Inducing High Spatial Correlation with Randomly Edge-Weighted Neighborhood Graphs[*]

Danna L. Cruz-Reyes[†], Renato M. Assunção[‡] and Rosangela H. Loschi[§]

Abstract. Traditional models for areal data assume a hierarchical structure where one of the components is the random effects that spatially correlate the areas. The conditional autoregressive (CAR) model is the most popular distribution to jointly model the prior uncertainty about these spatial random effects. A limitation of the CAR distribution is its inability to accommodate high correlations between neighboring areas. We propose a new model for areal data that alleviates this problem. We represent the map by an undirected graph where the nodes are the areas, and randomly-weighted edges connect nodes that are neighbors. The model is based on a spatially-structured, multivariate Normal/Independent(NI) distribution, in which the precision matrix is indirectly built assuming a multivariate distribution for the random edge effects. The joint distribution for the edge effects is a spatial multivariate NI distribution that induces another NI distribution for the areas' spatial effects, which inherit its capacity to accommodate outliers and heavy-tailed behavior. Most important, it can produce a higher marginal correlation between the spatial effects than the CAR model overcoming one of the main limitations of this model. We fit the proposed model to analyze real cancer maps and compared its performance with several state-of-art competitors. Our proposed model provides better fitting in almost all cases.

MSC2020 subject classifications: Primary 05C75, 62H11; secondary 46N30.

Keywords: spatial Bayesian inference, random edge effects, spatial autoregression, normal/independent distribution family.

1 Introduction

The conditional autoregressive (CAR) model introduced by Besag (1974) has been one of the main drivers of spatial models for areal or lattice data. It appeared again in Besag et al. (1991) in its intrinsic version (ICAR), gaining visibility and importance as the main framework to specify joint distributions through the set of the conditional distribution of each area given its neighbors (Martínez-Beneito and Botella-Rocamora,

2019). The spatially motivated Markov property enjoyed by this model, coupled with its computational ease for Bayesian analysis and the availability of fast computation, is responsible for its great appeal (Gelfand and Vounatsou, 2003).

Typically, the CAR and ICAR models are used to describe the joint behavior of random effects $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_n)$, where θ_i is the spatial effect associated with area i on a map. Such effects are latent factors representing the spatial dependence beyond the small area geographical boundary. A common application is the disease mapping problem where death or disease counts are modeled using a Poisson distribution with a log-link specification. The latent random effect θ_i is an additive term to the linear part of the model. A key point regarding the construction of CAR or ICAR models is the specification of an appropriate neighborhood structure. Usually, we take as neighbors any pair of areas sharing boundaries. The adjacency information is included in the model through a $n \times n$ neighborhood matrix $\boldsymbol{A}$ with binary entries $a_{ik} = 1$ if area units i and k share a common border (denoted $k \sim i$), and $a_{ik} = 0$, otherwise. Usually, $a_{ii} = 0$. This approach is popular because it can be easily calculated using GIS (Geographic Information System) routines. Denote by $N_n(a, B)$ the n-variate normal distribution with mean a and covariance matrix B (n is omitted in the univariate case). The CAR model is specified considering the n univariate full conditional distributions given by

$$\theta_i | \boldsymbol{\theta}_{-i} \sim N\left(\frac{\varsigma}{m_i} \sum_{k \sim i} \theta_k, \frac{\tau_\theta^2}{m_i}\right), \tag{1.1}$$

where $\boldsymbol{\theta}_{-i} = (\theta_1, ..., \theta_{i-1}, \theta_{i+1}, ..., \theta_n)$, $m_i = \sum_k a_{ik}$ is the number of neighbors of region i, and ς is a spatial autocorrelation parameter. This conditional distribution depends on the neighbors of area i only, which defines the Markovian property. The joint distribution of $\boldsymbol{\theta}$ is the n-variate normal distribution with mean $\boldsymbol{0}$ and covariance matrix $\tau_\theta^2(\boldsymbol{M} - \varsigma \boldsymbol{A})^{-1}$ denoted by

$$\boldsymbol{\theta} \sim N_n(\boldsymbol{0}, \tau_\theta^2(\boldsymbol{M} - \varsigma \boldsymbol{A})^{-1}), \tag{1.2}$$

where $M = \text{diag}\{m_1, \ldots, m_n\}$. The covariance matrix $\tau_\theta^2(\boldsymbol{M} - \varsigma \boldsymbol{A})^{-1}$ is positive-definite if $1/\lambda_1 < \varsigma < 1/\lambda_p$, where λ_1 and λ_p denote the smallest and the largest eigenvalues of the matrix $\boldsymbol{M}^{-1/2} \boldsymbol{A} \boldsymbol{M}^{-1/2}$. The Markov property defined in terms of this neighborhood structure induces a sparse precision matrix facilitating the Bayesian computational approaches.

Spatial statisticians extended the CAR and ICAR models in many different directions. Classes of space-time generalized linear models are proposed by Knorr-Held and Best (2001), MacNab and Dean (2000), Martínez-Beneito et al. (2008), and Silva et al. (2008). Carlin and Banerjee (2003) and Jin and Carlin (2005) extend the idea to model spatial survival data. Spatially-varying parameters models are introduced by Assunção (2003), Assunção et al. (2002), and Gelfand et al. (2003). Generalized additive models can be found in Fahrmeir and Lang (2001). Extensions incorporating two correlated sets of spatial effects are proposed by Jin and Carlin (2005), Gelfand and Vounatsou (2003), and Knorr-Held et al. (2006).

Despite their popularity, the CAR and ICAR models have several critical issues (Martínez-Beneito and Botella-Rocamora, 2019, p. 134). In one line of criticism, Wall

(2004) showed that there are many puzzling results involving the CAR model. For example, the correlation between any pair of neighboring areas is negatively associated with the number of neighbors of each region but this is not sufficient to explain the dependence structure. In addition, sites with equal numbers of neighbors have different variances. Even more puzzling, the spatial structure depends on the CAR spatial parameter ς in an unexpected way: a pair of areas more correlated than another one for $\varsigma > 0.5$, for instance, may become less correlated for some other value of ς. The ICAR model may exacerbate these problems. Wall (2004) concluded that the spatial correlation induced by the CAR model is not intuitive and does not follow a practical scheme. All these counterintuitive results were explained by Assunção and Krainski (2009). They showed that the correlation structure between two areas depends on the entire neighborhood graph structure and not only on their immediate neighborhood. The second-largest eigenvalue modulus of the neighborhood matrix used in the CAR or ICAR models heavily influences these puzzling results. A more serious concern is the lack of ability of CAR or ICAR spatial effects models to produce high pairwise spatial correlation even when the parameter ς is near or equal to 1, as in the ICAR (Carlin and Banerjee, 2003; Gelfand and Vounatsou, 2003).

The CAR model assumes that the random spatial effects share a common global level of spatial autocorrelation, ranging from independence to strong spatial smoothing, which depends on a single parameter ς. A uniform level of spatial smoothness for the entire region may be an unrealistic assumption. Brewer and Nolan (2007) proposed replacing the single ς parameter in (1.1) with a pairwise ς_{ik} that varies according to the pair i and k of adjacent regions. Rather than proposing a direct probability distribution for the large number of ς_{ik} pairwise parameters, they impose a substantial dimensionality reduction by assuming $\varsigma_{ik} = \lambda_i \lambda_k / (\lambda_i + \lambda_k)$ where λ_i and λ_k are the random effects associated with areas i and k, respectively. Lu et al. (2007) proposed another modification in the classical CAR model. They replaced the fixed, deterministic, and known adjacency matrix $\boldsymbol{A}$ in (1.2) with a random thinned version. The a_{ik} binary random variables of adjacent regions are defined by conditionally independent Bernoulli outcomes whose success probabilities are modeled with a logistic regression based on known features of the pair. Hence, its final matrix is the adjacency matrix $\boldsymbol{A}$ with some elements flipped to 0. Similar approaches have been proposed by Ma and Carlin (2007) and Ma et al. (2010), which replace the logistic regression with a prior second-stage Bernoulli intrinsic autoregression (or Ising) model. Lee and Mitchell (2012) modified Lu et al. (2007) by changing the logistic regression to create the thinned adjacency matrix, and Lee et al. (2014) by adopting a uniform distribution on the set of the reduced binary adjacency matrix.

Langford et al. (1999) considered a multilevel approach to model the spatial random effects. The random effect θ_i for each area i is a sum of a set of independent random effects θ_j^* weighted by a measure of proximity Z_{ij} between the areas. The effects θ_j^* measure the effect of area j on other areas, and they are the same for all neighboring areas of j. They are also independent (spatially unstructured). In contrast, in the model we propose in this paper, each neighboring area of i has a different random effect, the one associated with the edge $[ij]$, and they are spatially structured. In Langford et al. (1999), the proximity measure Z_{ij} is defined in terms of the adjacency matrix in which

entry ω_{ij} in the i-th row and j-th column is an exponential function of the Euclidean distance between the centroids of areas i and j. Leyland et al. (2000) extended this previous model to the multivariate response setting. Congdon (2004) used a model with random effects for space, time, and age. For the spatial component, he considered the same model as Langford et al. (1999).

Another approach to modeling areal data can be found in Leroux et al. (2000) and MacNab and Dean (2000). These models assume that the precision matrix is a linear combination of a diagonal matrix and the precision matrix of the ICAR model. They accommodate over-dispersion but inherit the lack of interpretability issue of the CAR model. More recent developments in this topic include Prates et al. (2012), Rodrigues and Assunção (2012), and Datta et al. (2019). The CAR structure may not be appropriate to describe the spatial correlation if some areas experience atypical spatial effects. In Prates et al. (2012), the normal distribution involved in the CAR model is replaced by distributions in the generalized skew-normal/independent class, a robust class of distributions that is able to simultaneously accommodate heavy-tail and asymmetry. Rodrigues and Assunção (2012) proposes a spatial model in which the neighborhood structure is a parameter that must be estimated. This model preserves the Markov property as it assumes that, given the neighborhood graph, the areal parameters follow a conditional autoregressive model. The directed acyclic graph autoregressive model (DAGAR) (Datta et al., 2019) constructs the spatial precision matrix considering a directed acyclic graph (with a defined order) derived from the original undirected graph associated with the map. Although DAGAR is an order-dependent model, the authors show that the ordering choice does not substantially affect the results. Compared to CAR, the DAGAR model better estimates the latent spatial surface if the spatial correlation is weak or moderate. They have similar performances when the data have a stronger spatial correlation. An order-free version of the DAGAR model is obtained by averaging over all the possible orderings resulting in a special case of the model proposed by Datta et al. (2016). The DAGAR model provides a different approach to model multivariate Gaussian data always providing a positive definite covariance matrix. Its use is particularly appealing to analyze large spatial datasets due to its induced sparsity.

We will focus on spatial data that may be represented by an undirected graph. Our goal is to build a model described in Section 3 that is able to alleviate one of the main constraints of CAR-based models: its incapacity to generate high marginal correlations. As usual, we represent the map with a graph where the nodes stand for the small areas, and edges link geographically neighboring areas. The novelty in our approach is that we assign spatial random effects to the *edges* of this neighborhood graph. The spatial random effect of each area is the sum of effects of its incident edges. The joint distribution for the incident edge effects is a multivariate Normal/Independent(NI) distribution where the spatial covariance matrix has a CAR-like structure. The NI distribution class, widely used to build more robust models (Lange and Sinsheimer, 1993), is obtained as a scale mixture of the normal distribution and contains the Student-t and Normal distributions as particular cases. The resulting spatial model for the area effects is a multivariate NI distribution inheriting the heavy-tailed behavior and robustness to outliers. Our model is called RENeGe, standing for a randomly edge-weighted neighborhood graph model. In the graph theory literature, these edge effects are called edge

weights. This is the sense in which we use the expression edge-weighted neighborhood graph.

The paper is organized in the following way. In Section 2 we provide the main definitions and notations. We introduce our new RENeGe model and study its properties in Section 3. An important RENeGe property is the induction of a higher marginal correlation than the CAR model, overcoming one of the main limitations of the CAR and ICAR models. We study the marginal and conditional correlation properties of this model in Section 3.3. In Section 4, we explore the eigenstructure features of the RENeGe covariance matrix and illustrate its use to analyze the spatial patterns of five types of cancer rates. We also show empirically the RENeGe capacity to recover the spatial random latent effects. We end up with conclusions in Section 5. Besides this main text, we provide an extensive additional analysis of RENeGein the Supplementary Material (Cruz et al., 2023), as described at the end of this paper.

2   Preliminary definitions

Consider a map with n contiguous geographical regions. The map is identified with an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{v_1, \dots, v_n\}$ is the set of vertices or nodes representing the areas, and $\mathcal{E}$ is the set of p edges connecting unordered pairs of distinct vertices and representing the adjacency relationship among regions. The edge connecting $v_i \in \mathcal{V}$ and $v_j \in \mathcal{V}$ is alternatively represented by $[ij]$ or (v_i, v_j). We assume that the edges are undirected, implying that $[ij] = [ji]$. If two nodes v_i and $v_j \in \mathcal{V}$ are connected, this will be denoted by $v_i \sim v_j$. When $v_i \in \mathcal{V}$ is a node in the edge $[ij] \in \mathcal{E}$, we say that the edge is incident on v_i. Usually, the graph is visualized by plotting each vertex on a typical spatial location inside the corresponding area, such as its geographical centroid. The spatial neighborhood structure is represented by the set $\mathcal{E}$ of edges. These edges will determine the stochastic dependence between the areas. The most common choice is to have an edge $[ij] = (v_i, v_j)$ when areas i and j share boundaries but other neighborhood structures may be assumed. A *simple path* from node v_1 to node v_m is a set of nodes $v_1, v_2, \dots, v_m$ in $\mathcal{V}$ which are connected by edges $(v_1, v_2), \dots, (v_{m-1}, v_m)$ such that v_1 and v_m may be the same (in this special case, it is called a *circuit*), and for each i from 1 to $m-1$, there exists an edge between v_i and v_{i+1}, and no vertex appears more than once in the sequence, except possibly for v_1 and v_m. A graph is said to be *connected* if, for any pair of nodes v_i and v_j, there is at least one path connecting them. Although it is not strictly necessary, we assume the graph is connected. We also assume the most common situation in practice, that the number p of edges is larger than the number n of nodes. The *adjacency matrix* $\boldsymbol{A}_v$ of $\mathcal{G}$ is an $n \times n$ binary matrix representing the neighborhood structure. That is, $\boldsymbol{A}_v(i, j) = a_{ij} = 1$, if $v_i \sim v_j$ (or, equivalently, if $[ij] \in \mathcal{E}$), and it is 0, otherwise. We assume that $a_{ii} = 0$ for all $i \in \mathcal{V}$.

Associated with the original graph $\mathcal{G}$, we define the *graph of edges* $\mathcal{L}(\mathcal{G})$, a fundamental tool for our purposes. The graph of edges represents the adjacency relations among the edges of the original graph $\mathcal{G}$. The nodes in $\mathcal{L}(\mathcal{G})$ are the edges $[ij] \in \mathcal{E}$ connecting the nodes v_i and v_j, with $i \neq j$. The edges in $\mathcal{L}(\mathcal{G})$ are also determined by the topology of $\mathcal{G}$. Two nodes $[ij]$ and $[kl]$ in $\mathcal{L}(\mathcal{G})$ are adjacent if, and only if, the edges $[ij]$ and $[kl]$

are incident on a common vertex, which means the pair of neighboring edges must be of the form $[ij]$ and $[jk]$ for some $v_j \in \mathcal{V}$. Let $\mathcal{I}_i = \{[ik] \in \mathcal{E}, v_k \in \mathcal{V}\}$ be the set of edges incident on area i. The adjacency matrix $\boldsymbol{A}_e$ associated with the graph of edges $\mathcal{L}(\mathcal{G})$ is a $p \times p$ matrix defined based on the edges' neighborhood structure in $\mathcal{L}(\mathcal{G})$. That is, $\boldsymbol{A}_e([ij], [jk]) = a_{[ij][jk]} = 1$, if $[ij]$ and $[jk]$ belong to $\mathcal{I}_j$. Otherwise, $\boldsymbol{A}_e([ij], [jk]) = 0$. Finally, the *incidence matrix* $\boldsymbol{C}$ associated with $\mathcal{G}$ is an $n \times p$ binary matrix such that $c_{ie} = 1$ if edge e is incident on node i, and $c_{ie} = 0$ otherwise. Figures 1-3 present a toy example of the graph of edges $\mathcal{L}(\mathcal{G})$.
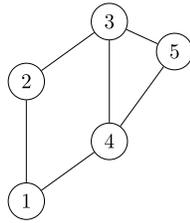
Figure 1: $\mathcal{G}$.
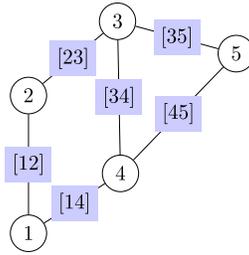
Figure 2: $\mathcal{E}$.
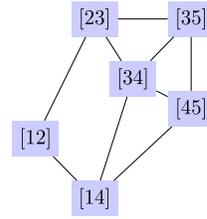
Figure 3: $\mathcal{L}(\mathcal{G})$ edges.

3   Randomly edge-weighted neighborhood graphs model RENeGe for spatial random effects

In a map partitioned into n contiguous regions, let $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_n)$ be a random vector where θ_i is a random variable associated with the i-th region. Assume that the coordinates in $\boldsymbol{\theta}$ are spatially correlated. The vector $\boldsymbol{\theta}$ may represent the spatial effects in a hierarchical model. The goal is to model the uncertainty about $\boldsymbol{\theta}$ accounting for this spatial correlation. Differently from the CAR-related models, we build the prior distribution for $\boldsymbol{\theta}$ in such a way that their correlation is induced by the prior correlation in spatial effects assigned to the neighborhood graph edges. The novelty of our approach is the use of a distribution over the pairs of neighboring areas to induce a distribution over individual areas. We look for a model able to generate a stronger marginal correlation between neighboring areas, overcoming one of the main limitations of the CAR model.

3.1   Modeling the edge random effects ρ

Assume that our map is represented by the connected graph $\mathcal{G}$. Each component of $\boldsymbol{\theta}$ is a node in $\mathcal{G}$. There are p undirected edges connecting pairs of neighboring nodes in $\mathcal{G}$. Let $\rho_{[ij]} \in \mathbb{R}$ be a random variable or edge effect associated with the edge $[ij]$ connecting nodes θ_i and θ_j. The vector of such edge effects is $\boldsymbol{\rho} \in \mathbb{R}^p$.

Spatial effects observed in areal data are surrogates for unknown or unobserved factors that vary on a scale extending beyond the geographical boundaries of the small areas. As these effects spread throughout the surrounding neighborhood of an area, they

act on the edges that connect neighbouring areas. Instead of directly modeling the θ_i effects in each small area, we can obtain them as a result of a random effects model for the neighborhood graph edges. The $[ij]$ edge contributes to θ_i and θ_j, both connected by this edge. If the edge effect $\rho_{[ij]}$ is large and positive (or large and negative), then θ_i and θ_j will tend to be large and positive (or large and negative). A second mechanism that induces a spatial correlation in $\boldsymbol{\theta}$ is the assumed distribution for the edge effects $\boldsymbol{\rho}$. Two edges, $[ij]$ and $[ik]$, incident on node i, must be correlated due to hidden factors that extrapolate the boundaries of the areas. For example, environmental factors may influence the spatial pattern to a geographical extent that is much larger than an area and its immediate neighbors. Hence, we can expect to observe similar values in the map on a geographical scale that goes beyond spatial local effects. Unlike CAR, the proposed model will allow neighboring areas i and j to be more strongly correlated than in the CAR or ICAR models due to the presence of these two mechanisms. This stronger correlation will also affect pairs of areas that are not directly connected in $\mathcal{G}$.

We represent the dependence structure of the components in $\boldsymbol{\rho}$ through the graph of edges $\mathcal{L}(\mathcal{G})$ and its associated $p \times p$ adjacency matrix $\boldsymbol{A}_e$. Denote by $m_{[ij]}$ the number of edges in $\mathcal{L}(\mathcal{G})$ that are neighbors of the edge $[ij]$ or, equivalently, the sum of the $[ij]$-th row elements of $\boldsymbol{A}_e$. Define the $p \times p$ diagonal matrix $\boldsymbol{M}_e$ which diagonal entries are the p elements $m_{[ij]}$.

Our goal is to build a flexible model for the spatial random effects $\boldsymbol{\theta}$ from the prior knowledge about the edge effects $\boldsymbol{\rho}$. To this end, consider a latent non-negative random variable U with pdf $f(.|\boldsymbol{\ell})$, where $\boldsymbol{\ell}$ is a vector of hyperparameters. Assume that the edge effects vector $\boldsymbol{\rho} \in \mathbb{R}^p$ has a centered p-variate Normal/independent (NI) distribution (Lange and Sinsheimer, 1993) with probability density function (pdf) given by

$$f(\boldsymbol{\rho}|\sigma_{\boldsymbol{\theta}}^2, \ell) = \int_0^\infty \left(\frac{u}{2\pi\sigma_{\boldsymbol{\theta}}^2}\right)^{p/2} |(\boldsymbol{M}_e - \gamma\boldsymbol{A}_e))|^{1/2} \exp\left\{-\frac{u}{2\sigma_{\boldsymbol{\theta}}^2}\boldsymbol{\rho}^t(\boldsymbol{M}_e - \gamma\boldsymbol{A}_e)\boldsymbol{\rho}\right\} f(u|\boldsymbol{\ell})du,$$
$$(3.1)$$

which, in a hierarchical representation, becomes

$$\boldsymbol{\rho}|U, \sigma_{\boldsymbol{\theta}}^2, \gamma \sim N_p(\boldsymbol{0}, \sigma_{\boldsymbol{\theta}}^2 U^{-1}(\boldsymbol{M}_e - \gamma\boldsymbol{A}_e)^{-1}) \qquad (3.2)$$
$$U|\boldsymbol{\ell} \sim f(u|\boldsymbol{\ell}),$$

where $\boldsymbol{\ell} > 0$ is the shape parameter, $\sigma_{\boldsymbol{\theta}}^2 > 0$ is the precision parameter, $(\boldsymbol{M}_e - \gamma\boldsymbol{A}_e)^{-1}$ is a $p \times p$ scale matrix representing the correlation among the edges' random effects, and γ is a constant assuming values in the interval $(1/\lambda_p, 1/\lambda_1)$, where λ_1 and λ_p are, respectively, the minimum and the maximum eigenvalues of $\boldsymbol{M}_e^{-1/2}\boldsymbol{A}_e\boldsymbol{M}_e^{-1/2}$. The matrix $(\boldsymbol{M}_e - \gamma\boldsymbol{A}_e)$ is a sparse matrix that induces conditional independence between unlinked edges. The constraint over γ guarantees that the matrix is a positive-definite matrix.

Different models for $\boldsymbol{\rho}$ are obtained if we assume different prior distributions for U (Lange and Sinsheimer, 1993). In (3.2), if the distribution of U is degenerate with $P(U = 1) = 1$, the joint distribution for the edge random effects is the Normal distribution

$$\boldsymbol{\rho} \mid \sigma_{\boldsymbol{\theta}}^2, \gamma \sim N_p\left(\boldsymbol{0}, \sigma_{\boldsymbol{\theta}}^2(\boldsymbol{M}_e - \gamma\boldsymbol{A}_e)^{-1}\right). \qquad (3.3)$$

Given $\sigma_{\boldsymbol{\theta}}^2$ and γ, the distribution in (3.3) implies a CAR correlation structure for the edge effects $\boldsymbol{\rho}$.

Alternatively, we obtain a heavy-tailed model for $\boldsymbol{\rho}$ if $U \mid \ell \sim \text{Gamma}(\ell/2, \ell/2)$ in (3.2), with $E(U \mid \ell) = 1$ and $V(U \mid \ell) = 2/\ell$. This implies that the edge effects vector $\boldsymbol{\rho}$ follows the centered p-variate Student-t distribution with scale matrix $\sigma_{\boldsymbol{\theta}}^2(\boldsymbol{M}_e - \gamma\boldsymbol{A}_e)^{-1}$ and $\ell > 0$ degree of freedom:

$$\boldsymbol{\rho} \mid \sigma_{\boldsymbol{\theta}}^2, \gamma \sim T_p\left(\boldsymbol{0}, \sigma_{\boldsymbol{\theta}}^2(\boldsymbol{M}_e - \gamma\boldsymbol{A}_e)^{-1}, \ell\right) , \tag{3.4}$$

with density function

$$p(\boldsymbol{\rho} \mid \sigma_{\boldsymbol{\theta}}^2, \gamma, \ell) = \frac{\Gamma\left(\frac{\ell+p}{2}\right)}{\Gamma\left(\frac{\ell}{2}\right)\pi^{p/2}(\ell)^{p/2}} \left|\frac{(\boldsymbol{M}_e - \gamma\boldsymbol{A}_e)}{\sigma_{\boldsymbol{\theta}}^2}\right|^{1/2} \left[1 + \boldsymbol{\rho}^t\frac{(\boldsymbol{M}_e - \gamma\boldsymbol{A}_e)}{\sigma_{\boldsymbol{\theta}}^2}\boldsymbol{\rho}\right]^{-(\ell+p)/2} .$$

The covariance matrix for the edge effects is given by

$$\text{Cov}[\boldsymbol{\rho} \mid \sigma_{\boldsymbol{\theta}}^2, \gamma, \ell] = \frac{\ell}{\ell - 2}\sigma_{\boldsymbol{\theta}}^2(\boldsymbol{M}_e - \gamma\boldsymbol{A}_e)^{-1}, \tag{3.5}$$

if $\ell > 2$. The same CAR structure in (3.3) appears also if we adopt an asymptotic argument by letting $\ell \to \infty$ in (3.5) as, in this case, we have $\ell/(\ell-2) \to 1$. Another way to obtain the same CAR structure as in (3.3) is to replace the scale matrix in (3.4) with $(\boldsymbol{M}_e - \gamma\boldsymbol{A}_e)^{-1}(\ell-2)/\ell$.

A non-centered approach of the proposed model can be easily obtained considering in (3.2) the linear transformation $\boldsymbol{\rho} + \boldsymbol{\mu}$, $\boldsymbol{\mu} \in \mathbb{R}^p$. If $\boldsymbol{\rho}|U, \sigma_{\boldsymbol{\theta}}^2, \gamma \sim N_p(\boldsymbol{0}, \sigma_{\boldsymbol{\theta}}^2 U^{-1}(\boldsymbol{M}_e - \gamma\boldsymbol{A}_e)^{-1})$ then, considering properties of the normal distribution, it follows that $\boldsymbol{\rho} + \boldsymbol{\mu} \mid U, \sigma_{\boldsymbol{\theta}}^2, \gamma \sim N_p(\boldsymbol{\mu}, \sigma_{\boldsymbol{\theta}}^2 U^{-1}(\boldsymbol{M}_e - \gamma\boldsymbol{A}_e)^{-1})$. However, for the disease mapping applications in which we are interested, the centered formulations in expressions (3.3) and (3.4) are more appropriate.

3.2 The induced distribution for the areas' random effects θ

The proposed model assumes that the random effect θ_i associated with area i is a linear combination of the effects of edges incident on area i, that is,

$$\theta_i = \sum_{[ik] \in \mathcal{I}_i} \rho_{[ik]} \;\Rightarrow\; \boldsymbol{\theta} = \boldsymbol{C}\,\boldsymbol{\rho}. \tag{3.6}$$

where $\boldsymbol{C}$ is the $n \times p$ incidence matrix and $\mathcal{I}_i$ is the set of edges incident on area i. Assuming that $\boldsymbol{\rho}$ has the NI distribution given in (3.2), we obtain that $\boldsymbol{\theta}$ also has a NI distribution which is hierarchically represented as

$$\begin{aligned}\boldsymbol{\theta}|U, \sigma_{\boldsymbol{\theta}}^2, \gamma &\sim N_p(\boldsymbol{0}, \sigma_{\boldsymbol{\theta}}^2 U^{-1}\boldsymbol{C}(\boldsymbol{M}_e - \gamma\boldsymbol{A}_e)^{-1}\boldsymbol{C}^t) \\ U|\boldsymbol{\ell} &\sim f(u|\boldsymbol{\ell}).\end{aligned} \tag{3.7}$$

The scale matrix $\boldsymbol{C}(\boldsymbol{M}_e - \gamma\boldsymbol{A}_e)^{-1}\boldsymbol{C}^t$ is definite-positive. As $\gamma \in (1/\lambda_p, 1/\lambda_1)$, this can be proved by assuming Theorem 4.2.1 in Golub and Van Loan (1996). In this proof,

we invoke Lemma 2.17 in Bapat (2014), which guarantees that, as $\mathcal{G}$ is not a bipartite graph, the rank of $\boldsymbol{C}$ is n.

As a consequence, if we assume the model in (3.3) for $\boldsymbol{\rho}$ by considering that $P(U = 1) = 1$, we prove that the linear transformation in (3.6) has the following n-variate normal distribution

$$\boldsymbol{\theta} \mid \sigma_{\boldsymbol{\theta}}^2, \gamma \sim N_n(\boldsymbol{0}, \sigma_{\boldsymbol{\theta}}^2 \boldsymbol{C}(\boldsymbol{M}_e - \gamma \boldsymbol{A}_e)^{-1} \boldsymbol{C}^t). \tag{3.8}$$

We name the model in (3.8) RENeGe-N. The covariance structure in (3.8) results from the transformation of the CAR covariance structure assumed for $\boldsymbol{\rho}$ in (3.3) by the incidence matrix $\boldsymbol{C}$. That is,

$$\mathrm{Cov}[\boldsymbol{\theta} \mid \sigma_{\boldsymbol{\theta}}^2, \gamma] = \boldsymbol{C} \, \mathrm{Cov}(\boldsymbol{\rho} \mid \sigma_{\boldsymbol{\theta}}^2, \gamma) \, \boldsymbol{C}^t. \tag{3.9}$$

Another model, called RENeGe-T, is obtained if we assume the heavy-tailed behavior for the edge random effects in (3.4). Taking $U \sim \mathrm{Gamma}(\ell/2, \ell/2)$ in (3.7), then $\boldsymbol{\theta}$ has the following n-variate Student-t distribution

$$\boldsymbol{\theta} \mid \sigma_{\boldsymbol{\theta}}^2, \gamma, \ell \sim T_n(\boldsymbol{0}, \sigma_{\boldsymbol{\theta}}^2 \boldsymbol{C}(\boldsymbol{M}_e - \gamma \boldsymbol{A}_e)^{-1} \boldsymbol{C}^t, \ell) \tag{3.10}$$

with covariance structure given by

$$\mathrm{Cov}[\boldsymbol{\theta} \mid \sigma_{\boldsymbol{\theta}}^2, \gamma, \ell] = \boldsymbol{C} \, \mathrm{Cov}(\boldsymbol{\rho} \mid \sigma_{\boldsymbol{\theta}}^2, \gamma) \, \boldsymbol{C}^t = \frac{\ell}{\ell - 2} \sigma_{\boldsymbol{\theta}}^2 \boldsymbol{C}(\boldsymbol{M}_e - \gamma \boldsymbol{A}_e)^{-1} \boldsymbol{C}^t. \tag{3.11}$$

Model (3.10) can also be obtained without making use of the hierarchical structure in (3.7). Starting directly from the t-distribution in expression (3.4), Kotz and Nadarajah (2004) show that $\boldsymbol{\theta}$ has the distribution given in (3.10). However, the hierarchical representation in (3.7) is more general, giving the option to build other heavy-tailed models for $\boldsymbol{\theta}$ spatial effects, and it is computationally more efficient.

The distribution in (3.10) is a special case of the generalized n-variate Student-t distribution (Arellano-Valle and Bolfarine, 1995) with location $\boldsymbol{0}$, dispersion matrix $\sigma_{\boldsymbol{\theta}}^2 \boldsymbol{C}(\boldsymbol{M}_e - \gamma \boldsymbol{A}_e)^{-1} \boldsymbol{C}^t$, scale parameter and degree of freedom both equal to ℓ, which is denoted by $\boldsymbol{\theta} \mid \sigma_{\boldsymbol{\theta}}^2, \gamma, \ell \sim \mathrm{T}_n\left(\boldsymbol{0}, \sigma_{\boldsymbol{\theta}}^2 \boldsymbol{C}(\boldsymbol{M}_e - \gamma \boldsymbol{A}_e)^{-1} \boldsymbol{C}^t, \ell, \ell\right)$. Let $\boldsymbol{K} = \boldsymbol{C}(\boldsymbol{M}_e - \gamma \boldsymbol{A}_e)^{-1} \boldsymbol{C}^t$. Consequently, it follows from results in Arellano-Valle and Bolfarine (1995) and Kotz and Nadarajah (2004) that

(i) The marginal distribution of each component θ_i of $\boldsymbol{\theta}$ is the univariate Student-t distribution $\theta_i \mid \sigma_{\boldsymbol{\theta}}^2, \gamma, \ell \sim T(0, \sigma_{\boldsymbol{\theta}}^2 \boldsymbol{K}_{ii}, \ell)$, a univariate centered t distribution where $\boldsymbol{K}_{ii}$ is an entry that lies in the ith row and ith column of matrix $\boldsymbol{K}$.

(ii) Let the column vectors $\boldsymbol{\theta}_A$ and $\boldsymbol{\theta}_B$ define a partition of $\boldsymbol{\theta}$ where $\boldsymbol{\theta}_A$ and $\boldsymbol{\theta}_B$ have dimensions n_A and $n_B = n - n_A$, respectively. Let the matrices $\boldsymbol{K}_{AA}$, $\boldsymbol{K}_{BB}$ and $\boldsymbol{K}_{AB}$, of order $n_A \times n_A$, $n_B \times n_B$ and $n_A \times n_B$, respectively, be the partition of $\boldsymbol{K}$ induced by $\boldsymbol{\theta}_A$ and $\boldsymbol{\theta}_B$. The conditional distribution of $\boldsymbol{\theta}_A$ given $\boldsymbol{\theta}_B$ is the generalized n_A-variate Student-t distribution

$$\boldsymbol{\theta}_A \mid \boldsymbol{\theta}_B, \sigma_{\boldsymbol{\theta}}^2, \gamma, \ell \sim \mathrm{T}_{n_A}\left(\boldsymbol{\mu}^\star, \boldsymbol{\Sigma}^\star, \phi^\star, \ell + n - n_A\right),$$

where the location is $\boldsymbol{\mu}^\star = \boldsymbol{K}_{AB}\boldsymbol{K}_{BB}^{-1}\boldsymbol{\theta}_B$, the dispersion matrix is $\boldsymbol{\Sigma}^\star = \sigma_\theta^2[\boldsymbol{K}_{AA} - \boldsymbol{K}_{AB}\boldsymbol{K}_{BB}^{-1}\boldsymbol{K}_{BA}]$, and the scale parameter is $\phi^\star = \ell + \boldsymbol{\theta}_B^t\boldsymbol{K}_{BB}^{-1}\boldsymbol{\theta}_B/\sigma_\theta^2$.

The choice of ℓ in (3.10) may inflate or deflate the variance of θ_i, being handy to establish more or less informative prior distributions for $\boldsymbol{\theta}$. It is a challenging task to define an appropriate value for ℓ since no prior information about ℓ is usually available. Different approaches to building objective priors for ℓ have been discussed in the literature (see discussion and references in Villa and Walker (2014)). Although non-objective, the Gamma prior distribution suggested by Juárez and Steel (2010) is proper, easy to implement and has a performance similar to the independent Jeffreys' prior introduced by Fonseca et al. (2008) for $\ell < 50$. In our analysis, following Juárez and Steel (2010), we choose a priori $\ell \sim \text{Gamma}(2, 1/10)$ with density

$$f(\ell) = \frac{\ell}{100}\exp\{-\ell/10\}. \tag{3.12}$$

There is another way to build a heavy-tailed Student-t model for the spatial effects $\boldsymbol{\theta}$ of the areas through a spatial model defined on the edge effects $\boldsymbol{\rho}$. In (3.2), the precision structure for $\boldsymbol{\rho}$ mixes a CAR-sparse matrix $\boldsymbol{M}_e - \gamma\boldsymbol{A}_e$ with a random positive scalar factor U. RENeGe-T and RENeGe-N are special cases obtained with different prior specifications for U. In Section 1 in the Supplementary Material (Cruz et al., 2023), we present a third model (RENeGe-TS) in which the randomness in the precision structure of $\boldsymbol{\rho}$ is not determined by a single scalar random variable and it is not sparse. Rather, the covariance structure of $\boldsymbol{\rho}$ will depend on an inverse-Wishart matrix $\boldsymbol{S}$ centered in a sparse CAR-type matrix.

3.3   The covariance structure of θ

By assuming a CAR-type structure for $\boldsymbol{\rho}$, the correlation among the edge effects inherits the counterintuitive results pointed out by Wall (2004) and Assunção and Krainski (2009). As $\text{Cov}(\boldsymbol{\theta}) = \boldsymbol{C}\,\text{Cov}(\boldsymbol{\rho})\,\boldsymbol{C}^t$, these results are passed on to the covariance structure of $\boldsymbol{\theta}$. Our goal in this section is to investigate the impact of the covariance of $\boldsymbol{\rho}$ in the marginal $\text{Cov}(\theta_i, \theta_j)$. To simplify, in this section, we set $\ell(\ell-2)^{-1}\sigma_{\boldsymbol{\theta}}^2 = 1$ in expression (3.11) and $\sigma_{\boldsymbol{\theta}}^2 = 1$ in expression (3.9). Results obtained here and in next section are also valid for RENeGe-TS presented in Section 1 in the Supplementary Material (Cruz et al., 2023). It is not possible to obtain the exact value for $\text{Cov}(\theta_i, \theta_j)$ as it requires the calculation of the inverse matrix $(\boldsymbol{M}_e - \gamma\boldsymbol{A}_e)^{-1}$. To investigate the $\text{Cov}(\rho)$ impact, we consider the relationship

$$(\boldsymbol{M}_e - \gamma\boldsymbol{A}_e)^{-1} = (\boldsymbol{I} - \gamma\boldsymbol{W}_e)^{-1}\boldsymbol{M}_e^{-1}, \tag{3.13}$$

where $\boldsymbol{W}_e$ is a $p \times p$ matrix with entries given by $\boldsymbol{W}_{e[ik][jl]} = a_{[ik][jl]}/m_{[ik]}$. The binary variable $a_{[ik][kj]}$ is equal to 1 if, and only if, $[ik]$ and $[kj]$ are incident edges on a same node k, $a_{[ik][ik]} = 0$ for all edges $[ik]$, and $m_{[ik]}$ denotes the total number of edges in $\mathcal{L}(\mathcal{G})$ that are neighbors of edge $[ik]$. The element $m_{[ik]}$ is the $[ik]$-th coordinate of the diagonal matrix $\boldsymbol{M}_e$.

The matrix $\boldsymbol{W}_e$ is a transition matrix of a random walk defined on the graph of edges $\mathcal{L}(\mathcal{G})$. Its elements are non-negative, its rows sum to 1, and the diagonal elements are equal to zero. A particle sitting on an edge $[ij]$ at time t moves to a different edge randomly selected among those in $\mathcal{I}_i$, the neighbors of $[ij]$ in $\mathcal{L}(\mathcal{G})$, with equal probability. The power matrix $\boldsymbol{W}_e^s$ represents the transition probabilities in s steps for this Markov chain. The non-zero elements in the $[ij]$-th row of $\boldsymbol{W}_e^s$ indicate the other edges to which one may visit s steps ahead. That is, $\boldsymbol{W}_{e[ik][jl]}^s > 0$ if there is at least one path composed of s edges in $\mathcal{L}(\mathcal{G})$ linking the initial edge $[ik]$ to the final edge $[jl]$.

For the matrix $\boldsymbol{W}_e$ to be ergodic and aperiodic, $\mathcal{L}(\mathcal{G})$ must be a connected graph. As we assume that $\mathcal{G}$ is a connected graph, the graph of edges $\mathcal{L}(\mathcal{G})$ is also connected. Furthermore, $\boldsymbol{W}_e$ and $\boldsymbol{M}_e^{-1/2}\boldsymbol{A}_e\boldsymbol{M}_e^{-1/2}$ are similar matrices and, therefore, they share the same eigenvalues. Thus, for $\gamma \in (1/\lambda_p, 1/\lambda_1)$, the matrix $(\boldsymbol{I} - \gamma\boldsymbol{W}_e)$ is non-singular and, from results in Assunção and Krainski (2009), its inverse is given by

$$(\boldsymbol{I} - \gamma\boldsymbol{W}_e)^{-1} = \boldsymbol{I} + \gamma\boldsymbol{W}_e + \gamma^2\boldsymbol{W}_e^2 + \gamma^3\boldsymbol{W}_e^3 + \cdots. \tag{3.14}$$

Substituting the right hand side of (3.14) into expressions (3.11) and (3.9), we have that

$$\mathrm{Cov}[\boldsymbol{\theta}] = \boldsymbol{C}(\boldsymbol{I} + \gamma\boldsymbol{W}_e + \gamma^2\boldsymbol{W}_e^2 + \gamma^3\boldsymbol{W}_e^3 + \cdots)\boldsymbol{M}_e^{-1}\boldsymbol{C}^t. \tag{3.15}$$

For $i \sim j$, the element $\mathrm{Cov}(\theta_i, \theta_j)$ in the matrix in (3.15) is given by

$$\frac{1}{m_{[ij]}} + \gamma \sum_{[jl]\in\mathcal{I}_j} \frac{1}{m_{[jl]}} \left(\sum_{\substack{[ik]\in\mathcal{I}_i \\ [ik]\sim[jl]}} \frac{1}{m_{[ik]}} \right) + \gamma^2 \sum_{[jl]\in\mathcal{I}_j} \frac{1}{m_{[jl]}} \left(\sum_{[ik]\in\mathcal{I}_i} \frac{1}{m_{[ik]}} \sum_{\substack{[lr]\sim[ik] \\ [lr]\sim[jl]}} \frac{1}{m_{[lr]}} \right) + \ldots \tag{3.16}$$

while, if $i = j$, we have $\mathrm{Cov}(\theta_i, \theta_i) = \mathrm{Var}(\theta_i)$ given by

$$\sum_{[ij]\in\mathcal{I}_i} \frac{1}{m_{[ij]}} + \gamma \sum_{[ij]\in\mathcal{I}_i} \frac{1}{m_{[ij]}} \left(\sum_{\substack{[ik]\in\mathcal{I}_i \\ [ik]\sim[ij]}} \frac{1}{m_{[ik]}} \right) + \gamma^2 \sum_{[ij]\in\mathcal{I}_i} \frac{1}{m_{[ij]}} \left(\sum_{[ik]\in\mathcal{I}_i} \frac{1}{m_{[ik]}} \sum_{\substack{[lr]\sim[ik] \\ [lr]\sim[ij]}} \frac{1}{m_{[lr]}} \right) + \ldots$$

For pairs of areas that are not adjacent, the formula is more convoluted, summing over all the paths connecting the two areas. The proof can be found in Supplementary Material?? S1.

The expansion in expression (3.16) shows that the correlation structure between θ_i and θ_j cannot be explained considering only the interaction between first-order neighbors. The covariance $\mathrm{Cov}(\theta_i, \theta_j)$ is a polynomial in γ where the m-th order coefficient is a weighted sum of all paths of order m connecting edges $[ik]$ and $[jl]$. If we consider only a first order approximation, $\mathrm{Cov}(\theta_i, \theta_j) \approx m_{[ij]}^{-1}$, inversely proportional to the number $m_{[ij]}$ of edges in $\mathcal{L}(\mathcal{G})$ that are neighbors of edge $[ij]$. Consequently, it is inversely proportional to the total number of neighbors in $\mathcal{G}$ of nodes θ_i and θ_j since $m_{[ij]} = m_i + m_j - 2$ where m_s is the number of neighbors of node θ_s, $s = i, j$. Including, for instance, the third term, the product $a_{[ik][uv]}a_{[uv][jl]}$ is equal to 1 only if the edges

$[ik]$ and $[jl]$ are 2nd-order neighbors in $\mathcal{L}(\mathcal{G})$. That implies that the edge $[uv]$ connects nodes i and j in $\mathcal{G}$, establishing a more complex dependence structure between the related nodes in the original graph $\mathcal{G}$ by imposing that (i) nodes k and j are 1st-order neighbor, (ii) nodes k and l and nodes i and j are 2nd-order neighbors and (iii) and nodes i and l are 3rd-order neighbors.

3.3.1 The regular lattice case

We obtain additional theoretical properties of our model by considering that the graph $\mathcal{G}$ is a regular square lattice with D^2 vertices symmetrically wrapped into a torus. In this simpler structure, the number of first-order neighbors of each node θ_i is constant and equal to four. To simplify the expressions, we take $\sigma_{\boldsymbol{\theta}}^2 = 1$. Denote by $\mathrm{Cov}(\rho_{[i*]} \overset{D}{\sim} \rho_{[j*]})$, for all $[i*] \in \mathcal{I}_i$ and $[j*] \in \mathcal{I}_j$, the covariance between the edge effects $\rho_{[i*]}$ and $\rho_{[j*]}$ whenever the edges $[i*]$ and $[j*]$ are Dth-order neighbors in $\mathcal{L}(\mathcal{G})$. The index D is omitted from the notation for the first-order neighbor case.

If $\mathcal{G}$ is a regular lattice, then the graph of edges $\mathcal{L}(\mathcal{G})$ is also a regular lattice. The covariance between any pair of nodes θ_i and θ_j in $\mathcal{G}$ depends on the effects related to the edges belonging to $\mathcal{I}_i$ and $\mathcal{I}_j$:

$$\mathrm{Cov}(\theta_i, \theta_j) = \begin{cases} \mathrm{Var}(\rho_{[ij]}) + 6\ \mathrm{Cov}(\rho_{[i*]} \sim \rho_{[j*]}) + 9\ \mathrm{Cov}(\rho_{[i*]} \overset{2}{\sim} \rho_{[j*]}), & \text{if } i \sim j, \\ \mathrm{Cov}(\rho_{[i*]} \overset{D}{\sim} \rho_{[j*]}) + 6\mathrm{Cov}(\rho_{[i*]} \overset{D+1}{\sim} \rho_{[j*]}) + 9\mathrm{Cov}(\rho_{[i*]} \overset{D+2}{\sim} \rho_{[j*]}), & \text{if } i \overset{D}{\sim} j, D > 1. \end{cases}$$
(3.17)

As $\mathrm{Cov}(\boldsymbol{\rho})$ is proportional to the CAR-type matrix $(\boldsymbol{M}_e - \gamma \boldsymbol{A}_e)^{-1}$ for all RE-NeGe models, then $\mathrm{Cov}(\rho_{[i*]} \overset{D}{\sim} \rho_{[j*]})$ decreases with the neighboring order n. Hence, $\mathrm{Cov}(\theta_i, \theta_j)$ decreases in our model. As a direct consequence of (3.17), the marginal correlation between the first-order neighbors θ_i and θ_j is given by

$$\mathrm{Corr}(\theta_i, \theta_j) = \frac{\mathrm{Var}(\rho_{[ij]}) + 6\mathrm{Cov}(\rho_{[i*]} \sim \rho_{[j*]}) + 9\mathrm{Cov}(\rho_{[i*]} \overset{2}{\sim} \rho_{[j*]})}{4\mathrm{Var}(\rho_{[ij]}) + 12\mathrm{Cov}(\rho_{[i*]} \sim \rho_{[j*]})}.$$
(3.18)

The RENeGe marginal and conditional correlations differ from that produced by the CAR model. The red arcs connecting nodes in Figures 4 (a), (b) and (c) show the first, second, and third-order (θ_i, θ_j) pairs of neighbors taking an inner row in a regular lattice with 400 nodes. The correlations are obtained assuming the RENeGe covariance matrix $\boldsymbol{C}(\boldsymbol{M}_e - \gamma \boldsymbol{A}_e)^{-1}\boldsymbol{C}^t$ with $\gamma = 0.8$. Under the CAR model, we assume $(\boldsymbol{M}_v - \varsigma \boldsymbol{A}_v)^{-1}$ with $\varsigma = 0.8$. Figures 4 (d), (e) and (f) show the values of $\mathrm{Corr}(\theta_i, \theta_j)$ under the RENeGe (blue dashed line) and the CAR (red solid line) models. Figures 4 (g), (h) and (i) show the conditional (or partial) correlation $\mathrm{Corr}(\theta_i, \theta_j | \boldsymbol{\theta}_{-(ij)})$, where $\boldsymbol{\theta}_{-(ij)}$ is the $(n-2)$-dimensional vector obtained by deleting the i-th and j-th coordinates from the n-dimensional vector $\boldsymbol{\theta}$. These conditional correlations were obtained numerically.

The grid analyzed in this example is not wrapped into a torus, so nodes belonging to the border of the lattice have a smaller number of neighbors. This produces a boundary
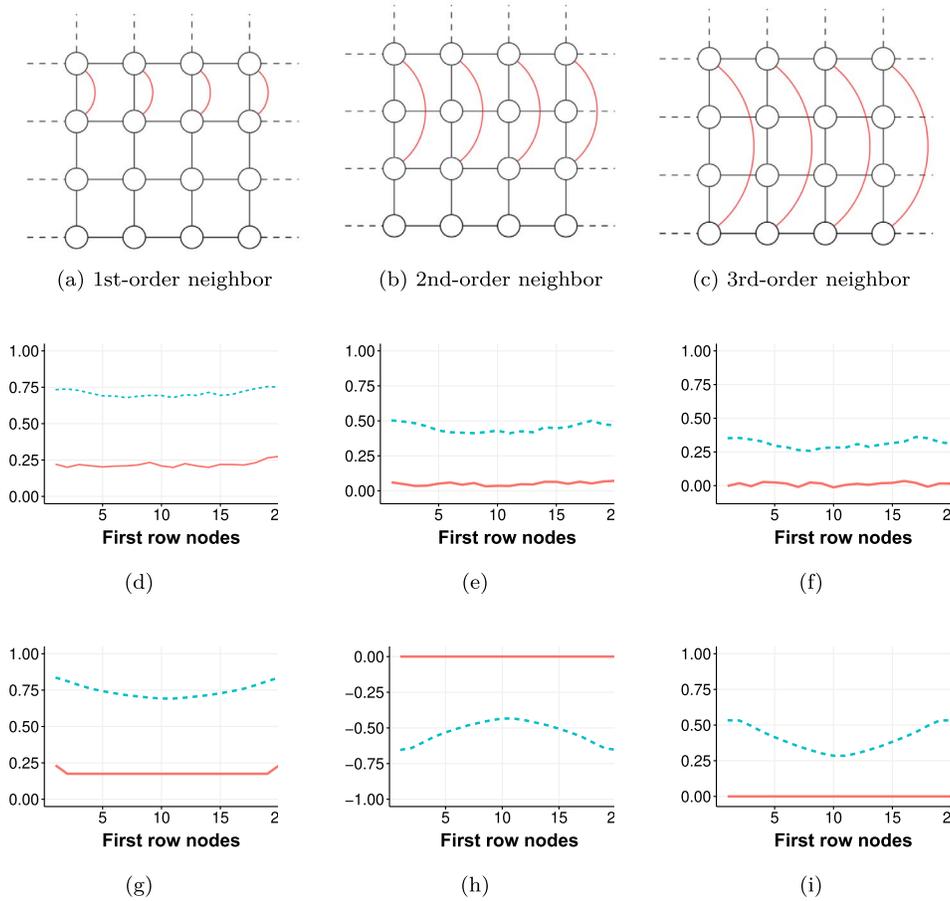
Figure 4: Marginal (plots (d), (e), and (f)) and conditional correlations (plots (g), (h), (i)) of first (d,g), second (e,h) and third (f,i) neighbor order for nodes $i = 1$ to 20, in a regular grid of order 20×20, under RENeGe (blue dashed line) and CAR (red solid line) models. The neighbor order is as indicated in the graphics on the first line.

effect indicating a higher conditional correlations near the boundaries when considering the first (Figures 4 (d) and (g)) and third (Figures 4 (f) and (i)) neighbor orders. The RENeGe model produces higher (in absolute value) marginal and conditional correlations between θ_i and θ_j than the CAR model, irrespective of their neighbor order. For first-order neighbors, the RENeGe marginal correlation in Figure 4 (d) is around 0.75, while under the CAR model, it is below 0.25. As expected, under both models, the marginal correlation is high for nodes near each other and decreases as the distance between the nodes increases. For instance, for 3rd-order neighbors in Figure 4 (f), the CAR model produces correlation approximately null while the RENeGe correlation is above 0.25.

(a) Conditional Correlation   (b) Time Series

(c) 1st order   (d) 2nd order   (e) 3rd order
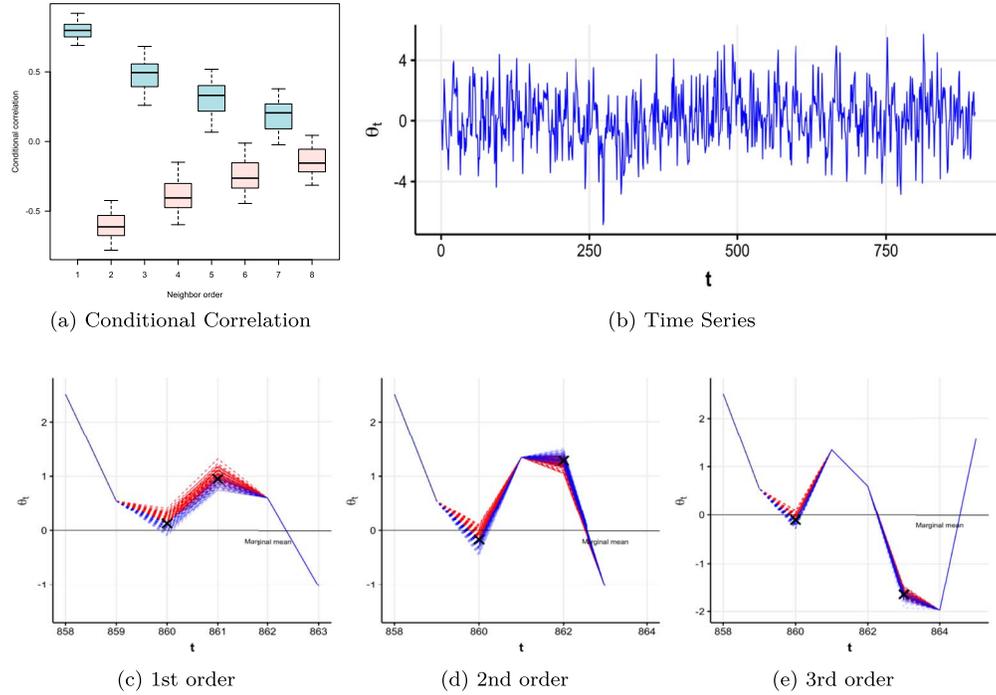
Figure 5: (a) Conditional correlations for different neighboring order of the RE-NeGe model in the 20×20 grid example. (b) RENeGe-based generated time series. The conditional expectations (black $\times$) and the generated values (dashed lines) removing observations that are neighbors of first (c), second (d) and third (e) orders.

The conditional correlations $\text{Corr}(\theta_i, \theta_j | \boldsymbol{\theta}_{-(ij)})$ in Figures 4 (g), (h), and (i) have a more intricate pattern. The CAR model shows the expected Markovian behavior. The low (smaller than 0.25) partial correlations of the first-order neighbors in Figure 4 (g) turn to zero in Figures 4 (h) and (i). The conditional correlations of the RENeGe model are more complex and seem puzzling. RENeGe produces a conditional correlation structure between θ_i and θ_j that is negative if these nodes are neighbors of even order (Figure 5 (a)). To understand this behavior, we simulated the RENeGe model in a one-dimensional situation, as we explain next.

Figure 5 (b) show a typical realization of a simulated RENeGe time series. Each θ_i is the sum of the antecedent and the subsequent edge effects: $\theta_i = \rho_{[i-1,i]} + \rho_{[i,i+1]}$. We assumed the parameter values $\gamma = 0.8$ and $\sigma_\theta^2 = 1$. We considered three different scenarios. In each of them, we deleted a pair (θ_i, θ_j) from the particular realization in Figure 5 (b) and kept all the other values fixed. Then we simulated a large number of (θ_i, θ_j) values conditioned on all the other values $\boldsymbol{\theta}_{-(ij)}$ in this time series. The three scenarios differ by the neighboring order of i and j. We used $i = 860$ and $j = 861, 862$, and 863. The bottom plots in Figure 5 show the simulated time series in each scenario considering only a small time window around the deleted (i, j) pair. The
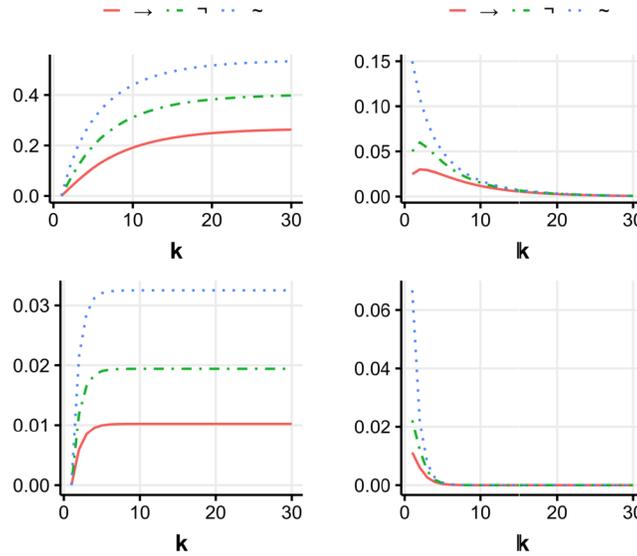
Figure 6: Cumulative summation $\sum_{j=1}^{k} \gamma^j \boldsymbol{C} \boldsymbol{W}_e^j \boldsymbol{M}_e^{-1} \boldsymbol{C}^t$ (left) and the values of $\boldsymbol{C}\gamma^k \boldsymbol{W}_e^k \boldsymbol{M}_e^{-1} \boldsymbol{C}^t$ (right) as a function of the neighboring order k and path types ($\sim$, $\rightarrow$, and $\neg$), for $\gamma = 0.9$ (top) and $\gamma = 0.4$ (bottom). ($i \sim j$: blue dotted line, $\rightarrow j$: red solid line, and $i\neg j$: green dot-dashed line).

black dots represent the conditional expectation of (θ_i, θ_j). Each line is a simulated series, and the values $\boldsymbol{\theta}_{-(ij)}$ are the same as in the original times series due to the conditioning. The randomness is associated only with the removed (θ_i, θ_j). Time series in which the generated θ_i is higher than the $E(\theta_i|\boldsymbol{\theta}_{-(ij)})$ are shown as red lines, and the others are shown in blue lines. Figure 5 (c) clearly shows the positive conditional correlation between the first-order neighbors θ_i and θ_{i+1}. The blue lines determined by $\theta_i < E(\theta_i|\boldsymbol{\theta}_{-(i,i+1)})$ typically are followed by $\theta_{i+1} < E(\theta_{i+1}|\boldsymbol{\theta}_{-(i,i+1)})$. This same pattern is observed for the third-order neighbors in Figure 5 (e). The negative correlation is evident between the second-order neighbors in Figure 5 (d). Indeed, typically, the blue lines determined by $\theta_i < E(\theta_i|\boldsymbol{\theta}_{-(i,i+2)})$ are followed by $\theta_{i+2} > E(\theta_{i+2}|\boldsymbol{\theta}_{-(i,i+2)})$.

Another interesting characteristic of the proposed model given by expression (3.16) is that, as γ is positive, the correlation increases monotonically with γ. The increasing rate differs depending on the paths connecting the nodes. Let $i \rightarrow j$ denote that nodes i and j are separated by a common neighbor and are in the same horizontal or vertical straight line in the grid. Denote by $i\neg j$ a pair of nodes separated by a single common neighbor but with a connecting path that is not a straight line.

Figure 6 shows the dependence on k of the terms $\gamma^k \boldsymbol{C} \boldsymbol{W}_e^k \boldsymbol{M}_e^{-1} \boldsymbol{C}^t$ in expression (3.15) and its cumulative sums $\sum_{j=1}^{k} \gamma^j \boldsymbol{C} \boldsymbol{W}_e^j \boldsymbol{M}_e^{-1} \boldsymbol{C}^t$, where k is the path order taking γ equal to 0.4 or $\gamma = 0.9$.

As this is a convergent series, the first terms in (3.16) are more relevant to obtain
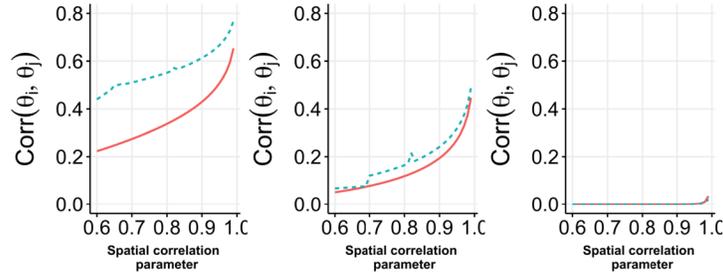
Figure 7: Correlation between first order (left), second order (middle), and third order (right) neighbors under the proposed (blue dashed line) and CAR (red solid line) models for different values of spatial correlation parameters γ and ς.

a good approximation for $\mathrm{Cov}(\theta_i, \theta_j)$. This relevance depends on the type of paths connecting the nodes. The higher-order neighbors terms in the $\mathrm{Cov}(\theta_i, \theta_j)$ expansion decay more slowly, requiring more of them in (3.16) to better approximate the true covariance. Hence, their influence increases with γ.

Figure 7 shows how the correlation between θ_i and θ_j changes for different values for γ and compares it with the $\mathrm{CAR}(\varsigma)$ model for different values for ς and different neighboring structures. In both models, the correlation increases with the γ and ς parameters. However, under RENeGe, the correlation between first-order neighbors is higher for any value of γ. This difference decreases as the neighboring order increases. Additional discussion of the correlation structure of the RENeGe model can be found in Section 2 in the Supplementary Material (Cruz et al., 2023).

4 Modeling spatial data using RENeGe

In this section, we empirically explore different features of the RENeGe model. First, we analyze its ability to induce a higher marginal correlation between neighboring areas than CAR models (Section 4.1). The eigenstructures of RENeGe and CAR models are explored in Section 4.2 following Hughes and Haran (2013) and Reich et al. (2006). In Section 4.3, we make a comparative analysis of five types of cancer mortality data using different versions of our RENeGe model and a set of spatial models previously proposed in the literature. The point estimates are the posterior mean in the case of Bayesian models. We also evaluate the performance of RENeGe, comparing it to some usual models considering simulated data and its capacity of recovering random effects (Section 4.4). In all our analysis, we fixed ℓ as the number of edges p plus 2 for RENeGe-TS model. A sensitivity analysis to the prior specification for the precision parameter $\sigma_{\boldsymbol{\theta}}^2$ and a toy example using RENeGe for image reconstruction can be found, respectively, in Sections 4 and 6 in the Supplementary Material (Cruz et al., 2023).

Under the RENeGe models, the posterior full conditional distributions have known closed forms for all parameters except for γ. To sample from the posterior distributions we propose a Gibbs sampler with a Metropolis step. The posterior full conditional

distributions are available in the Supplementary Material, Section 3, and the code is available on GitHub: https://github.com/DannaCruz/RENeGe-Code.

4.1 Inducing correlation between neighbors

One great limitation of the CAR model and its many variations is the small correlation this model induces between neighboring areas, even when a high value for the spatial parameter ς is assumed (Ver Hoef et al., 2017). To compare the marginal correlation induced by RENeGe and CAR models, we consider simulated datasets assuming that the areas are organized in a regular square lattice with $n = 15 \times 15$ nodes. The 225 nodes can be broken down into 169 nodes with four neighbors, 52 nodes with three neighbors, and 4 nodes with two neighbors, with a total of 420 edges.

Data are generated assuming that $Y_i \mid \boldsymbol{\theta}, \sigma_y^2 \overset{ind.}{\sim} N(\theta_i, \sigma_y^2)$, where $\sigma_y^2 \in \mathbb{R}_+$ is the variance parameter and $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_n) \in \mathbb{R}^n$. For RENeGe-N and RENeGe-T, we respectively generate the data assuming that $\boldsymbol{\theta} \sim N_n(\mathbf{0}; \boldsymbol{K})$ and $\boldsymbol{\theta} \sim T_n(\mathbf{0}; \boldsymbol{K}; \ell)$, where $\boldsymbol{K} = \boldsymbol{C}(\boldsymbol{M}_e - \gamma \boldsymbol{A}_e)^{-1} \boldsymbol{C}^t$. For the CAR model, we generate the data using $\boldsymbol{\theta} \sim N_n\left(\mathbf{0}, \sigma_{\boldsymbol{\theta}}^2 (\boldsymbol{D}_{\mathcal{G}} - \varsigma \boldsymbol{A}_{\mathcal{G}})^{-1}\right)$. We consider $\sigma_{\boldsymbol{\theta}}^2 = 1$ and choose two values for σ_y^2 (0.1 and 1.0). We select three high values for the spatial parameters γ and ς of the two models: $0.800, 0.900$, and 0.999.

We fit the models using the true values for the parameters σ_y^2, $\sigma_{\boldsymbol{\theta}}^2$, γ and ς. The RENeGe-T model is fitted fixing two values for ℓ. The first one is $\ell = 2.1$ to assume a heavy-tailed distribution for $\boldsymbol{\theta}$, and the second one is $\ell = 420$, the number of edges, which is a much larger value for this parameter in this model and provides an approximation to RENeGe-N. For RENeGe-TS, we assume $\ell = 420 + 2$, close to the minimum possible value for this parameter in this model, assuming that the degree of freedom is $\ell - n + 1 = 168$. We calculated the Moran's index $I = [\sum_{i=1}^n \sum_{j=1}^n a_{ij}(y_i - \bar{y})(y_j - \bar{y})]/[s_y^2 \sum_{i=1}^n \sum_{j=1}^n a_{ij}]$, in each simulated map, where s_y the sample standard deviation and a_{ij} is the element of the adjacency matrix defined by the torus-wrapped regular square lattice.

Table 1 shows the average value of the Moran's index calculated in each of 100 simulated datasets $\{y_i^{(k)}, i = 1, \ldots, n,$ and $k = 1, \ldots, 100\}$ for the four fitted models and shows the theoretical bounds limiting the possible values of this index. Table 1 presents also a different average. We randomly selected 100 pairs of neighboring areas. For each of those 100 pairs, we calculated the Pearson correlation index considering the 100 values generated. That is, for a pair composed of areas i and j, we calculated the Pearson correlation using the 100 values $(y_i^{(k)}, y_j^{(k)}), k = 1, \ldots, 100$. After this, we took the average of the Pearson correlation indexes for the selected 100 pairs of areas. For almost any value of the spatial parameters, RENeGe-N, RENeGe-T and RENeGe-ST induce a much higher correlation (Moran and Pearson) between the Y's of neighboring areas than the CAR model. If $\sigma_{\boldsymbol{\theta}}^2 = \sigma_y^2 = 1$, taking $\gamma = 0.9$ in RENeGe-N (respectively, RENeGe-T with $\ell = 2.1$), we have $I = 0.510$ (respectively, 0.550) while, in the CAR model, $I = 0.20$ even taking $\varsigma = 0.999$. If we take $\sigma_{\boldsymbol{\theta}}^2 = 1$ and $\sigma_y^2 = 0.1$, the Moran's I have higher values, but the previous conclusion is the same: RENeGe produces a higher

			Moran's Index			
Parameter	CAR	RENeGe-N	RENeGe-T	RENeGe-T	RENeGe-TS	Theoretical
γ or ς			$\ell = 2.1$	$\ell = 422$		Bounds
			$\sigma_{\boldsymbol{\theta}}^2 = 1,\ \sigma_y^2 = 1$			
0.8	0.07	0.260	0.290	0.269	0.270	$(-1.03, 1.02)$
0.9	0.06	0.510	0.550	0.510	0.557	$(-1.03, 1.02)$
0.999	0.20	0.590	0.620	0.590	0.584	$(-1.03, 1.02)$
			$\sigma_{\boldsymbol{\theta}}^2 = 1,\ \sigma_y^2 = 0.1$			
0.80	0.250	0.430	0.430	0.420	0.410	$(-1.03, 1.02)$
0.90	0.310	0.670	0.680	0.670	0.670	$(-1.03, 1.02)$
0.999	0.580	0.720	0.690	0.720	0.770	$(-1.03, 1.02)$
			Pearson Correlation			
			$\sigma_{\boldsymbol{\theta}}^2 = 1,\ \sigma_y^2 = 1$			
0.8	0.103	0.252	0.273	0.352	0.340	
0.9	0.135	0.603	0.586	0.623	0.535	
0.999	0.542	0.845	0.810	0.840	0.803	
			$\sigma_{\boldsymbol{\theta}}^2 = 1,\ \sigma_y^2 = 0.1$			
0.8	0.240	0.440	0.460	0.450	0.50	
0.9	0.340	0.730	0.710	0.730	0.730	
0.999	0.830	0.900	0.850	0.900	0.891	

Table 1: Average Moran's index and the average Pearson correlation index between pairs of areas under the CAR and RENeGe models for different values of γ and ς.

correlation than CAR. In general, the correlations induced by RENeGe-N, RENeGe-TS and RENeGe-T are comparable if we adopt $\ell = 422$ in RENeGe-T.

Although it is clear that RENeGe induces higher values for Moran's index than CAR, it is disappointing that the index does not come close to its maximum value even when the spatial parameter γ has a value very close to 1. The Moran's index reaches only 0.384 when $\gamma = 0.999$. A possible explanation for this behavior is that Moran's index I is a combination of a spatially smooth surface and the random noise. It is possible that conditionally specified spatial models may have some intrinsic limitation making it impossible to reach very high values for this marginal correlation. While the global Moran's I measures the overall clustering of the spatial data based on the entire map, the behavior of the pairwise correlation between neighboring areas given by the Pearson index shows a very similar pattern. In this case, maps generated by RENeGe show a much higher similarity of values between neighboring areas than those generated by CAR.

4.2 Comparison of RENeGe and CAR eigenstructure

Inspired by the eigenstructure analysis of spatial models in Hughes and Haran (2013) and Reich et al. (2006), we compare CAR, RENeGe-N, RENeGe-T and RENeGe-TS with respect to the covariance eigenstructure of the spatial effects $\boldsymbol{\theta}$ given, respectively, in (1.2), (3.8), (3.10) and in expression (1.8) in the Supplementary Material (Cruz
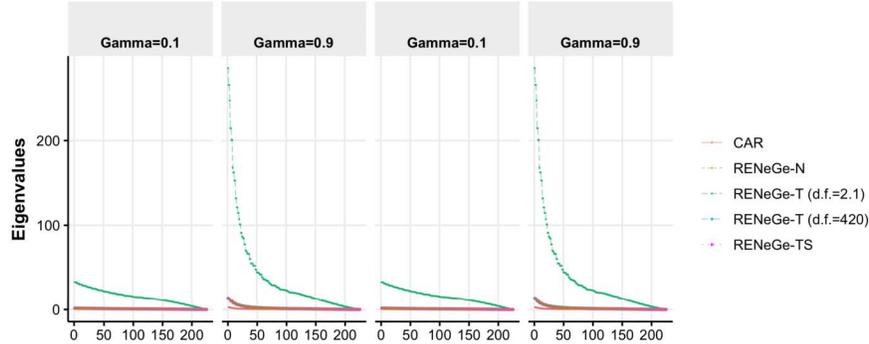
Figure 8: Eigenvalues for CAR and RENeGe models taking $\gamma = \varsigma$ and using $\gamma = 0.1$ and 0.9 for São Paulo State map (two graphics on the left) and a regular 15×15 lattice (two graphics on the right).

et al., 2023), and the respective eigenvalues assuming $\sigma_\theta = \tau_\theta = 1$. Two different spatial structures are considered: The São Paulo State map divided into 63 micro-regions and a regular square lattice with $n = 15 \times 15$ nodes. We assume $\gamma = \varsigma = 0.1$ and 0.9. For RENeGe-T, we consider two values for the degree of freedom: $\ell = 2.1$ and ℓ equal to the number of edges in the neighborhood graphs (157 for São Paulo State map and 420 for the regular lattice).

Figure 8 shows that, for both spatial structures analyzed, the eigenvalues associated with all models are positive but the first ones are much higher than the corresponding eigenvalues associated with the CAR model, under all RENeGe models.

The eigenvalues increase with γ, but this growth is much larger in the case of RE-NeGe than in the CAR model. Thus, compared to the CAR model, the fit for any RE-NeGe model depends on the first eigenvectors more strongly. This dependence is stronger if ℓ is small. As a reviewer pointed out, that may help explain why the neighbor-pair correlations are higher for RENeGe. Comparing the RENeGe models, the eigenvalues for RENeGe-N, RENeGe-TS and RENeGe-T with high value for ℓ are comparable and smaller than the eigenvalues for the heavy-tailed RENeGe-T which assumes $\ell = 2.1$. This behavior is expected since RENeGe-N is a limiting case of RENeGe-T when $\ell \to \infty$. This different eigenstructure between the models explains the behavior of the Moran's and Pearson correlation indexes shown in the previous section.

Figure 9 shows some eigenvectors related to the covariance matrix of RENeGe-N, RENeGe-T with $\ell = 2.1$ and CAR models assuming $\gamma = \varsigma = 0.9$ for São Paulo State map. The 1st (respectively, 2nd) and 4th (respectively, 5th) rows show, respectively, the first and last five eigenvectors maps for RENeGe-T model with $\ell = 2.1$ (respectively, RENeGe-N) while 3rd and 6th rows show the corresponding maps for the CAR model. Section 4 in the Supplementary Material (Cruz et al., 2023) shows the results when $\gamma = \varsigma = 0.1$ (Figure 2). It also shows (Figure 3) the eigenvector structures for a regular square lattice 15×15 by fitting RENeGe-T with $\ell = 2.1$ and CAR models also taking $\gamma = \varsigma = 0.9$ and 0.1.

(a) RENeGe-T ($\ell = 2.1$), first five eigenvectors

(b) RENeGe-N, first five eigenvectors

(c) CAR, first five eigenvectors

(d) RENeGe-T ($\ell = 2.1$), last five eigenvectors

(e) RENeGe-N, last five eigenvectors

(f) CAR, last five eigenvectors
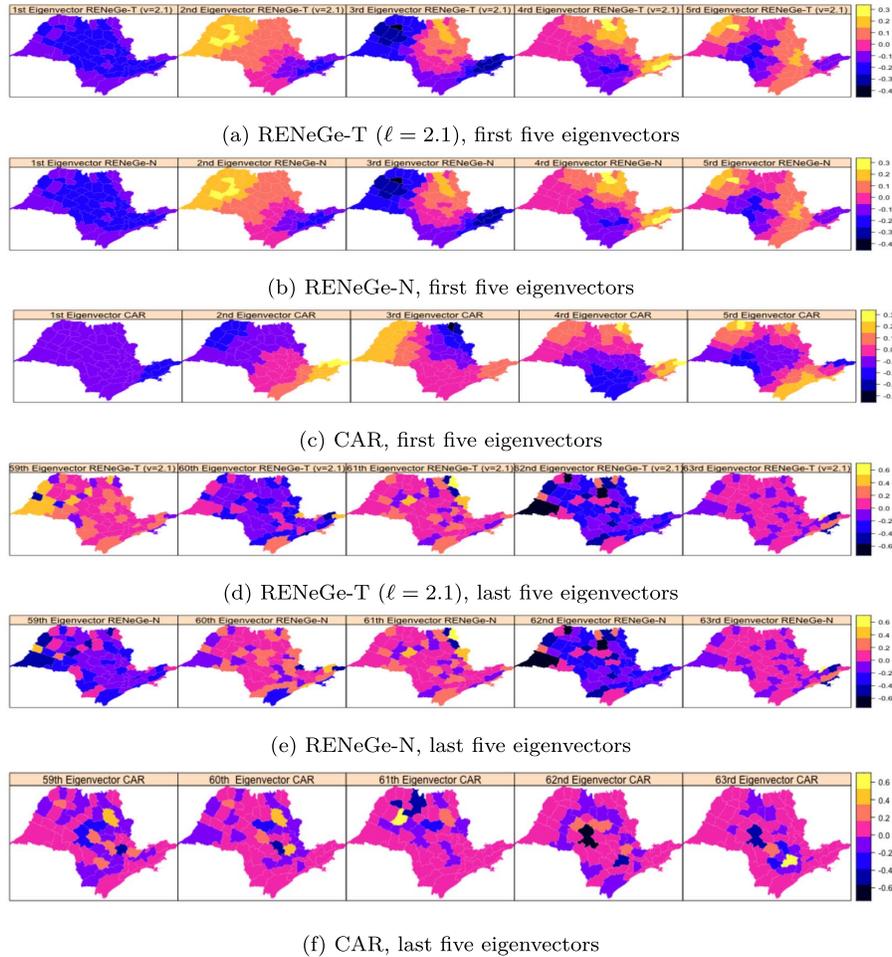
Figure 9: Eigenvectors for RENeGe-N, RENeGe-T with $\ell = 2.1$, and CAR models taking the spatial parameters equal to each other $\gamma = \varsigma = 0.9$ for São Paulo State map.

The eigenvector for RENeGe-N (2nd row) and RENeGe-T (1st row) with $\ell = 2.1$ are identical, which is expected since the covariance matrices for these models are proportional. Comparing RENeGe-T (1st row), RENeGe-N (2nd row) and CAR (3rd row) eigenvectors, there is a similar spatial pattern between the models, but there are also clear differences. Fixing $\gamma = \varsigma = 0.9$, the CAR patterns are much smoother than the RENeGe ones, especially in the first eigenvector. The spatial patterns in the last eigenvectors are more diffuse and do not show similarity. This is expected since these last eigenvectors are not relevant to explain the covariance matrix. Figure 2, Section 4 in the Supplementary Material (Cruz et al., 2023)) shows the spatial patterns in the eigenvectors using $\gamma = 0.1$. It shows that there is less spatial structure in the eigenvectors maps in all models if compared with the $\gamma = 0.9$ case, but this is more pronounced in the

CAR model. We do not show the last eigenvectors for $\gamma = 0.1$ because their eigenvalues are very close to zero but they have a similar spatial pattern as in the $\gamma = 0.9$.

The regular lattice removes the irregularities of the real map and displays the underlying spatial pattern in the eigenvectors more clearly (Figure 3, Section 4 in the Supplementary Material (Cruz et al., 2023)). The first five eigenvectors spatial patterns of RENeGe and CAR for $\gamma = \varsigma = 0.9$ are generally very similar. The largest difference is on the first eigenvector, where CAR has a more pronounced valley in the center while RENeGe has tiny clusters of small values in the four corners. The other four eigenvectors also show small differences due to border effects in RENeGe. Moving to the $\gamma = \varsigma = 0.1$ case, the spatial pattern is flatter, especially in the CAR model. The five last eigenvector maps show Scottish-like textures with alternating small clumps of different colors.

4.3   Case study: cancer maps

We analyze the spatial pattern of deaths caused by five types of cancer: Lung/Bronchial, Colon/Rectal, Stomach, female Breast cancer, and male Prostate cancer. They are selected because they are the most important causes of death by cancer. We collect the total number of deaths that occurred in the 2008–2019 periods in areas from four states located in the south of Brazil: Rio Grande do Sul, Santa Catarina, Paraná, and São Paulo. This region has 73 million inhabitants (around 35% of the Brazilian population), and it is partitioned into $n = 159$ administrative areas or micro-regions. In this region, data has good quality, having little underreporting or cause of death misreporting problems. Cancer and population data were collected from the DATASUS (http://datasus.saude.gov.br/), the official Brazilian Health Department data repository.

As a covariate, we consider the Municipal Human Development Index (MHDI), a composite index measuring social deprivation calculated by the United Nations. These data were obtained from https://www.br.undp.org/. MHDI is a combination of three indicators capturing different dimensions of human development: longevity, education, and economic well-being. The index varies from 0 to 1. The closer to 1, the greater the human development. The MHDI in the southern region of Brazil has a clustered spatial pattern, as shown in Figure 4, Section 4 in the Supplementary Material (Cruz et al., 2023).

Let Y_i and E_i denote the observed and the expected counts of death by cancer at micro-region i, respectively. The expected value E_i is calculated using the age-gender population distribution and assuming that the age-specific risk is constant in the entire map. The data are analyzed assuming that

$$Y_i|MHDI_i, \theta_i, \beta_0, \beta \overset{ind}{\sim} \text{Poisson}(E_i\lambda_i), \quad i = 1, \ldots, 159. \tag{4.1}$$

The λ_i follows the log-linear regression structure such that $\log(E_i\lambda_i) = \beta_0 + \beta MHDI_i + \theta_i + \log(E_i)$, where θ_i is the spatial random effects at area i.

To fit RENeGe-N, RENeGe-T, and RENeGe-TS models, we used the following vague prior distributions: $\beta \overset{D}{=} \beta_0 \sim N(0, 10^2)$; $\gamma \sim U(0, 1)$ and $\sigma_\theta^2 \sim IG(0.001, 0.001)$. Additionally, for RENeGe-T, we choose $\ell \sim \text{Gamma}(2, 1/10)$, and for RENeGe-TS, we fixed $\ell = p + 2 = 424$ and the number of degrees of freedom is 266.

We compare these proposed models with the plain CAR model (Besag et al., 1995), the generalized skew-normal (GSN) model (Prates et al., 2012), the Leroux model (Leroux et al., 2000), the BYM model (Besag et al., 1991), the model introduced by (Rodrigues and Assunção, 2012) that extends the BYM and Leroux models by considering a higher-neighborhood dependence (HND) and the DAGAR model (Datta et al., 2019). CAR and Leroux models use a Gaussian Markov random field (GMRF) assuming $\boldsymbol{\theta} \sim N(\mathbf{0}, \sigma_{\boldsymbol{\theta}}^2 \boldsymbol{Q}(\varsigma)^{-1})$ where the precision matrix $\boldsymbol{Q}(\varsigma)$ is a function of the spatial correlation parameter ς specific for each model. We choose a prior $\varsigma \sim U(0,1)$. For BYM, the spatial component was an ICAR model with an additional unstructured iid component. The two variance components are $IG(1, 0.01)$, as recommended in the CARBayes R package. For the HND model, we consider the particular model of three components $\boldsymbol{Q}^{-1} = \sigma_\theta^2(\lambda_1 \boldsymbol{I} + \lambda_2 \boldsymbol{W}_1 + \lambda_3 \boldsymbol{W}_2)$. We adopt a gamma distribution with parameters equal to 0.01 for all inverse variance parameters and a uniform distribution in the two-dimensional simplex for the weights $(\lambda_1, \lambda_2, \lambda_3)$. The prior distributions of parameters β_0, β, and $\sigma_{\boldsymbol{\theta}}^2$ in these four models are the same as those considered in RENeGe. In the DAGAR model, the $\boldsymbol{Q}(\varsigma)$ corresponds to an autoregressive model induced by a directed acyclic graph. We used the order based on the sum of the coordinates of the vertex assignments to define the directed graph. Additionally, we choose $\beta \stackrel{D}{=} \beta_0 \sim N(0, 10^2)$, $\varsigma \sim U(0,1)$, and $\sigma_{\boldsymbol{\theta}}^2 \sim IG(2,1)$. We use the Nimble R-package (de Valpine et al., 2021) to fit the model. For the GSN model, the random vector $\boldsymbol{\theta}$ follows the generalized skew-Gaussian spatial field (GSGSF) $\boldsymbol{\theta} \sim GSNI_n(\sqrt{2/\pi}\kappa\lambda\mathbf{1}_n, \boldsymbol{\Sigma}, \gamma\boldsymbol{I}, H_n(\cdot; \ell))$ with a CAR structure for $\boldsymbol{\Sigma}$, $\kappa = E[U^{-1/2}]$ and U is a positive random variable with distribution $H(\cdot; \ell)$. Different specifications of $H(\cdot; \ell)$ define different GSGSF (see Prates et al. (2012) for discussion). In our applications, $H_n(\cdot; \ell)) \equiv \mathrm{Gamma}(\ell/2, \ell/2)$ where $\ell = 0.1$, $\lambda \sim N(0, 10)$, $\sigma_{\boldsymbol{\theta}}^2 \sim IG(0.001, 0.001)$, and the spatial dependence parameter $\varsigma \sim U(-1, 1)$. We also use the Nimble package from R to fit this last model (de Valpine et al. (2021)).

To compare the models, we consider the following model selection criteria: the Deviance Information Criteria (DIC, Spiegelhalter et al. (2002)), the Extended Bayesian Information Criterion (EBIC, Brooks et al. (2002)), the extended Akaike information criterion (EAIC, Carlin and Louis (1996)), and the Watanabe-Akaike information criterion(WAIC, Watanabe (2010)). Small values for the DIC, EBIC, EAIC, and WAIC imply a better fit of the model. DIC, EAIC and EBIC are given, respectively, by

$$D\hat{I}C = \bar{D} + \bar{\rho_D}, \quad E\hat{B}IC = \bar{D} + p\log(n), \quad \text{and} \quad E\hat{A}IC = \bar{D} + 2p,$$

where $\bar{D} = E(D(\boldsymbol{\Psi})|\boldsymbol{y})$, $D(\boldsymbol{\Psi}) = -2\sum_{i=1}^n \log(f(y_i|\boldsymbol{\Psi}))$ is the deviance, $f(y_i|\boldsymbol{\Psi})$ is the conditional distribution of y_i given $\boldsymbol{\Psi}$, p is the number of parameters in the model, n is the total number of observations, and Ψ denotes all model parameters. The effective number of parameters ρ_D (Spiegelhalter et al., 2002) is given by $\rho_D = E(D(\boldsymbol{\Psi})|\boldsymbol{y}) - D(E(\boldsymbol{\Psi}|\boldsymbol{y}))$. Estimates for $D(E(\boldsymbol{\Psi}|\boldsymbol{y}))$ and $\bar{D}$ are easily obtained using the MCMC iterations. The WAIC is an extension of AIC and is given by$-2(lppd - pWAIC)$, where $lppd$ is the log point-wise predictive density and $pWAIC$ is the effective number of parameters (see Gelman et al., 2004, for details).

To assess the predictive accuracy of the models, we removed $n^* = 31$ adjacent microregions from the map to obtain a connected graph representing the remaining areas.

	RENeGe								
	TS	T	N	CAR	GSN	Leroux	BYM	HND*	DAGAR*
Lung/Bronchial Cancer	470.54	26.81	12.21	13.34	142.52	12.21	9.85	5.33	5.48
Breast cancer	470.42	28.12	13.91	14.06	142.23	12.51	10.18	5.11	5.11
Colon/rectal cancer	496.90	26.41	13.51	13.30	141.62	11.8	9.802	9.82	4.5
Stomach cancer	470.70	27.83	12.82	13.77	141.79	11.97	10.41	5.24	5.91
Prostate cancer	470.46	26.95	12.92	13.59	142.16	12.64	9.92	4.94	5.72

Table 2: Computational time (in seconds) for all fitted models. The asterisk symbol indicates a model implemented in the NIMBLE language.

We considered 5 sets of removed areas, varying their location in the map to account for the different spatial features in the risk structure. The removed areas in each of the 5 sets are shown in Figure 5, Section 4, in the Supplementary Material (Cruz et al., 2023). The RENeGe and the other models are fitted to each set of fully observed data. Next, we generated $m = 50$ values $\hat{Y}$ from the posterior predictive distribution for the omitted Y values and compared Y and $\hat{Y}$ using the following BIAS and Root Mean Squared Error (RMSE) metrics:

$$BIAS = \sum_{j=1}^{5} \sum_{i=1}^{n^*} \sum_{k=1}^{m} (\hat{Y}_{ijk} - Y_i)/5mn^* \qquad RMSE = \sqrt{\sum_{j=1}^{5} \sum_{i=1}^{n^*} \sum_{k=1}^{m} (\hat{Y}_{ijk} - Y_i)^2/5mn^*},$$

where Y_i is the observed count in area i and $\hat{Y}_{ijk}$ is a random value from the posterior predictive distribution at area i in replication k for j-th set of removed areas.

Table 2 shows the computational time needed to obtain the posterior samples under all fitted models. It shows that the times to run RENeGe-T and RENeGe-N are close to that of the competitor models. However, the RENeGe-TS model does not come without a cost. The more complex model structure requires more computing time. In part, this extra computational time is due to our non-optimized implementation of RENeGe model. We should also emphasize that the comparison with HND and DAGAR is not fair, as they were implemented in NIMBLE.

Although no model is unanimously indicated as the best by the adopted selection criteria, the RENeGe class of models out performs the competitor models (Table 3). Only RENeGe-TS is selected, by some criterion, as the best model for all types of cancers. RENeGe-TS fits the data better of all kinds of cancers according to EBIC and EAIC, except for Stomach cancer, where EBIC and EAIC select the RENeGe-N model. DIC also chooses a RENeGe model to fit the data, except for Lung/Bronchial cancer where the BYM model is selected as the best. WAIC is the only criterion that never selected a RENeGe model as the best. It is also important to mention that, for most of the cancers, the RENeGe-TS model produces the most biased predictions. In general, DAGAR and GSN are the models with the poorest performance if we use DIC, EBIC and EAIC criteria.

Figures 10 and 11 respectively show the relative risk (λ_i) and spatial random effect (θ_i) estimates for area i, $i = 1, \ldots, n = 159$, of the two most prevalent types of cancers, female Breast cancer (bottom) and Lung/Bronchial cancer (top), under all fitted models. Results for other cancers can be found in Section 4 of the Supplementary Material (Cruz et al., 2023).

Although Table 3 shows differences among the models, these differences cannot be clearly seen in Figure 10. CAR, GSN, Leroux, BYM, HND, DAGAR, and the RE-NeGe models provide very similar maps for the relative risk of Lung/Bronchial cancer mortality, and the estimates are comparable to the naive SMR estimate. Lung cancer is the most prevalent cancer, and its large numbers allow for stable rates that are not much different from the Bayesian smoothed rates for this cancer. The relative risks (λ_i) are grouped in relatively large clusters indicating that neighboring areas tend to have similar risk. All these models, except for the GLM, indicate a very clear North-South gradient, with lung cancer risk increasing as we move toward the south. This gradient is the most distinctive spatial pattern in this map, with small differences across the East-West direction. A similar clustering structure was identified in lung cancer data in the South region of Brazil by applying the Bayesian model for spatio-temporal cluster detection introduced by Teixeira et al. (2019). Disturbing the smooth trend along the North-South gradient in this map, we have three high-risk spatial clusters in the South. Two of them are on the border between Brazil and Uruguay, while the third contains Porto Alegre, the largest city in the South region. The relative risks in these clusters are around 2, meaning that they have a risk twice as large as the average risk in the entire region. The two clusters on the border are also detected by RENeGe-TS, but the estimates for the relative risks in these areas are smaller than in the other RENeGe models. Another aspect is that the spatial pattern in the spatial effect $\hat{\theta}_i$ is very similar to that found in the relative risk $\hat{\lambda}_i$ for the majority of the models. The smooth North-South gradient is almost the same in both maps, being the spatial effect positively (negatively) high for regions experiencing the highest (smallest) mortality risk.

The estimates for the relative mortality risk of breast cancer are spatially more heterogeneous than the lung cancer risk. Except for the GLM, all other models provide visually similar estimates for the mortality risk. The maps do not show any striking differences between the spatial models. In fact, they look identical, and differences between the models must be ascertained through the model selection metrics from Table 3. Returning to the breast cancer maps, they point to the same four spatially unconnected areas with the highest relative risk, around twice the average risk. The maps of $\hat{\theta}_i$ and $\hat{\lambda}_i$ for breast cancer have practically the same spatial pattern.

This similarity between the spatial patterns of the relative risks and the spatial effects for both type of cancers is an indication that the quality-of-life index, as measured by MHDI, has little correlation with Lung/Bronchial and Breast cancers mortality risks. In fact, this is reinforced by the results in Table 4, which shows the posterior means and the 95% highest posterior density (HPD) intervals. Using a non-spatial GLM Poisson regression model, we find that the effect of the covariate MHDI is significant for both cancers (Table 4). It is positive for breast cancer but negative for lung cancer. There is a possible explanation for this. Lung cancer occurs mainly due to cigarette consumption,

Model	DIC	EBIC	EAIC	WAIC	RMSE	BIAS
			Lung/Bronchial Cancer			
RENeGe-TS	1625.51	**1346.7**	**1331.15**	2021.68	6.40	0.15
RENeGe-T	1635.92	1367.15	1351.59	2023.13	**6.21**	0.02
RENeGe-N	1623.91	1356.95	1341.39	2021.92	6.34	0.16
CAR	1514.37	2204.80	1684.06	**1457.24**	6.36	**0.01**
GSN	1663.19	2542.00	1879.18	1563.84	6.86	0.05
Leroux	1509.94	2208.67	1681.66	1540.59	8.69	0.05
BYM	**1508.60**	2173.84	1672.09	1457.52	8.63	0.02
HND	1561.27	2440.08	1777.25	1461.92	8.31	0.05
DAGAR	1599.71	2212.58	1750.34	1558.77	8.68	0.04
			Breast cancer			
RENeGe-TS	1381.52	**1105.27**	**1089.71**	1508.79	2.86	−0.03
RENeGe-T	**1380.98**	1120.66	1105.10	1523.73	4.14	0.07
RENeGe-N	1382.21	1115.07	1099.51	1523.26	3.06	−0.08
CAR	1388.86	2042.80	1549.58	1335.89	2.91	**0.02**
GSN	1492.47	2146.58	1653.22	1442.46	2.89	**0.02**
Leroux	1390.25	2025.94	1546.48	1344.21	**2.40**	0.03
BYM	1390.10	2053.67	1553.18	**1335.38**	2.92	−0.03
HND	1390.54	2044.65	1551.30	1340.53	3.98	**-0.02**
DAGAR	1483.81	2117.22	1639.48	1437.95	3.40	**0.02**
			Colon/Rectal cancer			
RENeGe-TS	1533.92	**1248.96**	**1233.41**	1819.81	**2.84**	0.06
RENeGe-T	1532.73	1276.29	1260.73	1829.38	8.99	**0.00**
RENeGe-N	**1525.79**	1267.29	1251.74	1829.08	7.39	−0.06
CAR	1535.04	2181.42	1693.90	**1481.71**	6.84	0.05
GSN	1644.51	2279.93	1800.68	1606.46	7.48	0.05
Leroux	1609.60	2459.86	1818.57	1529.40	7.41	0.01
BYM	1540.23	2207.45	1704.21	1482.29	7.02	0.07
HND	1542.59	2178.01	1698.75	1504.54	6.93	0.07
DAGAR	1711.52	2561.79	1920.49	1631.33	6.92	0.05
			Stomach cancer			
RENeGe-TS	**1496.10**	1217.30	1201.74	1783.44	**2.39**	0.23
RENeGe-T	1508.02	1215.21	1199.65	1783.45	2.49	**0.01**
RENeGe-N	1498.87	**1214.78**	**1199.23**	1783.69	**2.39**	0.17
CAR	1514.37	2204.80	1684.06	1457.24	4.25	0.18
GSN	1663.19	2542.00	1879.18	1563.84	5.27	−0.04
Leroux	1509.94	2208.67	1681.66	**1440.59**	5.16	0.11
BYM	1508.60	2173.84	1672.09	1457.52	3.85	**0.01**
HND	1561.27	2440.08	1777.25	1461.92	3.88	−0.11
DAGAR	1711.52	2561.79	1920.49	1631.33	3.88	−0.03
			Prostate cancer			
RENeGe-TS	1523.99	**1236.51**	**1220.96**	1792.16	**2.15**	0.04
RENeGe-T	**1514.66**	1243.14	1227.58	1790.56	2.53	−0.03
RENeGe-N	1520.39	1238.92	1223.36	1792.64	2.42	−0.04
CAR	2411.65	2216.26	1696.49	1468.43	5.20	**−0.01**
GSN	2511.65	1329.18	3548.18	1725.84	4.95	**−0.01**
Leroux	1530.53	2296.37	1718.75	**1431.95**	4.08	0.02
BYM	1526.38	2200.23	1691.99	1476.42	4.01	**−0.01**
HND	2409.72	1327.26	3446.25	1623.92	4.58	0.03
DAGAR	1632.45	2398.29	1820.67	1533.88	4.05	**−0.01**

Table 3: Model selection criteria for all fitted models.
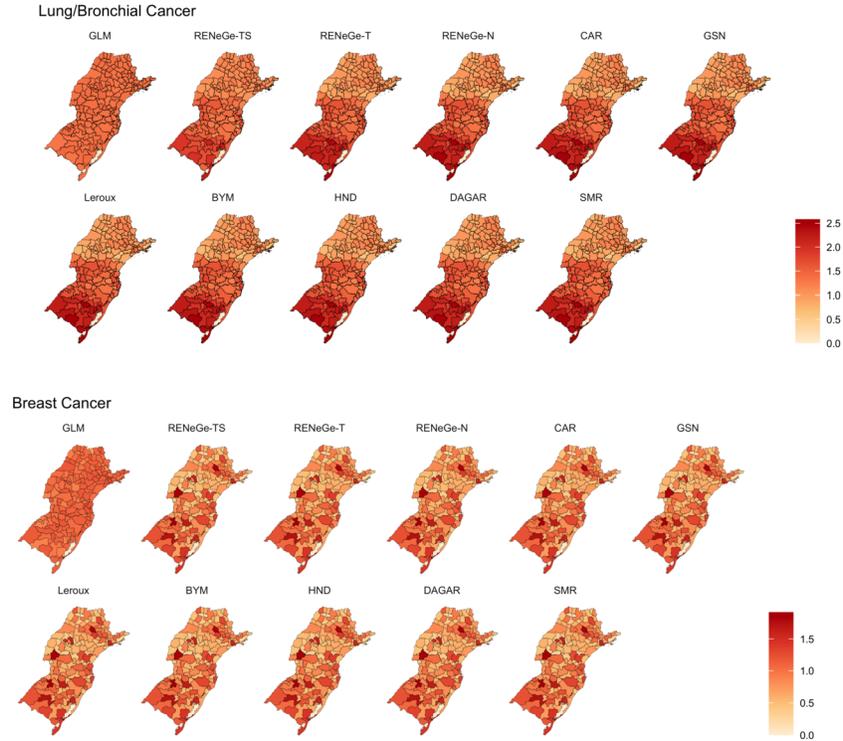
Figure 10: Posterior means for the relative risk $\boldsymbol{\lambda}$ of Lung/Bronchial (left) and Breast (right) cancer mortality in the southern region of Brazil under all fitted models.

which used to be higher in Brazil in those areas with higher income levels and therefore, higher MHDI. Breast cancer mortality may be related to the lack of preventive exams, a shortage that is more intense in poorer areas. However, the HPD intervals related to the covariate coefficients include zero for most spatial models. The exceptions occur for the RENeGe-T, GSN and HND models when analyzing Lung/Bronchial cancer data. That is, as soon as we allow for random spatial effects, the covariate is no longer relevant. This puzzling result may be caused by the confounding between the strongly spatially patterned covariate and the unobserved random spatial effects. Several recent papers have been dedicated to this thorny issue in spatial statistics (Hodges and Reich, 2010; Hughes and Haran, 2013; Hanks et al., 2015; Prates et al., 2019; Nobre et al., 2021). Although we do not pursue this issue further in this paper, some kind of controlling for spatial confounding may be necessary to analyze these cancers in Brazil if the MHDI covariate is used.

In Section 4 in the Supplementary Material (Cruz et al., 2023), we present the posterior means and the 95% HPD intervals for the spatial random effects of 32 areas selected as follows: 10 areas with the lowest SMR (bottom), 11 areas with intermediate SMR, and 11 areas with the highest SMR (top). For both types of cancers, the higher the

	Mean	95% HPD	Mean	95% HPD	Mean	95% HPD	Mean	95% HPD	Mean	95% HPD
				Lung/Brochial Cancer						
Coefficients	RENeGe-TS		RENeGe-T		RENeGe-N		CAR		GSN	
Intercept	0.13	(0.08,0.27)	0.51	(0.42, 0.57)	0.12	(−0.04,0.29)	0.17	(0.03,0.30)	0.61	(0.45, 0.76)
MHDI	0.16	(−0.10,0.40))	0.17	(0.01,0.24)	0.16	(−0.10,0.40)	0.10	(−0.11,0.31)	−0.62	(−0.85,-0.37)
σ_θ^2	0.07	(0.03,0.12)	0.01	(0.00, 0.01)	0.07	(0.03,0.12)	0.10	(0.08,0.12)	0.00	(0.00,0.01)
γ/ς	0.51	(0.04,0.96)	0.51	(0.04,0.96)	0.51	(0.04,0.96)	1.00	(0.98,1.00)	0.06	(0.06,0.06)
ℓ			148.66	(145.20,153.80)						
	Leroux		BYM		HND		DAGAR		GLM	
Intercept	0.10	(−0.04,0.29)	0.17	(−0.07,0.35)	0.61	(0.45, 0.76)	0.17	(−0.07,0.35)	0.53	(0.43,0.63)
MHDI	0.19	(−0.09,0.42)	0.10	(−0.18,0.47)	−0.62	(−0.85,-0.37)	0.10	(−0.18,0.47)	−0.35	(−0.45 ,-0.25)
σ_θ^2	0.10	(0.07,0.12)	0.09	(0.07,0.12)	0.00	(0.00,0.01)	0.09	(0.07,0.12)		
γ/ς	0.97	(0.91,1.00)	0.00	(0.00,0.01)	0.06	(0.06,0.06)	0.00	(0.00,0.01)		
				Breast cancer						
Coefficients	RENeGe-TS		RENeGe-T		RENeGe-N		CAR		GSN	
Intercept	−0.48	(−0.91,-0.18)	−0.49	(−0.75,-0.16)	−0.47	(−0.91,-0.18)	−0.68	(−1.34,-0.26)	−0.29	(−0.49,-0.08)
MHDI	0.24	(−0.20, 0.91)	0.27	(−0.22,0.69)	0.24	(−0.20,0.91)	0.60	(−0.05, 1.60)	−0.03	(−0.36,0.28)
σ_θ^2	0.44	(0.32,0.59)	0.40	(0.28,0.60)	0.44	(0.32,0.59)	1.39	(1.10,1.77)	0.00	(0.00,0.00)
γ/ς	0.39	(0.02,0.95)	0.39	(0.02,0.95)	0.39	(0.02,0.95)	0.28	(0.03,0.66)	0.06	(0.06,0.06)
ℓ			113.91	(101.05,145.26)						
	Leroux		BYM		HND		DAGAR		GLM	
Intercept	−0.45	(−1.06,0.23)	−0.52	(−0.94,-0.13)	−0.29	(−0.49,-0.08)	−0.52	(−0.94,-0.13)	−0.60	(−0.70 , −0.50)
MHDI	0.23	(−0.81,1.18)	0.34	(−0.25,1.01)	−0.03	(−0.36,0.28)	0.34	(−0.25, 1.01)	1.04	(0.94,1.14)
σ_θ^2	0.41	(0.27,0.71)	0.16	(0.05,0.35)	0.00	(0.00,0.00)	0.16	(0.05,0.35)		
γ/ς	0.13	(0.02,0.38)	0.21	(0.14,0.28)	0.06	(0.06,0.06)	0.21	(0.14,0.28)		

Table 4: Posterior means and 95% HPD intervals for some model parameters under all models, Lung/Bronchial and Breast cancers.
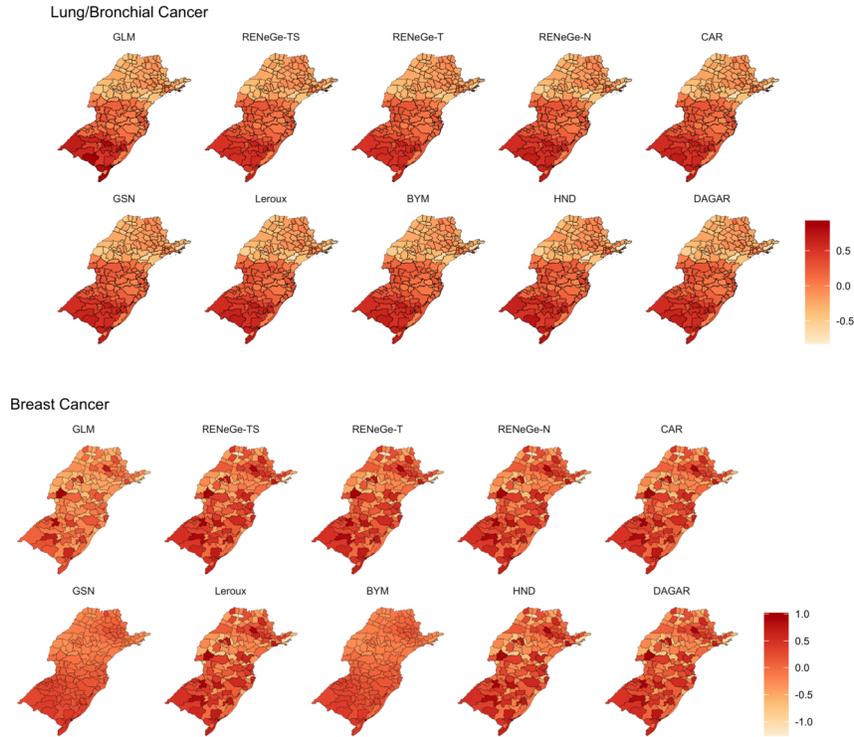
Figure 11: Posterior means for the spatial effects $\boldsymbol{\theta}$ in Lung/Bronchial (left) and Breast (right) data in the southern region of Brazil under all fitted models.

SMR, the higher the spatial effect. Areas with the intermediate and low values for the SMR tend to experience negative spatial effects. These areal effects approach zero for areas with intermediate SMR in both types of cancers. In general, the uncertainty about the random effects is higher in areas with small SMR. This is more clearly perceived for Breast cancer data. There is little difference between the methods and no clear pattern for interval lengths between the methods. For instance, RENeGe-T has the largest length for Rio Negro and Auriflama, but it is the CAR model that provides the largest length for São Paulo and Cascavel. If we compare RENeGe models, RENeGe-T generates intervals with a greater range than RENeGe-N, but for Cerro Azul, the opposite occurs.

4.4 Analyzing simulated data

Do RENeGe models recover the spatial effects θ? To investigate this question, we simulate a data set using the map of São Paulo State in Brazil. This map is composed of 63 micro-regions (nodes in the neighborhood graph) and a total of 157 edges connecting the adjacent micro-regions. For this graph structure, we generate a vector
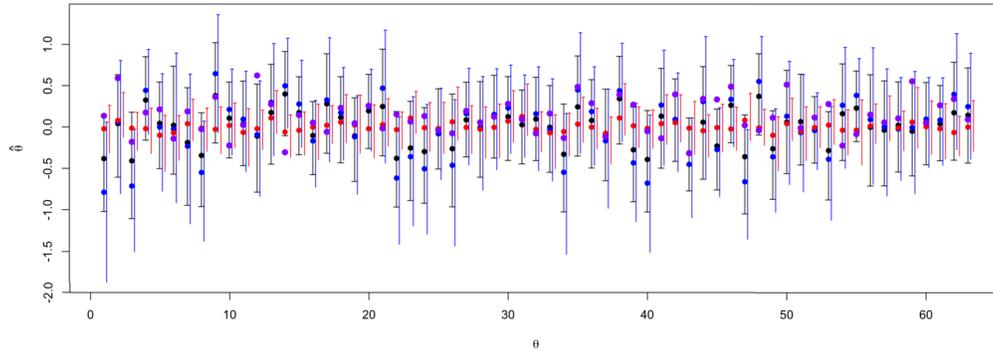
Figure 12: Posterior means (circles) and 95% HPD intervals for $\hat{\boldsymbol{\theta}}$ under RENeGe-N(blue), RENeGe-T(black), and RENeGe-TS(red). The purple circles are the true values $\boldsymbol{\theta}$.

$\boldsymbol{\theta}$ with dimension $n = 63$ using RENeGe-N and RENeGe-T, respectively, using the distributions: $N_n(\mathbf{0}; \boldsymbol{K})$ and $T_n(\mathbf{0}; \boldsymbol{K}; \ell)$, where $\boldsymbol{K} = \boldsymbol{C}(\boldsymbol{M}_e - \gamma \boldsymbol{A}_e)^{-1}\boldsymbol{C}^t$ with $\gamma = 0.5$, $\sigma_{\boldsymbol{\theta}}^2 = 0.05$ and $\ell = 159$. For RENeGe-TS, we generate a vector $\boldsymbol{\rho}$, with dimension $p = 157$ as follows. A matrix $\boldsymbol{S}$ is generated from the Inverse-Wishart distribution $\boldsymbol{S} \sim IW_p(\ell, (\boldsymbol{M}_e - \gamma \boldsymbol{A}_e)^{-1})$, with $\gamma = 0.5$. The vector $\boldsymbol{\rho}$ is thus generated from the multivariate normal distribution $\boldsymbol{\rho}|\boldsymbol{S} \sim N(\mathbf{0}, (\ell-p-1)\sigma_{\boldsymbol{\theta}}^2\boldsymbol{S})$ with $\ell = 159$ and $\sigma_{\boldsymbol{\theta}}^2 = 0.05$. Each coordinate of vector $\boldsymbol{\theta}$ is obtained as a linear combination of the incident edges in that node, as defined in Section 3.1. The response Y_i is generated from a Poisson distribution with mean $\exp\{0.1X_i + \theta_i\}$, and the X_i covariate is independently generated from a $N(0, 1)$ distribution.

To analyze the data, we fit all the RENeGe models assuming $\beta \sim N(0, 10^2)$, $\gamma \sim$ Uniform$(0, 1)$, and $\sigma_{\boldsymbol{\theta}}^2 \sim IG(0.001, 0.001)$. For RENeGe-T and RENeGe-TS, the parameter ℓ is fixed in the true value. A total of 10000 MCMC iterations are generated, and the first 1000 are discarded as burn-in.

The posterior means (respectively, the 95% HPD intervals) for $\sigma_{\boldsymbol{\theta}}^2$ and γ are, respectively, 0.04 ([0.01, 0.1]) and 0.40 ([0.05, 0.65]) under the RENeGe-N model, 0.05 ([0.04, 0.27]) and 0.62 ([0.10, 0.72]) under the RENeGe-T model, and 0.04 ([0.01, 0.08]) and 0.52 ([0.03, 0.86]) under the RENeGe-TS model. Such posterior summaries for the spatial effects $\boldsymbol{\theta}$ are given in Figure 12. This simple example shows that all RENeGe models provide reasonable estimates for all parameters and recover the true spatial effects in almost all microregions.

Comparing the models: To evaluate the performance of the RENeGe models and compare them with the performance of the models considered in Section 4.3, we simulate 12 scenarios with different spatial complexity and variability among the spatial effects $\boldsymbol{\theta}$. To define the spatial structure of the simulated data (coordinates, shapes, and spatial adjacency), we consider the geographical neighborhood structures of the southern Brazilian regions presented in Section 4.3. Considering this map, our neighboring graph

is composed of $n = 159$ nodes connected by $p = 422$ edges. We considered four different spatial patterns or scenarios.

The models considered were: RENeGe-TS, RENeGe-T, RENeGe-N, CAR, GSN, Leroux, BYM, HND, and DAGAR. The scenarios as well as additional information on the models specification and detailed results are presented in Section 5 in the Supplementary Material (Cruz et al., 2023). Based on the results shown in this Supplementary Material, we conclude that no model is unanimously chosen as the best by the criteria used. However, the RENeGe model is deemed the best by at least one criterion in all scenarios.

5   Conclusions

The CAR model, the most popular approach to handling spatially correlated data, is unable to generate a high spatial correlation between neighboring areas even when we take extreme values for its spatial correlation parameter. We propose a new spatial model that alleviates this limitation of the CAR model. We consider a Normal/Independent (NI) class of distributions for the random effects generating a wide class of spatial models. The NI class includes the Normal family and heavy-tailed distributions such as the Student-t distribution. Another approach to building a heavy-tailed Student-t spatial model (RENeGe-TS) was also introduced assuming a conditional normal distribution for the random effects where the random covariance matrix has an Inverse-Wishart distribution. Independently of the model used for the spatial random effects, the novelty in our approach is the way the spatial correlation is induced. We assign spatial random effects to the *edges* of this neighborhood graph. The spatial random effect of each area is a linear combination of incident edge effects. For the incident edge effects, we assigned a multivariate NI (or the Student-t for RENeGe-TS) distribution where the spatial covariance matrix has a CAR-like structure inducing a heavy-tailed behavior for the spatial random effects. We prove that the proposed RENeGe model induces a higher marginal correlation than the CAR model, alleviating one of the main limitations of the CAR and ICAR models.

The proposed model, compared to the CAR model, better accounts for heterogeneity, providing a better reconstruction of the image. For the cancer datasets, the proposed model outperforms many other spatial models previously introduced in the literature showing that it is a competitive model to account for spatial correlation.

Supplementary Material

Supplementary Material for "Inducing High Spatial Correlation with Randomly Edge-Weighted Neighborhood Graphs" (DOI: 10.1214/23-BA1390SUPP; .pdf). In Supplementary Material, we present (i) another heavy-tailed RENeGe model (called RENeGe-TS model). We also provide (ii) further results related to the correlation structure of RENeGe ; (iii) the full conditional distributions for Normal and Poisson spatial models under the RENeGe structure; (iv) a sensitivity analysis to the prior specification for σ_θ^2;

(v) additional results related to the analysis of Cancer and simulated data and, (vi) an application of RENeGe for image reconstruction, (vii) and details, figures, and results associated with the simulated data analysis under four spatial scenarios.

References

Arellano-Valle, R. B. and Bolfarine, H. (1995). "On some characterizations of the *t*-distribution." *Statistics & Probability Letters*, 25(1): 79–85. MR1364821. doi: https://doi.org/10.1016/0167-7152(94)00208-P.   1255

Assunção, R. M. (2003). "Space varying coefficient models for small area data." *Environmetrics*, 14(5): 453–473.   1248

Assunção, R. M. and Krainski, E. (2009). "Neighborhood dependence in Bayesian spatial models." *Biometrical Journal*, 51(5): 851–869. MR2751717. doi: https://doi.org/10.1002/bimj.200900056.   1249, 1256, 1257

Assunção, R. M., Potter, J. E., and Cavenaghi, S. M. (2002). "A Bayesian space varying parameter model applied to estimating fertility schedules." *Statistics in Medicine*, 21(14): 2057–2075.   1248

Bapat, R. B. (2014). *Graphs and Matrices*. Springer. London. MR3289036. doi: https://doi.org/10.1007/978-1-4471-6569-9.   1255

Besag, J. (1974). "Spatial interaction and the statistical analysis of lattice systems." *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 36(2): 192–236. MR0373208.   1247

Besag, J., Green, P., Higdon, D., and Mengersen, K. (1995). "Bayesian computation and stochastic systems." *Statistical Science*, 10(1): 3–41. MR1349818.   1268

Besag, J., York, J., and Mollié, A. (1991). "Bayesian image restoration, with two applications in spatial statistics." *Annals of the Institute of Statistical Mathematics*, 43(1): 1–20. MR1105822. doi: https://doi.org/10.1007/BF00116466.   1247, 1268

Brewer, M. J. and Nolan, A. J. (2007). "Variable smoothing in Bayesian intrinsic autoregressions." *Environmetrics*, 18(8): 841–857. MR2416406. doi: https://doi.org/10.1002/env.844.   1249

Brooks, S., Smith, J., Vehtari, A., Plummer, M., Stone, M., Robert, C., Titterington, D., Nelder, J., Atkinson, A., Dawid, A., Lawson, A., Clark, A., Bernardo, J., Sahu, S., Richardson, S., Green, P., Burnham, K., DeIorio, M., Robert, C., Draper, D., Gelfand, A., Trevisani, M., Hodges, J., Lee, Y., De Luna, X., and Meng, X. (2002). "Discussion on the paper by Spiegelhalter, Best, Carlin and van der Linde." *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, 64(4): 616–639.   1268

Carlin, B. P. and Banerjee, S. (2003). "Hierarchical multivariate CAR models for spatio-temporally correlated survival data." In Bernardo, J. M., Dawid, A. P., Berger, J. O., West, M., Heckerman, D., Bayarri, M., and Smith, A. F. (eds.), *Bayesian Statistics 7*, 45–63. Oxford University Press. MR2003166.   1248, 1249

Carlin, B. P. and Louis, T. A. (1996). *Bayes and Empirical Bayes Methods for Data Analysis*. London: Chapman & Hall. MR1427749.   1268

Congdon, P. (2004). "Modelling trends and inequality in small area mortality." *Journal of Applied Statistics*, 31(6): 603–622. MR2109291. doi: https://doi.org/10.1080/1478881042000214695.   1250

Cruz, D., Assunção, R. M., and Loschi, R. H. (2023). "Supplementary Material for "Inducing high spatial correlation with randomly edge-weighted neighborhood graphs"." *Bayesian Analysis.* doi: https://doi.org/10.1214/23-BA1390SUPP.   1251, 1256, 1262, 1264, 1265, 1266, 1267, 1269, 1270, 1272, 1276

Datta, A., Banerjee, S., Finley, A. O., and Gelfand, A. E. (2016). "On nearest-neighbor Gaussian process models for massive spatial data." *WIREs Computational Statistics*, 8(5): 162–171. MR3544254. doi: https://doi.org/10.1002/wics.1383.   1250

Datta, A., Banerjee, S., and Hodges, J. S. (2019). "Spatial disease mapping using directed acyclic graph auto-regressive (DAGAR) models." *Bayesian Analysis*, 14(8): 1221–1244. MR4044851. doi: https://doi.org/10.1214/19-BA1177.   1250, 1268

de Valpine, P., Paciorek, C., Turek, D., Michaud, N., Anderson-Bergman, C., Obermeyer, F., Wehrhahn Cortes, C., Rodrìguez, A., Temple Lang, D., and Paganin, S. (2021). *NIMBLE: MCMC, Particle Filtering, and Programmable Hierarchical Modeling*. R package version 0.12.1.   1268

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. MR1833273. doi: https://doi.org/10.1111/1467-9876.00229.   1248

Fonseca, T., Ferreira, M., and Migon, H. (2008). "Objective Bayesian analysis for the Student-*t* regression model." *Biometrika*, 95(2): 325–333. MR2521587. doi: https://doi.org/10.1093/biomet/asn001.   1256

Gelfand, A. E., Hyon-Jung, K., Sirmans, C. F., and Banerjee, S. (2003). "Spatial modeling with spatially varying coefficient processes." *Journal of the American Statistical Association*, 98(462): 387–396. MR1995715. doi: https://doi.org/10.1198/016214503000170.   1248

Gelfand, A. E. and Vounatsou, P. (2003). "Proper multivariate conditional autoregressive models for spatial data analysis." *Biostatistics*, 4(1): 11–15.   1248, 1249

Gelman, A., Carlin, B. P., Stern, H. S., and Rubin, D. B. (2004). *Bayesian Data Analysis*. Chapman and Hall/CRC, 2nd ed. edition. MR2027492.   1268

Golub, G. H. and Van Loan, C. F. (1996). *Matrix Computations*. The Johns Hopkins University Press, third edition. MR1417720.   1254

Hanks, E. M., Schliep, E. M., Hooten, M. B., and Hoeting, J. A. (2015). "Restricted spatial regression in practice: Geostatistical models, confounding, and robustness under model misspecification." *Environmetrics*, 26(4): 243–254. MR3340961. doi: https://doi.org/10.1002/env.2331.   1272

Hodges, J. S. and Reich, B. J. (2010). "Adding spatially-correlated errors can mess up the fixed effect you love." *The American Statistician*, 64(4): 325–334. MR2758564. doi: https://doi.org/10.1198/tast.2010.10052.   1272

Hughes, J. and Haran, M. (2013). "Dimension reduction and alleviation of confounding for spatial generalized linear mixed models." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 75(1): 139–159. MR3008275. doi: https://doi.org/10.1111/j.1467-9868.2012.01041.x.   1262, 1264, 1272

Jin, X. and Carlin, B. P. (2005). "Multivariate parametric spatiotemporal models for county level breast cancer survival data." *Lifetime Data Analysis*, 11(1): 5–27. MR2139479. doi: https://doi.org/10.1007/s10985-004-5637-1.   1248

Juárez, M. and Steel, M. (2010). "Model-based clustering of non-Gaussian panel data based on skew-*t* distributions." *Journal of Business and Economic Statistics*, 28(1): 52–66. MR2650600. doi: https://doi.org/10.1198/jbes.2009.07145.   1256

Knorr-Held, L. and Best, N. G. (2001). "A shared component model for detecting joint and selective clustering of two diseases." *Journal of the Royal Statistical Society, Series A*, 164(1): 73–85. MR1819023. doi: https://doi.org/10.1111/1467-985X.00187.   1248

Knorr-Held, L., Graziano, G., Frank, C., and Rue, H. (2006). "Joint spatial analysis of gastrointestinal infectious diseases." *Statistical Methods in Medical Research*, 15(5): 465–480. MR2307627. doi: https://doi.org/10.1177/0962280206071642.   1248

Kotz, S. and Nadarajah, S. (2004). *Multivariate T-Distributions and Their Applications*. Cambridge University Press. MR2038227. doi: https://doi.org/10.1017/CBO9780511550683.   1255

Lange, K. and Sinsheimer, J. (1993). "Normal/independent distributions and their applications in robust regression." *Journal of Computational and Graphical Statatistics*, 2(2): 175–198. MR1272391. doi: https://doi.org/10.2307/1390698.   1250, 1253

Langford, I., Leyland, A. H., Rasbash, J., and Goldstein, H. (1999). "Multilevel modelling of the geographical distributions of diseases." *Journal of the Royal Statistical Society, Serie C (Applied Statistics)*, 48(2): 253–268.   1249, 1250

Lee, D. and Mitchell, R. (2012). "Boundary detection in disease mapping studies." *Biostatistics*, 13: 415–426.   1249

Lee, D., Rushworth, A., and Sahu, S. K. (2014). "A Bayesian localized conditional autoregressive model for estimating the health effects of air pollution." *Biometrics*, 70(2): 419–429. MR3258046. doi: https://doi.org/10.1111/biom.12156.   1249

Leroux, B. G., Lei, X., and Breslow, N. (2000). "Estimation of disease rates in small areas: A new mixed model for spatial dependence." In Halloran, M. E. and Berry, D. (eds.), *Statistical Models in Epidemiology, the Environment, and Clinical Trials*, 179–191. New York, NY: Springer New York. MR1731684. doi: https://doi.org/10.1007/978-1-4612-1284-3_4.   1250, 1268

Leyland, A. H., Langford, I., Rasbash, J., and Goldstein, H. (2000). "Multivariate spatial models for event data." *Statistics in Medicine*, 19(17-18): 2469–2478.   1250

Lu, H., Reilly, C. S., Banerjee, S., and Carlin, B. P. (2007). "Bayesian areal wombling via adjacency modeling." *Environmental and Ecological Statistics*, 14(4): 433–452. MR2405556. doi: https://doi.org/10.1007/s10651-007-0029-9.   1249

Ma, H. and Carlin, B. P. (2007). "Bayesian multivariate areal wombling for multiple disease boundary analysis." *Bayesian Analysis*, 2(2): 281–302. MR2312282. doi: https://doi.org/10.1214/07-BA211.   1249

Ma, H., Carlin, B. P., and Banerjee, S. (2010). "Hierarchical and joint site-edge methods for medicare hospice service region boundary analysis." *Biometrics*, 66(2): 355–364. MR2758815. doi: https://doi.org/10.1111/j.1541-0420.2009.01291.x.   1249

MacNab, Y. and Dean, C. B. (2000). "Parametric bootstrap and penalized quasi-likelihood inference in conditional autoregressive models." *Statistics in Medicine*, 19(17-18): 2421–2435.   1248, 1250

Martínez-Beneito, M. A. and Botella-Rocamora, P. (2019). *Disease Mapping: From Foundations to Multidimensional Modeling*. CRC Press.   1247, 1248

Martínez-Beneito, M. A., Lopez-Quilez, A., and Botella-Rocamora, P. (2008). "An autoregressive approach to spatio-temporal disease mapping." *Statistics in Medicine*, 27(15): 2874–2889. MR2521570. doi: https://doi.org/10.1002/sim.3103.   1248

Nobre, W. S., Schmidt, A. M., and Pereira, J. B. (2021). "On the effects of spatial confounding in hierarchical models." *International Statistical Review*, 89(2): 302–322. MR4411907. doi: https://doi.org/10.1111/insr.12407.   1272

Prates, M. O., Assunção, R. M., and Rodrigues, E. C. (2019). "Alleviating spatial confounding for areal data problems by displacing the geographical centroids." *Bayesian Analysis*, 14(2): 623–647. MR3959875. doi: https://doi.org/10.1214/18-BA1123.   1272

Prates, M. O., Dey, D. K., and Lachos, V. H. (2012). "A dengue fever study in the state of Rio de Janeiro with the use of generalized skew-normal/independent spatial fields." *Chilean Journal of Statistics.*, 3(2): 143–155. MR2982067.   1250, 1268

Reich, B. J., Hodges, J. S., and Zadnik, V. (2006). "Effects of residual smoothing on the posterior of the fixed effects in disease-mapping models." *Biometrics*, 62(4): 1197–1206. MR2307445. doi: https://doi.org/10.1111/j.1541-0420.2006.00617.x.   1262, 1264

Rodrigues, E. C. and Assunção, R. M. (2012). "Bayesian spatial models with a mixture neighborhood structure." *Journal of Multivariate Analysis*, 109: 88–102. MR2922856. doi: https://doi.org/10.1016/j.jmva.2012.02.017.   1250, 1268

Silva, G. L., Dean, C. B., Niyonsenga, T., and Vanasse, A. (2008). "Hierarchical Bayesian spatiotemporal analysis of revascularization odds using smoothing splines." *Statistics in Medicine*, 27(13): 2381–2401. MR2432495. doi: https://doi.org/10.1002/sim.3094.   1248

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. MR1979380. doi: https://doi.org/10.1111/1467-9868.00353.   1268

Teixeira, L. V., Assunção, R. M., and Loschi, R. H. (2019). "Bayesian space-time partitioning by sampling and pruning spanning trees." *Journal of Machine Learning Research*, 20(85): 1–35. MR3960939.   1270

Ver Hoef, J., Peterson, E., Hooten, M., Hanks, E., and Fortin, M.-J. (2017). "Spatial autoregressive models for statistical inference from ecological data." *Ecological Monographs*.   1263

Villa, C. and Walker, S. (2014). "Objective prior for the number of degrees of freedom of a t distribution." *Bayesian Analysis*, 9(1): 197–220. MR3188305. doi: https://doi.org/10.1214/13-BA854.   1256

Wall, M. M. (2004). "A close look at the spatial structure implied by the CAR and SAR models." *Journal of Statistical Planning and Inference*, 121(2): 311–324. MR2038824. doi: https://doi.org/10.1016/S0378-3758(03)00111-3.   1248, 1249, 1256

Watanabe, S. (2010). "Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory." *Journal of Machine Learning Research*, 11: 3571–3594. MR2756194.   1268