal unit based data to high-resolution grids with application to vaccination coverage mapping. *Statistical Methods in Medical Research*, *28*(10–11), 3226–3241.

Van Niekerk, J., Bakka, H., Rue, H., & Schenk, O. (2021). New frontiers in Bayesian modeling using the INLA package in R. *Journal of Statistical Software*, *100*, 1–28.

Vicente, G., Adin, A., Goicoa, T., & Ugarte, M. D. (2023). High-dimensional order-free multivariate spatial disease mapping. *Statistics and Computing*, *33*(5), 104.

Wakefield, J. (2007). Disease mapping and spatial regression with count data. *Biostatistics*, *8*(2), 158–183.

Watanabe, S. (2010). Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research*, *11*(116), 3571–3594.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Orozco-Acosta, E., Riebler, A., Adin, A., & Ugarte, M. D. (2023). A scalable approach for short-term disease forecasting in high spatial resolution areal data. *Biometrical Journal*, *65,* 2300096. https://doi.org/10.1002/bimj.202300096

APPENDIX

A.1 | Predictive distribution of the counts

R-INLA provides posterior marginal estimates for the mortality rates λ_{it}, for $i = 1, \ldots, n$ and $t = 1, \ldots T$, after setting the arguments `control.predictor=list(compute=TRUE, link=1)` and `control.inla=list(return.marginals.predictor=TRUE)` of the main `inla()`-function call. It also provides predicted distribution estimates of mortality rates λ_{it^*} by setting the observed counts at time point t^* as `NA` and giving the corresponding offset N_{it^*} (in our case, the number of population at risk at municipality i and predicted year t^*).

Using the law of iterated expectation, the expected value for the posterior predictive counts is given by $\mu_{it^*} = E[E[O_{it^*}|\lambda_{it^*}]] = E[N_{it^*} \cdot \lambda_{it^*}] = N_{it^*} \cdot E[\lambda_{it^*}]$. We can also compute the posterior quantiles for the predicted counts by sampling from the marginal posterior of λ_{it^*} (Martino & Riebler, 2019). Our sampling scheme proceeds in two steps. First, we generate $S = 5000$ samples from the posterior marginal distribution of $\lambda_{it\star}$ using the function `inla.rmarginal()` function. Then, we generate values of the mortality counts $O^s_{it\star}$ from a Poisson distribution with mean $N_{it\star} \cdot \lambda^s_{it\star}$, for $s = 1, \ldots, S$, in order to compute its empirical quantiles.

A.2 | Additional tables

TABLE A1 Average values of prediction evaluation scores for models fitted with INLA (simplified Laplace approximation strategy) in the municipalities with the proportion of zero observed cases during the study period less or equal than 0.2.

	Space-time	1-year ahead			2-year ahead			3-year ahead		
Model	**interaction**	**$IS_{0.05}$**	**MAE**	**RMSE**	**$IS_{0.05}$**	**MAE**	**RMSE**	**$IS_{0.05}$**	**MAE**	**RMSE**
Classical	Type I	12.96	2.20	2.67	13.22	2.25	2.72	13.18	2.30	2.76
	Type II	–	–	–	–	–	–	–	–	–
	Type III	13.10	2.21	2.68	13.14	2.26	2.73	13.50	2.31	2.78
	Type IV	–	–	–	–	–	–	–	–	–
Disjoint	Type I	12.93	2.18	2.65	13.09	2.23	2.70	13.21	2.27	2.74
by provinces	Type II	12.65	2.13	2.59	12.93	2.16	2.63	13.25	2.21	2.67
	Type III	12.76	2.17	2.64	12.92	2.22	2.69	13.08	2.26	2.72
	Type IV	12.51	2.12	2.59	12.79	2.17	2.64	13.09	2.21	2.68
First-order	Type I	12.88	2.17	2.64	13.02	2.21	2.69	13.14	2.26	2.73
neighborhood	Type II	12.57	2.13	2.58	12.88	2.16	2.63	13.18	2.21	2.67
by provinces	Type III	12.73	2.16	2.63	12.92	2.21	2.68	12.96	2.25	2.71
	Type IV	12.48	2.12	2.58	12.72	2.16	2.63	12.96	2.21	2.68

Abbreviations: IS, score; MAE, mean absolute error; RMSE, root mean square error.

TABLE A2 Average values of prediction evaluation scores for models fitted with INLA (simplified Laplace approximation strategy) in the municipalities lying at the boundary between two or more provinces with at least two observed cases per 100,000 inhabitants during the whole study period.

	Space-time	1 year ahead			2-year ahead			3-year ahead		
Model	**interaction**	**$IS_{0.05}$**	**MAE**	**RMSE**	**$IS_{0.05}$**	**MAE**	**RMSE**	**$IS_{0.05}$**	**MAE**	**RMSE**
Classical	Type I	16.14	2.62	3.20	16.40	2.66	3.25	15.97	2.65	3.23
	Type II	–	–	–	–	–	–	–	–	–
	Type III	16.36	2.62	3.20	16.01	2.67	3.26	16.01	2.66	3.23
	Type IV	–	–	–	–	–	–	–	–	–
Disjoint	Type I	15.82	2.61	3.15	15.92	2.65	3.20	15.79	2.64	3.18
by provinces	Type II	15.87	2.63	3.18	16.07	2.64	3.21	16.19	2.64	3.19
	Type III	15.58	2.61	3.15	15.79	2.65	3.20	15.77	2.64	3.19
	Type IV	15.71	2.62	3.17	15.98	2.65	3.20	15.99	2.64	3.19
First-order	Type I	15.70	2.60	3.15	15.67	2.64	3.21	15.54	2.64	3.19
neighborhood	Type II	15.70	2.62	3.17	15.69	2.64	3.21	15.83	2.63	3.18
by provinces	Type III	15.50	2.60	3.15	15.50	2.65	3.22	15.36	2.65	3.20
	Type IV	15.60	2.60	3.15	15.63	2.64	3.20	15.62	2.63	3.18

Abbreviations: IS, score; MAE, mean absolute error; RMSE, root mean square error.