




## Bayesian Multivariate Spatial Models for Lattice Data with INLA

Francisco Palmí-Perales   
Universidad de Castilla-La Mancha

Virgilio Gómez-Rubio   
Universidad de Castilla-La Mancha

Miguel A. Martínez-Beneito   
Universitat de València

---

### Abstract

The **INLAMSM** package for the R programming language provides a collection of multivariate spatial models for lattice data that can be used with the **INLA** package for Bayesian inference. The multivariate spatial models implemented include different structures to model the spatial variation of the variables and the between-variables variability. In this way, fitting multivariate spatial models becomes faster and easier. The use of the different models included in the package is illustrated using two different datasets: the well-known North Carolina SIDS data and mortality by three causes of death in Comunidad Valenciana (Spain).

*Keywords:* INLA, spatial statistics, multivariate modeling, R.

---

### 1. Introduction

The integrated nested Laplace approximation (INLA, [Rue, Martino, and Chopin 2009](#)) provides an alternative to traditional Markov chain Monte Carlo (MCMC, [Gilks, Gilks, Richardson, and Spiegelhalter 1996](#)) for Bayesian inference. The INLA methodology focuses on estimating the posterior marginals of the model parameters instead of their joint posterior distribution. INLA is implemented in the **INLA** package ([Lindgren and Rue 2015](#)) for the R programming language, that provides a simple way to fit models via the `inla()` function, which works in a similar way as other functions to fit regression models such as `glm()` or `gam()`.

The **INLA** package implements several likelihoods, priors and latent effects that can be used to build models. It is also capable of fitting models with several likelihoods, which can be

useful for multivariate modeling. However, multivariate spatial models are not included. The **INLAMSM** package (Palmí-Perales, Gómez-Rubio, and Martínez-Beneito 2021) adds a number of multivariate latent effects that implement well-known multivariate spatial models for areal data. By fitting these models with INLA, instead of MCMC, computing times should be reduced.

Lattice data are made of  $I$  areas, usually related to some administrative boundaries, where data are collected. We will assume that values of  $K$  variables are obtained from each area. In addition, an adjacency structure is defined by considering that areas with a shared boundary are neighbors. The analysis of this type of datasets is often made by resorting to multivariate regression models.

This paper is organized as follows. The remainder of this introduction includes a review of current software for multivariate spatial modeling. Section 2 gives a description of the different multivariate spatial models implemented in the **INLAMSM** package. Two detailed applications of the **INLAMSM** package and the corresponding results can be found in Section 3. Finally, a summary and discussion are provided in Section 4.

The R programming language provides a number of standalone packages for multivariate spatial analysis. In the specific case of analyzing spatial point patterns, some R packages are available, such as **spatstat** (Baddeley and Turner 2005) and **spatialkernel** (Gómez-Rubio *et al.* 2017). Package **spatstat** is able to model multitype point patterns as well as handling a good deal of the models and methods for the analysis of spatial and spatio-temporal point patterns such as estimators of the space-time inhomogeneous  $K$ -function and pair correlation function. Package **spatialkernel** performs edge-corrected kernel density estimation and binary kernel regression estimation for multivariate spatial point patterns.

Regarding geostatistical data, **gstat** (Pebesma and Wesseling 1998; Pebesma 2004) provides functions to fit both univariate and multivariate models using different types of kriging. Package **spBayes** (Finley, Banerjee, and Carlin 2007) is able to fit a wide variety of Gaussian spatial process models for univariate as well as multivariate point-referenced data using efficient MCMC algorithms.

Finally, in the case of areal data, R package **CARBayes** (Lee 2013) offers the possibility of fitting a wide class of CAR models using MCMC methods. This package provides a number of univariate and multivariate likelihoods, and it also includes a multivariate CAR model with an inverse-Wishart and a CAR prior proposed by Leroux, Lei, and Breslow (2000) to estimate the variability between the different variables and the spatial variation, respectively.

In addition to these packages, the BUGS language (Lunn, Spiegelhalter, Thomas, and Best 2009) is a flexible framework for the implementation of multivariate spatial models for lattice data. Package **R2WinBUGS** (Sturtz, Ligges, and Gelman 2005) provides a set of functions to call the WinBUGS software (Lunn, Thomas, Best, and Spiegelhalter 2000) from R. Recently, Stan (Carpenter *et al.* 2017) develops a flexible language for Bayesian inference that could also be used to fit (multivariate) spatial models (Morris, Wheeler-Martin, Simpson, Mooney, Gelman, and DiMaggio 2019). The **RStan** package (Stan Development Team 2019) is a convenient interface between R and Stan to fit models with ease. Both BUGS and Stan rely on MCMC algorithms for model fitting and inference.

A computationally efficient alternative to software based on MCMC algorithms is described in Rue *et al.* (2009). This has been implemented in an R package called **INLA**, which is often referred to as **R-INLA** to distinguish it from the main INLA methodology. INLA advantages

include an easy way of implementing hierarchical models and short computing times to fit spatial or spatio-temporal models (Lindgren and Rue 2015; Bakka *et al.* 2018). This software is designed for Bayesian inference on latent Gaussian Markov random field (GMRF) models which include (generalized) linear mixed and spatial and spatio-temporal models. **INLA** can deal with lattice data as well as geostatistical data by means of an approximation to continuous spatial processes based on stochastic partial differential equations (SPDE, Lindgren, Rue, and Lindström 2011). This approach can be used to fit log-Gaussian Cox processes (Simpson, Illian, Sørbye, and Rue 2016) for spatial and spatio-temporal point patterns as well (Lindgren *et al.* 2011). Multivariate models can be fit by considering different likelihoods with shared terms. See, for example, Lindgren *et al.* (2011), Krainski *et al.* (2018) and Gómez-Rubio (2020).

Recently, some R packages have been developed on top of **R-INLA**. For instance, **inlabru** (Bachl, Lindgren, Borchers, and Illian 2019) uses **INLA** to fit spatial density surfaces and estimating abundance in a spatial point process, gridded and georeferenced context.

As our review has shown, there are only a few packages that can fit multivariate spatial models in a Bayesian framework. Furthermore, in the case of analyzing multivariate areal data using Bayesian inference, there are options based on MCMC and INLA. INLA is a widely used alternative to MCMC and often provides shorter computing times when fitting spatial models. However, there are not a set of functions to fit multivariate spatial models for lattice data using **INLA**. Developing these latent effects is the main goal of this paper, which will enable users to include multivariate spatial effects in their models to consider different ways of measuring spatial cross-correlation (Sain and Cressie 2002).

Some of the models implemented in the **INLAMSM** package have been described by several authors (Gelfand and Vounatsou 2003; Carlin and Banerjee 2003). Other authors (Jin, Banerjee, and Carlin 2007; Martínez-Beneito 2013; MacNab 2018) provide different reviews of multivariate spatial models that describe the models implemented and many others that could be added to the package in the future. Furthermore, a recent review of multivariate spatial models in the context of disease mapping is provided in Martínez-Beneito and Botella-Rocamora (2019).

## 2. Models

**INLAMSM** implements different functions which correspond to different multivariate spatial latent effects. These differ in structure, complexity and number of hyperparameters. However, all of them are defined in a multivariate spatial context in which  $i = 1, \dots, I$  represents the spatial areas and  $k = 1, \dots, K$  is used to index the variables measured in region  $i$ .

Random effects can be represented using a matrix  $\Theta$  with entries  $\theta_{ik}$ ,  $i = 1, \dots, I$ ,  $k = 1, \dots, K$ . Hence, the  $k$ -th column of  $\Theta$  represent the spatial random effects associated to variable  $k$ .

A particular application of these multivariate spatial models is disease mapping, as described in the examples in Section 3. In this context, the number of the observed cases of  $k$ -th disease in the  $i$ -th spatial area,  $Y_{ik}$ , is modeled as a Poisson random variable:

$$Y_{ik} \sim Po(E_{ik} \cdot R_{ik})$$

where  $E_{ik}$  and  $R_{ik}$  represent the expected cases and the relative risk for the  $i$ -th spatial area

and the  $k$ -th disease, respectively. Then, the logarithm of the relative risk is the sum of two terms:

$$\log(R_{ik}) = a_k + \theta_{ik}$$

where  $a_k$  is the intercept of the  $k$ -th disease and  $\theta_{ik}$  is the term that models the spatial variability. Note that other covariates can be included in the linear predictor on the right hand side of the previous equation.

Following [Martinez-Beneito \(2013\)](#), the variability of a multivariate spatial latent effect can be divided in two terms: variability between the values of the variables measured in the same area and the spatial variability corresponding to the values of different areas for a particular variable. Both sources of variability are modeled with their respective variance matrices, that have different number of hyperparameters.

Finally, before explaining any model, some useful specifications are defined. Let  $\Theta_i$ . ( $i = 1, \dots, I$ ) denote the  $i$ -th row of matrix  $\Theta$  and  $\Theta_{\cdot j}$  ( $j = 1, \dots, J$ ) denote the  $j$ th column of matrix  $\Theta$ . In addition, on a matrix  $A = [A_{\cdot 1} : \dots : A_{\cdot J}]$ , operator  $\text{vec}(\cdot)$  is defined as

$$\text{vec}(A) = (A_{\cdot 1}^\top, \dots, A_{\cdot J}^\top)^\top.$$

Note that all latent effects will be defined using the `rgeneric` latent effect, which defines the latent effect using the representation of the Gaussian Markov random field. This is described in the **INLA** documentation, that can be accessed with `inla.doc("rgeneric")`, and also in Chapter 11 of [Gómez-Rubio \(2020\)](#). This representation is essentially a multivariate Gaussian vector with a sparse precision matrix. Hence,  $\text{vec}(\Theta)$  will be defined using a Normal distribution with zero mean and a highly structured variance matrix  $\Sigma$ , i.e.,

$$\text{vec}(\Theta) \sim N(0, \Sigma).$$

Note that **INLA** works with the precision and that this is why  $\Sigma^{-1}$  will be required instead of the variance matrix  $\Sigma$ . Furthermore,  $\Sigma$  will be created so that it accounts for the association between the different variables within areas as well as association between areas for the same variable or disease, as explained above. The new latent effects included in the **INLAMSM** can be used as a guidance to develop other types of multivariate spatial latent effects.

## 2.1. Independent intrinsic MCAR

One of the easiest options to model the variability within- and between-variables is using an intrinsic CAR distribution ([Besag, York, and Mollié 1991](#)) for the spatial structure and a diagonal matrix for the covariance matrix between variables, respectively. Specifically,  $\Theta$  is modeled as

$$\text{vec}(\Theta) \sim N\left(0, \Lambda^{-1} \otimes (D - W)^{-1}\right)$$

where  $D = \text{diag}(n_1, \dots, n_I)$  is a diagonal matrix with values  $n_i$  (the number of neighbors of region  $i$ ) and  $W$  is an adjacency matrix with entries  $W_{ij}$ . Each entry  $W_{ij}$  is equal to 1 if units  $i$  and  $j$  are neighbors and 0 otherwise.

Let  $\tau_k$  be the marginal precision of the  $k$ -th variable. When variables are independent of each

other, the between-variables precision matrix  $\Lambda$  is a  $k \times k$  diagonal matrix defined as

$$\Lambda = \begin{bmatrix} \tau_1 & 0 & \cdots & \cdots & 0 \\ 0 & \tau_2 & 0 & & \vdots \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & \tau_K \end{bmatrix}.$$

Following [Jin \*et al.\* \(2007\)](#), matrix  $\Lambda$  can be regarded as a (precision) matrix for the association between variables within the same area and (precision) matrix  $(D - W)$  to account for the association between areas for a given variable. In this case it is assumed that the different variables are independent and this is why  $\Lambda$  is a diagonal matrix, but in other latent random effects introduced below  $\Lambda$  will be a dense matrix to model association between the different variables.

Thus, this model has  $K$  hyperparameters,  $\{\tau_k\}_{k=1}^K$ , equal to the total number of variables. This model is computationally fast because the number of hyperparameters is the lowest among the multivariate spatial models implemented in the **INLAMSM** package, as it will be seen below. Therefore, this model can be used as a baseline to compare a naïve model with independent spatial terms with more complex alternatives.

Hyperparameters are internally represented in **INLA** using the log-scale so that they are not bounded, and the vector of hyperparameters is  $(\log(\tau_1), \dots, \log(\tau_K))$ . Prior distributions are assigned to the standard deviations instead of the precisions ([Gelman 2006](#)). In particular, a uniform improper distribution between 0 and  $+\infty$  is assigned to each standard deviation:

$$\sigma_k = \frac{1}{\sqrt{\tau_k}} \sim Un(0, +\infty); \quad k = 1, \dots, K.$$

Note that **INLA** will report the results in the internal scale but we have included several functions in the package to transform from the log-precision into some other more meaningful scale. See [Section 3](#) for details. Furthermore, **INLA** provides a set of functions that can be used to make this transformation (see, for example, [Gómez-Rubio 2020](#), Chapter 2).

Finally, given that this is an improper distribution, it will be necessary to add further constraints when fitting the model, similarly to the one dimensional intrinsic CAR distribution ([Besag and Kooperberg 1995](#)). In this case, a sum-to-zero constraint will be added to the random effects associated to each variable. This implies that a variable-specific intercept should be added to the model as well.

## 2.2. Independent proper MCAR

A proper CAR distribution can be used to model the within-variables variability instead of the intrinsic version. Here, a common spatial autocorrelation parameter  $\alpha$  for all the variables is introduced. Thus  $\Theta$  is modeled as

$$\text{vec}(\Theta) \sim N\left(0, \Lambda^{-1} \otimes (D - \alpha \cdot W)^{-1}\right)$$

where matrices  $D$  and  $W$  are defined similarly as above. The matrix used to model the between-variables variability keeps the same structure, thus  $\Lambda$  is defined as a diagonal matrix with entries  $(\tau_1, \dots, \tau_K)$ .

The proper CAR specification described above is not implemented in the **INLA** package. In our opinion this is an additional contribution of our **INLAMSM** package. It is worth noting that the way in which the proper CAR distribution is implemented here differs from the proper latent random effects implemented in the **INLA** package as defined in latent effects `besagproper` and `besagproper2`. The `besagproper` effect is based on adding a constant to the diagonal entries of the precision matrix of the intrinsic CAR distribution, while the `besagproper2` effect is actually the model proposed by Leroux *et al.* (2000).

Now, this model has  $K + 1$  hyperparameters which are the spatial autocorrelation parameter,  $\alpha$ , and the  $K$  precisions,  $\{\tau_k\}_{k=1}^K$ . Internally,  $(\alpha^*, \log(\tau_1), \dots, \log(\tau_K))$  are the hyperparameters. Hyperparameter  $\alpha^*$  is defined by transforming  $\alpha$  as follows:

$$\alpha^* = \text{logit} \left( \frac{\alpha - \alpha_{\min}}{\alpha_{\max} - \alpha_{\min}} \right),$$

where  $\alpha_{\min}$  and  $\alpha_{\max}$  define the bounds of the domain of  $\alpha$  (Sun, Tsutakawa, and Speckman 1999). However, the proper CAR distribution has been aduced to show weird artifacts when the spatial correlation parameter takes negative values (Wall 2004). For this reason, it is common to assume  $\alpha_{\min} = 0$  and  $\alpha_{\max} = 1$ .

As in the previous model, uniform prior distributions are set on the standard deviations, and a uniform prior on  $\alpha$  is considered:

$$\begin{aligned} \sigma_k &= \frac{1}{\sqrt{\tau_k}} \sim Un(0, +\infty), & k = 1, \dots, K, \\ \alpha &\sim Un(\alpha_{\min}, \alpha_{\max}). \end{aligned}$$

Once again, the independent proper MCAR distribution is a naïve implementation of independent spatial patterns as the independent intrinsic MCAR. Nevertheless, this model could also be used as a benchmark for comparing multivariate proper CAR models assuming dependence between variables (as described below).

### 2.3. Improper MCAR model

A diagonal precision matrix  $\Lambda$  is a naïve way to model the between-variables variability which leads to short computing times. However, setting the off-diagonal elements of  $\Lambda$  to zero assumes that variables are independent of each other. The off-diagonal elements model dependence between every pair of variables, which could be useful when looking for similar spatial patterns of different variables.

Therefore, a dense precision matrix and an intrinsic conditional autoregressive distribution can be chosen to model the between- and within-variables variability, respectively. This can be regarded as a generalization of the univariate intrinsic conditional autoregressive model. In particular,  $\Theta$  is modeled similarly as the independent IMCAR:

$$\text{vec}(\Theta) \sim N \left( 0, \Lambda^{-1} \otimes (D - W)^{-1} \right)$$

but now  $\Lambda$  is a dense, symmetric and positive-definite matrix. The parameterization of  $\Lambda$  follows that of other latent effects implemented in **INLA** and the hyperparameters are the precisions of the variables and the correlations between any pair of variables. Hence, instead

of dealing with the structure of  $\Lambda$ ,  $\Lambda^{-1}$  is defined as follows:

$$\Lambda^{-1} = \begin{bmatrix} 1/\tau_1 & \rho_{12}/\sqrt{\tau_1\tau_2} & \cdots & \cdots & \rho_{1K}/\sqrt{\tau_1\tau_K} \\ \rho_{12}/\sqrt{\tau_1\tau_2} & 1/\tau_2 & \rho_{23}/\sqrt{\tau_2\tau_3} & & \vdots \\ \vdots & \rho_{23}/\sqrt{\tau_2\tau_3} & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \rho_{(K-1)K}/\sqrt{\tau_{K-1}\tau_K} \\ \rho_{1K}/\sqrt{\tau_1\tau_K} & \cdots & \cdots & \rho_{(K-1)K}/\sqrt{\tau_{K-1}\tau_K} & 1/\tau_K \end{bmatrix}.$$

Here,  $\rho_{jk}$  is the correlation coefficient between variables  $j$  and  $k$  and  $\tau_k$  is the marginal precision of the  $k$ -th variable. Therefore, the set of hyperparameters contains  $K$  marginal precisions and  $K(K-1)/2$  correlation parameters. In this case, a joint prior distribution is considered for  $\Lambda$  instead of setting a prior distribution for each hyperparameter. Thus,  $\Lambda$  follows a Wishart distribution as

$$\Lambda \sim \text{Wishart}_K(r, R^{-1}),$$

where  $r$  is the number of degrees of freedom and  $R^{-1}$  is a fixed symmetric positive definite matrix of size  $K \times K$ . In our implementation,  $r$  is equal to  $K$  (following [Carlin and Banerjee 2003](#)) and  $R$  is the  $K \times K$  identity matrix.

The vector of hyperparameters contains now the precisions plus the correlations in the lower triangular matrix of  $\Lambda^{-1}$  (columnwise). However, these parameters are re-scaled so that the log-precisions are used and the correlation parameters are transformed using

$$\rho^* = \text{logit}((\rho + 1)/2)$$

for any correlation  $\rho$ . Hence, the vector of hyperparameters in the internal scale is now defined as  $(\log(\tau_1), \dots, \log(\tau_K), \rho_{21}^*, \rho_{31}^*, \dots)$ .

As stated above, given that this distribution is improper, a sum-to-zero constraint on the effects for each variable will be added when fitting the model. This also means that variable-specific intercepts are required in the model.

## 2.4. Proper MCAR model

A proper CAR distribution can be used instead of the intrinsic CAR in order to model the within-variables variability. This alternative corresponds to the multivariate generalization of the univariate proper conditional autoregressive model. As in the previous case, a dense  $\Lambda$  matrix is used to model the between-variables variability. Specifically,  $\Theta$  is modeled as

$$\text{vec}(\Theta) \sim N\left(0, \Lambda^{-1} \otimes (D - \alpha \cdot W)^{-1}\right)$$

where  $\alpha$  is a common spatial autocorrelation parameter and  $\Lambda$  is a dense symmetric positive-definite matrix with  $\Lambda^{-1}$  defined as in the previous model.

In this case, the set of hyperparameters comprehends one common spatial autocorrelation parameter  $\alpha$ ,  $K$  marginal precisions,  $\{\tau_k\}_{k=1}^K$ , and  $K(K-1)/2$  correlation parameters,  $\{\rho_{jk}\}_{j,k=1}^K$ . Again, a Wishart distribution is considered as a joint prior distribution for the  $\Lambda$  matrix, while a uniform prior is considered for  $\alpha$ , i.e.,

$$\alpha \sim \text{Un}(\alpha_{\min}, \alpha_{\max}).$$



Latent effect	Wrapper function	rgeneric function	# Hyperparameters
Indep. IMCAR	<code>inla.INDIMCAR.model</code>	<code>inla.rgeneric.indep.IMCAR.model</code>	$K$
Indep. PMCAR	<code>inla.INDMCAR.model</code>	<code>inla.rgeneric.indep.MCAR.model</code>	$K + 1$
IMCAR	<code>inla.IMCAR.model</code>	<code>inla.rgeneric.IMCAR.model</code>	$K(K + 1)/2$
PMCAR	<code>inla.MCAR.model</code>	<code>inla.rgeneric.MCAR.model</code>	$K(K + 1)/2 + 1$
M-model	<code>inla.Mmodel.model</code>	<code>inla.rgeneric.Mmodel.model</code>	$K^2$

Table 1: Summary of the latent effects implemented in the **INLAMSM** package.

In this case, the vector of hyperparameters contains the spatial autocorrelation parameter  $\alpha$ , the precisions and the correlations in the lower triangular elements of  $\Lambda^{-1}$  (columnwise). As in previous models, all hyperparameters are transformed to take values in the  $(-\infty, +\infty)$  interval. Hence, the vector of hyperparameters is  $(\alpha^*, \log(\tau_1), \dots, \log(\tau_K), \rho_{21}^*, \rho_{31}^*, \dots)$ .

## 2.5. M-model

Botella-Rocamora, Martinez-Beneito, and Banerjee (2015) describe a unifying modeling framework for multivariate disease mapping when the number of diseases is potentially large. Here, spatial effects are linear combinations of proper CAR spatial effects. In particular, we will consider  $K$  underlying proper CAR spatial effects defined by

$$\phi_k \sim N\left(0, (D - \alpha_k W)^{-1}\right), \quad k = 1, \dots, K.$$

Here,  $\phi_k$  is a vector of length  $I$ . The value of the spatial random effect  $\Theta_{.j}$  for variable  $j$  is defined as

$$\Theta_{.j} = \phi_1 m_{1j} + \dots + \phi_K m_{Kj}.$$

Hence, matrix  $M$  with entries  $m_{ij}$  defines the loadings of the different underlying CAR spatial effects for each disease or variable. The distribution of these random effects is given by

$$\text{vec}(\Theta) \sim N\left(0, (M^\top \otimes I) \text{diag}((\Sigma_w)_1, \dots, (\Sigma_w)_K) (M \otimes I)\right).$$

Here, matrices  $(\Sigma_w)_k$  are the variance matrices of the  $K$  underlying proper CAR spatial effects, i.e.,

$$(\Sigma_w)_k = (D - \alpha_k W)^{-1}, \quad k = 1, \dots, K.$$

In addition, Botella-Rocamora *et al.* (2015) show that for the separable case, with  $\alpha_1 = \dots = \alpha_K$ , the between-variables variance matrix is  $M^\top M$ . The prior of this model is on  $M^\top M$ , and it follows a Wishart with parameters  $K$  and  $\tau I$ . Parameter  $\tau$  is a fixed precision which is set to 0.001, but it could also be considered as another hyperparameter to be estimated (Corpas-Burgos, Botella-Rocamora, and Martinez-Beneito 2019).

The vector of hyperparameters in this model is made of the  $K$  autocorrelation parameters (conveniently transformed) followed by the columns of matrix  $M$ , for which no transformation is required.

## 2.6. Summary of the models

To summarize the different models implemented and their associated functions, Table 1 displays some basic information about them. Note how all functions names follow a similar pattern for consistency, and that the latent models differ in the number of hyperparameters.



These depend on the number of variables  $K$ . Furthermore, there are two different functions to define the new latent effects: the one that implements the latent effects using the `rgeneric` framework and a wrapper function. The wrapper function is intended to provide a simpler way to define the latent effect and avoid calling function `inla.rgeneric.define` (see, for example, Chapter 11 in [Gómez-Rubio 2020](#)). This will be further discussed in the examples in Section 3 below.

### 3. Examples

In this section, two examples are developed with the **INLAMSM** package. The first one is on the well-known North Carolina sudden infant death syndrome (SIDS) data ([Cressie and Read 1985](#)), which is used to show how the different models implemented in the **INLAMSM** package are fit. This example is also intended to show that multivariate spatial models can not only be used to model several diseases but also the same disease across different time periods so that dependence across different periods can be estimated. The second example is based on the study of three causes of death in Comunidad Valenciana (Spain), a region that comprises 540 municipalities and provides a more challenging dataset to fit spatial models than the North Carolina SIDS data (with only 100 areas). In this second example, the focus is on investigating dependence among the different diseases.

#### 3.1. North Carolina SIDS data

In the first example, the North Carolina SIDS data ([Cressie and Read 1985](#)) have been considered to test the methods implemented in the **INLAMSM** package. This dataset includes counts of number of live births and number of deaths from sudden infant death syndrome for the 100 counties of North Carolina for two time periods: from July 1st, 1974 to Jun 30th, 1978 and from July 1st, 1979 to June 30th, 1984. This dataset is available in a shapefile in R package **spData** ([Bivand, Nowosad, and Lovelace 2019b](#)), which will be loaded using package **rgdal** ([Bivand, Keitt, and Rowlingson 2019a](#)).

```
R> library("rgdal")
R> nc.sids <- readOGR(system.file("shapes/sids.shp", package = "spData")[1],
+   verbose = FALSE)
R> proj4string(nc.sids) <- CRS("+proj=longlat +ellps=clrk66")
```

Next, the adjacency matrix of the 100 counties in North Carolina is obtained with function `poly2nb` in the **spdep** package ([Bivand, Pebesma, and Gómez-Rubio 2013](#); [Bivand and Wong 2018](#)) and converted into a sparse matrix of type `Matrix` ([Bates and Maechler 2019](#)).

```
R> library("spdep")
R> adj <- poly2nb(nc.sids)
R> W <- as(nb2mat(adj, style = "B"), "Matrix")
```

The model that will be fit in this case is

$$O_{ik} \sim Po(\mu_{ik}); \log(\mu_{ik}) = \log(E_{ik}) + a_k + \theta_{ik}, \quad i = 1, \dots, 100, \quad k = 1, 2.$$

Here,  $O_{ik}$  is the number of SIDS cases in county  $i$  and time period  $k$ ,  $E_{ik}$  the expected counts,  $a_k$  a period-specific intercept and  $\theta_{ik}$  the multivariate spatial effect, which is defined using

one of the models implemented in the **INLAMSM** package. Note that other covariates and effects could be included in the linear predictor as well.

The expected number of cases  $E_{ik}$  are computed by multiplying the period-specific mortality rate  $r_k$  by the number of births  $N_{ik}$ :

$$E_{ik} = r_k N_{ik}; \quad r_k = \frac{\sum_{i=1}^{100} O_{ik}}{\sum_{i=1}^{100} N_{ik}}, \quad i = 1, \dots, I, \quad k = 1, 2.$$

In the next lines of R code the expected counts for both time periods are computed, as well as the standardized mortality ratio (SMR),  $O_{ik}/E_{ik}$ , and the proportion of non-white births, which could be used as a covariate for both time periods.

```
R> r74 <- sum(nc.sids$SID74) / sum(nc.sids$BIR74)
R> nc.sids$EXP74 <- r74 * nc.sids$BIR74
R> nc.sids$SMR74 <- nc.sids$SID74 / nc.sids$EXP74
R> nc.sids$NWPROP74 <- nc.sids$NWBIR74 / nc.sids$BIR74
R> r79 <- sum(nc.sids$SID79) / sum(nc.sids$BIR79)
R> nc.sids$EXP79 <- r79 * nc.sids$BIR79
R> nc.sids$SMR79 <- nc.sids$SID79 / nc.sids$EXP79
R> nc.sids$NWPROP79 <- nc.sids$NWBIR79 / nc.sids$BIR79
```

In order to prepare the data to fit the models with **INLA**, a new object is created by appending the data from the first time period to the data from the second one. In addition to the observed, expected and proportion of non-white births, an index is created to identify the counties and time periods.

```
R> d <- data.frame(OBS = c(nc.sids$SID74, nc.sids$SID79),
+   PERIOD = c(rep("74", 100), rep("79", 100)),
+   NWPROP = c(nc.sids$NWPROP74, nc.sids$NWPROP79),
+   EXP = c(nc.sids$EXP74, nc.sids$EXP79))
R> d$idx <- 1:length(d$OBS)
```

Now, we will fit the models. This requires a two step process:

1. Define the multivariate spatial latent effect using one of the functions in the **INLAMSM** package. This is done via the `inla.rgeneric.define` function, that will take the function of the multivariate spatial model to be included plus any other required arguments, such as the adjacency matrix, the number of time periods, etc. Alternatively, the associated wrapper function can be used (see below).
2. Fit the model with **INLA** (Rue, Riebler, Sørbye, Illian, Simpson, and Lindgren 2017) using a formula that includes the newly defined latent effect.

For example, for the independent IMCAR model the latent effect will be defined as follows:

```
R> library("INLAMSM")
R> library("INLA")
R> k <- 2
R> model.indimcar <- inla.rgeneric.define(inla.rgeneric.indep.IMCAR.model,
+   k = k, W = W)
```

The previous definition of the latent effect uses a call to function `inla.rgeneric.define`, which takes function `inla.rgeneric.indep.IMCAR.model` that defines the new latent effect plus any other required arguments (i.e., a named list with the number of diseases `k` and adjacency matrix `W`). However, as explained in Section 2.6 there are wrapper functions associated to each new latent effect, that will be used from now on. For this particular case, the definition of the new latent effect using a wrapper function is:

```
R> model.indimcar <- inla.INDIMCAR.model(k = k, W = W)
```

As explained in Section 2, given that this model includes intrinsic CAR specifications, it is necessary to include some constraints so that the random effects for each disease sum up to zero. Constraints in **INLA** are added using equation  $Ax^\top = e^\top$ , where matrix  $A$  and vector  $e$  are used to define linear constraints on the (multivariate) vector of random effects  $x$ . These two elements are defined below:

```
R> A <- kronecker(Diagonal(k, 1), Matrix(1, ncol = nrow(W), nrow = 1))
R> e = rep(0, k)
```

The constraints are specified using argument `extraconstr` in the call to the `f` function that defines the latent effect (see below).

The next step fits the model with **INLA**:

```
R> IIMCAR <- inla(OBS ~ 0 + PERIOD + f(idx, model = model.indimcar,
+   extraconstr = list(A = as.matrix(A), e = e)),
+   data = d,
+   E = EXP, family = "poisson", control.predictor = list(compute = TRUE),
+   control.compute = list(dic = TRUE, waic = TRUE))
```

Note that linear constraints have been added so that random effects associated to each disease sum up to zero, as explained above. Furthermore, argument `control.predictor` has been set to compute the marginals of the linear predictor (i.e., by setting `compute` equal to `TRUE`) and argument `control.compute` has been set to compute two model selection criteria. In particular, the deviance information criterion (Spiegelhalter, Best, Carlin, and Van der Linde 2002, DIC) and Watanabe-Akaike information criterion (Watanabe 2013, WAIC) will be computed. See below for a discussion on how to use these criteria for model choice.

This model can be summarized as follows:

```
R> summary(IIMCAR)
```

Call:

```
c("inla(formula = OBS ~ 0 + PERIOD + f(idx, model =
model.indimcar, ", " extraconstr = list(A = as.matrix(A), e =
e)), family = \"poisson\", ", " data = d, E = EXP,
control.compute = list(dic = TRUE, waic = TRUE), ", "
control.predictor = list(compute = TRUE))")
```

Time used:

```
Pre = 1.78, Running = 5.51, Post = 0.143, Total = 7.44
```

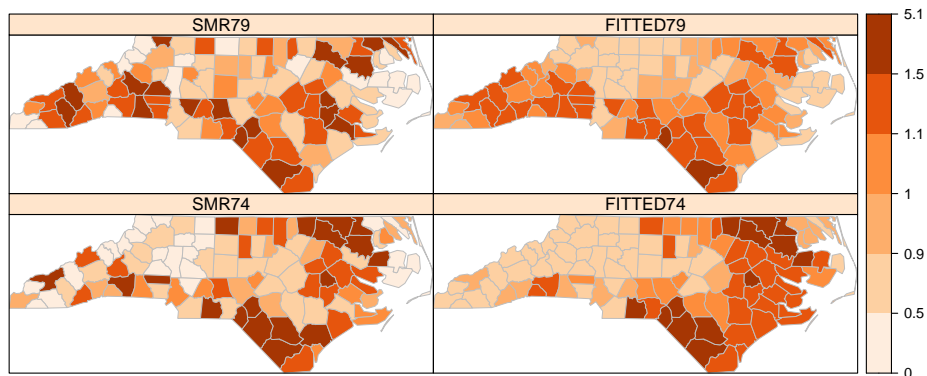


Figure 1: Standardized mortality ratio and posterior means of the fitted values with the IIMCAR model for both time periods for the North Carolina SIDS data.

Fixed effects:

	mean	sd	0.025quant	0.5quant	0.975quant	mode	kld
PERIOD74	-0.071	0.056	-0.184	-0.070	0.036	-0.067	0
PERIOD79	-0.017	0.046	-0.110	-0.016	0.071	-0.015	0

Random effects:

Name	Model
idx	RGeneric2

Model hyperparameters:

	mean	sd	0.025quant	0.5quant	0.975quant	mode
Theta1 for idx	0.774	0.339	0.129	0.764	1.46	0.731
Theta2 for idx	1.429	0.408	0.664	1.414	2.27	1.361

Expected number of effective parameters(stdev): 65.92(8.94)

Number of equivalent replicates : 3.03

Deviance Information Criterion (DIC) .....: 909.59

Deviance Information Criterion (DIC, saturated) ....: 271.91

Effective number of parameters .....: 67.35

Watanabe-Akaike information criterion (WAIC) ...: 910.58

Effective number of parameters .....: 53.61

Marginal log-Likelihood: -484.51

Posterior marginals for the linear predictor and  
the fitted values are computed

Note that hyperparameters  $\Theta_1$  and  $\Theta_2$  correspond to the log-precisions of the latent effect, i.e., the two independent ICAR spatial effects. Because the two effects are independent of each other, this model is equivalent to fitting two models using an intrinsic CAR model

(implemented as the `besag` latent effect in **INLA**).

Next, the fitted values of the relative risks of the independent IMCAR (IIMCAR) model can be compared with the SMR for both time periods as seen in Figure 1. Package **RColorBrewer** (Neuwirth 2014) has been used to set the colors in the palette. Note how the fitted values show smoothed spatial patterns as compared with the raw SMR values. The other models implemented in the **INLAMSM** package can be fit in a similar way, as listed below. For example, the independent proper MCAR model requires the values of  $\alpha_{\min}$  and  $\alpha_{\max}$  to be passed in addition to `k` and `W` when defining the latent effect. This is exemplified below. We have constrained  $\alpha$  to take just positive values since the proper CAR distribution has been said to have a counterintuitive performance when this parameter takes negative values (Wall 2004). Anyway, our implementation of the proper CAR distribution admits, if wanted, the whole range of admissible values for this parameter (Sun *et al.* 1999).

```
R> alpha.min <- 0
R> alpha.max <- 1
R> model.indmcar <- inla.INDMCAR.model(k = k, W = W, alpha.min = alpha.min,
+   alpha.max = alpha.max)
R> IPMCAR <- inla(OBS ~ 0 + PERIOD + f(idx, model = model.indmcar), data = d,
+   E = EXP, family = "poisson", control.predictor = list(compute = TRUE),
+   control.compute = list(dic = TRUE, waic = TRUE))
```

The remainder of the models in the **INLAMSM** package are defined and fit as:

```
R> model.imcar <- inla.IMCAR.model(k = k, W = W)
R> IMCAR <- inla(OBS ~ 0 + PERIOD + f(idx, model = model.imcar,
+   extraconstr = list(A = as.matrix(A), e = e)),
+   data = d,
+   E = EXP, family = "poisson", control.predictor = list(compute = TRUE),
+   control.compute = list(dic = TRUE, waic = TRUE))
```

As discussed above, linear constraints have been added when fitting this model too. This way, the intercepts in the model can be identified as well.

```
R> model.mcar <- inla.MCAR.model(k = k, W = W, alpha.min = alpha.min,
+   alpha.max = alpha.max)
R> PMCAR <- inla(OBS ~ 0 + PERIOD + f(idx, model = model.mcar), data = d,
+   E = EXP, family = "poisson", control.predictor = list(compute = TRUE),
+   control.compute = list(dic = TRUE, waic = TRUE))
R> model.mmodel <- inla.Mmodel.model(k = k, W = W, alpha.min = alpha.min,
+   alpha.max = alpha.max)
R> Mmodel <- inla(OBS ~ 0 + PERIOD + f(idx, model = model.mmodel), data = d,
+   E = EXP, family = "poisson", control.predictor = list(compute = TRUE),
+   control.compute = list(dic = TRUE, waic = TRUE))
```

Once we have fit all models, we can compare point estimates of the relative risks. Posterior means of the relative risks from different models are displayed in Figure 2.

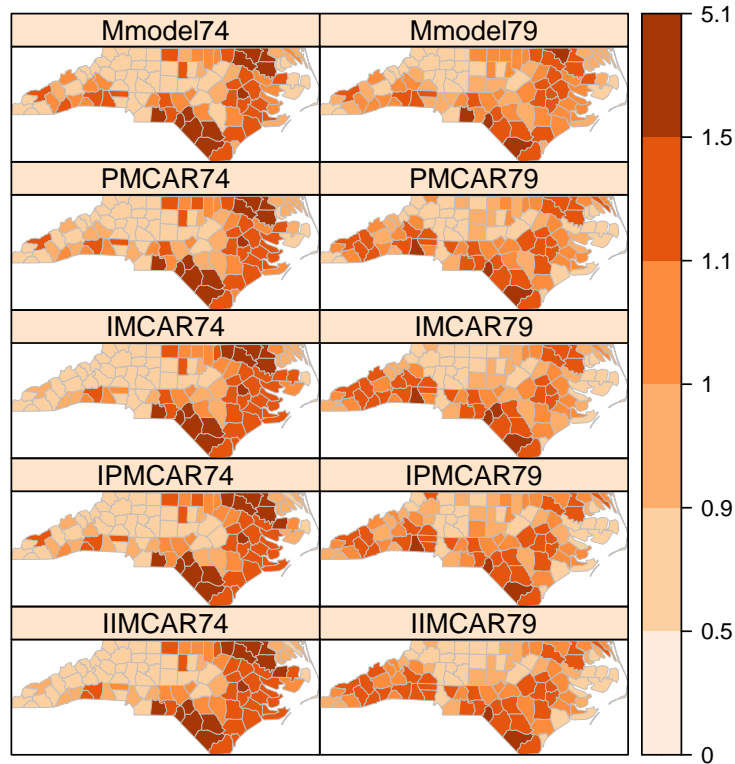


Figure 2: Posterior means of the relative risks for the different multivariate spatial effects fit to the North Carolina SIDS data.

```
R> nc.sids$IIMCAR74 <- IIMCAR$summary.fitted[1:100, "mean"]
R> nc.sids$IIMCAR79 <- IIMCAR$summary.fitted[100 + 1:100, "mean"]
R> nc.sids$IPMCAR74 <- IPMCAR$summary.fitted[1:100, "mean"]
R> nc.sids$IPMCAR79 <- IPMCAR$summary.fitted[100 + 1:100, "mean"]
R> nc.sids$IMCAR74 <- IMCAR$summary.fitted[1:100, "mean"]
R> nc.sids$IMCAR79 <- IMCAR$summary.fitted[100 + 1:100, "mean"]
R> nc.sids$PMCAR74 <- PMCAR$summary.fitted[1:100, "mean"]
R> nc.sids$PMCAR79 <- PMCAR$summary.fitted[100 + 1:100, "mean"]
R> nc.sids$Mmodel74 <- Mmodel$summary.fitted[1:100, "mean"]
R> nc.sids$Mmodel79 <- Mmodel$summary.fitted[100 + 1:100, "mean"]
```

#### *Including covariates in the model*

The proportion of non-white births shows a very similar pattern to that of the relative risk. These proportions for the time periods are shown in Figure 3, which can be compared with the estimates of the relative risk in Figure 1 to appreciate the similar patterns. Note the similar patterns between the proportion of non-white births and the standardized mortality ratio for both time periods. For this reason, several authors (see, for example, [Cressie and Read 1985](#)) have mentioned the importance of including this covariate in the model. Hence, the same models are fitted now including the covariate so that a different coefficient is estimated for each time period.

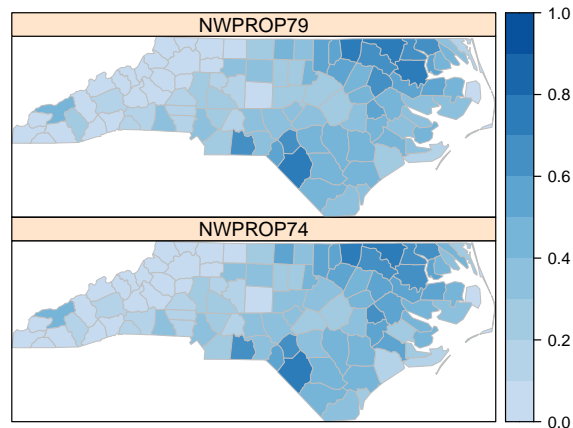


Figure 3: Proportion of non-white births per county for both time periods in the North Carolina SIDS data.

In order to have the covariate in the model so that a different coefficient is estimated for each time period, the covariate must be structured in a two-column as follows:

```
R> n <- nrow(W)
R> NWPROP <- matrix(NA, ncol = 2, nrow = 2 * n)
R> NWPROP[1:n, 1] <- nc.sids$NWP74
R> NWPROP[n + 1:n, 2] <- nc.sids$NWP79
```

Next, data used to fit the previous models are converted into a list so that the covariate can be added:

```
R> d <- as.list(d)
R> d$NWPROP <- NWPROP
```

Finally, models are fit again:

```
R> IIMCAR2 <- inla(OBS ~ 0 + PERIOD + NWPROP + f(idx, model = model.indimcar,
+   extraconstr = list(A = as.matrix(A), e = e)), data = d,
+   E = EXP, family = "poisson", control.predictor = list(compute = TRUE),
+   control.compute = list(dic = TRUE, waic = TRUE))
R> IPMCAR2 <- inla(OBS ~ 0 + PERIOD + NWPROP + f(idx, model = model.indmcar),
+   data = d,
+   E = EXP, family = "poisson", control.predictor = list(compute = TRUE),
+   control.compute = list(dic = TRUE, waic = TRUE))
R> IMCAR2 <- inla(OBS ~ 0 + PERIOD + NWPROP + f(idx, model = model.imcar,
+   extraconstr = list(A = as.matrix(A), e = e)), data = d,
+   E = EXP, family = "poisson", control.predictor = list(compute = TRUE),
+   control.compute = list(dic = TRUE, waic = TRUE))
R> PMCAR2 <- inla(OBS ~ 0 + PERIOD + NWPROP + f(idx, model = model.mcar),
+   data = d, E = EXP, family = "poisson",
+   control.predictor = list(compute = TRUE),
+   control.compute = list(dic = TRUE, waic = TRUE))
```



```
R> Mmodel2 <- inla(OBS ~ 0 + PERIOD + NWPROP + f(idx, model = model.mmodel),
+   data = d, E = EXP, family = "poisson",
+   control.predictor = list(compute = TRUE),
+   control.compute = list(dic = TRUE, waic = TRUE))
```

Note that, in a similar manner that sum-to-zero constraints have been imposed for intrinsic CAR distributions, further constraints could be also imposed above assuming the spatial random effects to be orthogonal to the covariate. In this manner, confounding between the covariate and spatial random effects could be alleviated (Hodges and Reich 2010).

In order to show the estimates of the coefficients of the covariates, the summary of the independent IMCAR model with covariates is shown below:

```
R> summary(IIMCAR2)
```

Call:

```
c("inla(formula = OBS ~ 0 + PERIOD + NWPROP + f(idx, model =
model.indimcar, ", " extraconstr = list(A = as.matrix(A), e =
e)), family = \"poisson\", ", " data = d, E = EXP,
control.compute = list(dic = TRUE, waic = TRUE), ", "
control.predictor = list(compute = TRUE))")
```

Time used:

```
Pre = 1.81, Running = 5.87, Post = 0.112, Total = 7.78
```

Fixed effects:

	mean	sd	0.025quant	0.5quant	0.975quant	mode	kld
PERIOD74	-0.687	0.130	-0.949	-0.684	-0.438	-0.679	0
PERIOD79	-0.211	0.124	-0.458	-0.210	0.031	-0.208	0
NWPROP1	1.978	0.351	1.297	1.974	2.681	1.966	0
NWPROP2	0.622	0.365	-0.103	0.623	1.336	0.627	0

Random effects:

Name	Model
idx	RGeneric2

Model hyperparameters:

	mean	sd	0.025quant	0.5quant	0.975quant	mode
Theta1 for idx	1.91	0.587	0.867	1.87	3.16	1.72
Theta2 for idx	1.54	0.437	0.733	1.53	2.44	1.46

Expected number of effective parameters(stdev): 49.28(9.58)

Number of equivalent replicates : 4.06

Deviance Information Criterion (DIC) .....: 899.49

Deviance Information Criterion (DIC, saturated) ....: 261.81

Effective number of parameters .....: 51.51

Watanabe-Akaike information criterion (WAIC) ....: 904.67

Effective number of parameters .....: 46.14

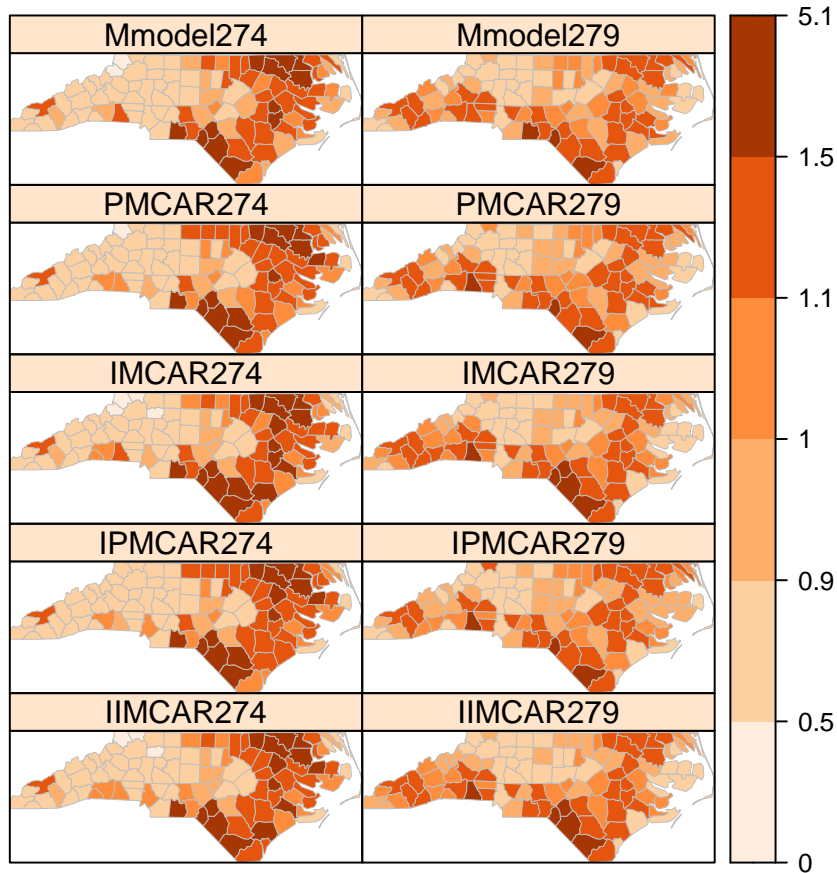


Figure 4: Posterior means of the relative risks from the different multivariate spatial effects fit to the North Carolina SIDS data. These models include the proportion of non-white births as a covariate.

Marginal log-Likelihood: -479.10

Posterior marginals for the linear predictor and  
the fitted values are computed

These results show the positive association between the relative risk and the proportion of non-white births. Other models have similar estimates of the coefficients of this covariate. In addition, Figure 4 shows the posterior means of the relative risks estimated with these models. They can be compared to those in Figure 2 to notice that the estimates are usually very similar.

#### *Model selection*

Given the large number of models fit to the data it is necessary to select the best model among all of them. The deviance information criterion (Spiegelhalter *et al.* 2002, DIC) and Watanabe-Akaike information criterion (Watanabe 2013, WAIC) have been computed when the models were fit and they are shown in Table 2.

Model	DIC		WAIC	
	Covariate		Covariate	
	No	Yes	No	Yes
Independent IMCAR	909.59	899.49	910.58	904.67
Independent PMCAR	910.59	897.87	903.75	895.61
IMCAR	904.26	894.96	905.96	897.38
PMCAR	905.47	893.80	898.34	888.76
M-model	911.47	893.30	917.71	898.17

Table 2: Information criteria computed for model selection for the North Carolina SIDS data.

According to the values obtained of the DIC and WAIC, it is clear that the covariate should be included in the model because it produces a decrease in the DIC and WAIC of about 10 for all models. Secondly, models with independent effects in general do not perform as well as the other models. This means that there is some within-area association between the different variables that needs to be accounted for.

### 3.2. Mortality in Comunidad Valenciana (Spain)

The next example is based on simulated data of the mortality by cirrhosis, lung and oral cancer in Comunidad Valenciana (Spain). This dataset has been obtained from [Martinez-Beneito and Botella-Rocamora \(2019\)](#) and it has been generated to mimic the spatial pattern of the real data, that cannot be provided due to confidentiality constraints. The original files are available at <https://github.com/MigueBeneito/DisMapBook>.

Here, the number of deaths by these three causes are available at the municipality level in Comunidad Valencia (Spain), as well as the expected number of cases that have been computed using internal standardization. Hence, the aim now is to estimate the spatial pattern of the different diseases as well as their possible correlations.

Given that in the previous example we have already described how to fit all the new models in the R package, we will focus now on the models that include a term to model the covariance for several diseases. That is, only models IMCAR, PMCAR and M-model will be fit.

Comunidad Valenciana data are available in the package as two separate objects: `CV` and `CV.nb`. `CV` is an object of class ‘SpatialPolygonsDataFrame’ ([Bivand \*et al.\* 2013](#)) that contains the data and boundaries of the municipalities in Comunidad Valenciana, while `CV.nb` is an object of class ‘nb’ ([Bivand and Wong 2018](#)) with the neighborhood structure. Both objects can be loaded as:

```
R> data("CV", package = "INLAMSM")
```

The adjacency matrix is computed using a sparse representation ([Bates and Maechler 2019](#)) as follows:

```
R> W <- as(nb2mat(CV.nb, style = "B"), "Matrix")
```

Next, a `data.frame` is created with the observed and expected data, together with an index variable to be passed to the definition of the latent effects and a disease-specific intercept:

```
R> d <- data.frame(OBS = c(CV$Obs.Cirrhosis, CV$Obs.Lung, CV$Obs.Oral),
+   EXP = c(CV$Exp.Cirrhosis, CV$Exp.Lung, CV$Exp.Oral))
R> d$Intercept <- rep(c("Cirrhosis", "Lung", "Oral"), each = nrow(W))
R> d$idx <- 1:length(d$OBS)
```

In order to fit the different models, the following parameters are defined:

```
R> k <- 3
R> alpha.min <- 0
R> alpha.max <- 1
```

Similarly to the SIDS example, some of the multivariate spatial effects require additional constraints. The matrices required are defined below:

```
R> A <- kronecker(Diagonal(k, 1), Matrix(1, ncol = nrow(W), nrow = 1))
R> e = rep(0, k)
```

Then, the latent effects are defined and the models are fit:

```
R> model <- inla.IMCAR.model(k = k, W = W)
R> IMCAR.cval <- inla(OBS ~ 0 + Intercept + f(idx, model = model,
+   extraconstr = list(A = as.matrix(A), e = e)),
+   data = d, E = EXP, family = "poisson",
+   control.compute = list(config = TRUE, dic = TRUE, waic = TRUE),
+   control.mode = list(theta = c(1, 1, 1, 0, 0, 0), restart = TRUE),
+   control.inla = list(h = 0.001),
+   control.predictor = list(compute = TRUE))
R> model <- inla.MCAR.model(k = k, W = W, alpha.min = alpha.min,
+   alpha.max = alpha.max)
R> PMCAR.cval <- inla(OBS ~ 0 + Intercept + f(idx, model = model), data = d,
+   E = EXP, family = "poisson",
+   control.compute = list(config = TRUE, dic = TRUE, waic = TRUE),
+   control.predictor = list(compute = TRUE))
R> model <- inla.Mmodel.model(k = k, W = W, alpha.min = alpha.min,
+   alpha.max = alpha.max)
R> Mmodel.cval <- inla(OBS ~ 0 + Intercept + f(idx, model = model), data = d,
+   E = EXP, family = "poisson",
+   control.compute = list(config = TRUE, dic = TRUE, waic = TRUE),
+   control.predictor = list(compute = TRUE))
```

Table 3 shows the computing times for the models fit to both examples. All models have been run on a Mac OS X computer with an Intel Core i5 processor (2,7 GHz), 4 cores and 16GB of RAM. Now the models take longer to run than in the previous example because there are three diseases and about 5 times more areas (as Comunidad Valenciana has 540 municipalities in the cartography that we have used). Botella-Rocamora *et al.* (2015) report times of about 16 minutes to fit the M-model with **WinBUGS** on this same dataset. Hence, **INLA** can fit the same models in a fraction of the time.

Dataset	# Diseases	# Areas	Model				
			Ind. IMCAR	Ind. PMCAR	IMCAR	PMCAR	M-model
NC SIDS	2	100	7.44	5.45	13.73	17.66	23.03
C. Valenciana	2	540	–	–	14.23	23.36	78.45
C. Valenciana	3	540	–	–	45.48	58.59	496.41

Table 3: Computing times (in seconds) for the models fit to the two examples.

In addition, Table 3 also shows the computing times of models fit to the Comunidad Valenciana dataset using two diseases (cirrhosis and lung cancer). These models have been computed as follows:

```
R> d2 <- subset(d, Intercept != "Oral")
R> d2$idx <- 1:nrow(d2)
R> k <- 2
R> A <- kronecker(Diagonal(k, 1), Matrix(1, ncol = nrow(W), nrow = 1))
R> e = rep(0, k)
R> model <- inla.IMCAR.model(k = k, W = W)
R> IMCAR.cval2 <- inla(OBS ~ 0 + Intercept + f(idx, model = model,
+   extraconstr = list(A = as.matrix(A), e = e)),
+   data = d2, E = EXP, family = "poisson",
+   control.compute = list(config = TRUE, dic = TRUE, waic = TRUE),
+   control.predictor = list(compute = TRUE))
R> model <- inla.MCAR.model(k = k, W = W, alpha.min = alpha.min,
+   alpha.max = alpha.max)
R> PMCAR.cval2 <- inla(OBS ~ 0 + Intercept + f(idx, model = model),
+   data = d2, E = EXP, family = "poisson",
+   control.compute = list(config = TRUE, dic = TRUE, waic = TRUE),
+   control.predictor = list(compute = TRUE))
R> model <- inla.Mmodel.model(k = k, W = W, alpha.min = alpha.min,
+   alpha.max = alpha.max)
R> Mmodel.cval2 <- inla(OBS ~ 0 + Intercept + f(idx, model = model),
+   data = d2, E = EXP, family = "poisson",
+   control.compute = list(config = TRUE, dic = TRUE, waic = TRUE),
+   control.predictor = list(compute = TRUE))
```

The fitted values (i.e., posterior means of the relative risks) can be added to the map in the ‘SpatialPolygonsDataFrame’ object as follows:

```
R> n <- nrow(W)
R> CV$IMCAR.CIR <- IMCAR.cval$summary.fitted[1:n, "mean"]
R> CV$IMCAR.LUN <- IMCAR.cval$summary.fitted[n + 1:n, "mean"]
R> CV$IMCAR.ORA <- IMCAR.cval$summary.fitted[2 * n + 1:n, "mean"]
R> CV$PMCAR.CIR <- PMCAR.cval$summary.fitted[1:n, "mean"]
R> CV$PMCAR.LUN <- PMCAR.cval$summary.fitted[n + 1:n, "mean"]
R> CV$PMCAR.ORA <- PMCAR.cval$summary.fitted[2 * n + 1:n, "mean"]
R> CV$Mmodel.CIR <- Mmodel.cval$summary.fitted[1:n, "mean"]
R> CV$Mmodel.LUN <- Mmodel.cval$summary.fitted[n + 1:n, "mean"]
R> CV$Mmodel.ORA <- Mmodel.cval$summary.fitted[2 * n + 1:n, "mean"]
```

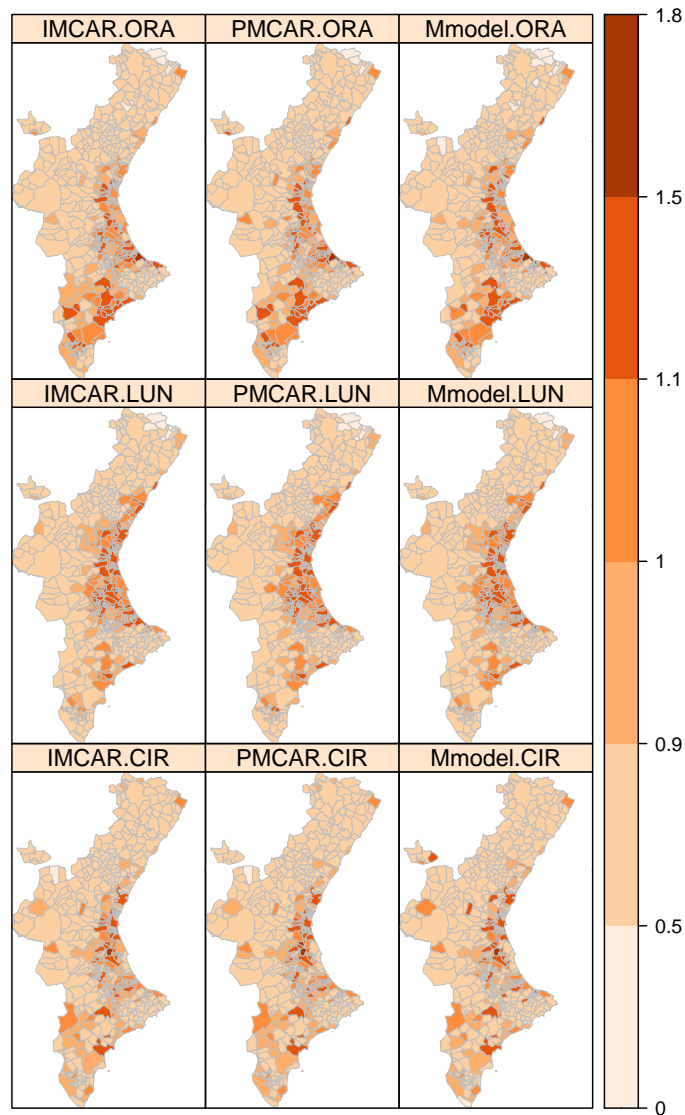


Figure 5: Posterior means of the relative risks of cirrhosis, lung cancer and oral cancer in Comunidad Valenciana (Spain).

The maps in Figure 5 show the different posterior means of the relative risks from the models fitted for the different causes of death. In general, the three models produce similar point estimates of the relative risks.

The **INLAMSM** package includes a few functions to transform the marginals and summary statistics of the model hyperparameters in the internal scale into the original scale in the model. For the models fitted, this back-transformation can be obtained as follows:

```
R> hyper.imcar <- inla.MCAR.transform(IMCAR.cval, 3)
R> hyper.pmcar <- inla.MCAR.transform(PMCAR.cval, 3, model = "PMCAR",
+   alpha.min = alpha.min, alpha.max = alpha.max)
R> hyper.mmodel <- inla.Mmodel.transform(Mmodel.cval, 3,
+   alpha.min = alpha.min, alpha.max = alpha.max)
```

These will provide a transformation of the model parameters so that they are not in the internal scale anymore and make inference easier. Spatial autocorrelation parameters are transformed to be in the range between  $\alpha_{\min}$  and  $\alpha_{\max}$ , and log-precisions are transformed to be variances. Correlation hyperparameters are transformed to be between  $-1$  and  $1$ .

Hence, summary statistics for the models are:

```
R> hyper.imcar$summary.hyperpar
```

	mean	sd	quant0.025	quant0.25	quant0.5	quant0.75
Theta1 for idx	0.243	0.0399	0.1745	0.215	0.240	0.268
Theta2 for idx	0.127	0.0161	0.0977	0.115	0.126	0.137
Theta3 for idx	0.226	0.0462	0.1492	0.193	0.221	0.254
Theta4 for idx	0.470	0.0824	0.2996	0.415	0.473	0.528
Theta5 for idx	0.559	0.0969	0.3510	0.497	0.566	0.628
Theta6 for idx	0.707	0.0634	0.5660	0.668	0.714	0.753
	quant0.975					
Theta1 for idx	0.331					
Theta2 for idx	0.161					
Theta3 for idx	0.330					
Theta4 for idx	0.622					
Theta5 for idx	0.729					
Theta6 for idx	0.813					

```
R> hyper.pmcars$summary.hyperpar
```

	mean	sd	quant0.025	quant0.25	quant0.5	quant0.75
Theta1 for idx	0.989	0.00633	0.974	0.986	0.991	0.994
Theta2 for idx	0.262	0.04351	0.190	0.231	0.257	0.288
Theta3 for idx	0.139	0.01805	0.106	0.126	0.138	0.151
Theta4 for idx	0.255	0.05188	0.169	0.217	0.249	0.286
Theta5 for idx	0.457	0.08193	0.290	0.403	0.459	0.515
Theta6 for idx	0.546	0.09598	0.341	0.484	0.552	0.614
Theta7 for idx	0.688	0.06531	0.547	0.647	0.693	0.735
	quant0.975					
Theta1 for idx	0.998					
Theta2 for idx	0.361					
Theta3 for idx	0.177					
Theta4 for idx	0.372					
Theta5 for idx	0.611					
Theta6 for idx	0.716					
Theta7 for idx	0.802					

```
R> hyper.mmodel$summary.hyperpar
```

	mean	sd	quant0.025	quant0.25	quant0.5	quant0.75
Theta1 for idx	0.2638	0.18431	0.0231	0.11257	0.2255	0.3810



Theta2 for idx	0.9963	0.00337	0.9873	0.99507	0.9973	0.9986
Theta3 for idx	0.9938	0.01251	0.9945	0.99995	1.0000	1.0000
Theta4 for idx	0.5245	0.05295	0.4234	0.48815	0.5232	0.5596
Theta5 for idx	0.1963	0.04075	0.1153	0.16882	0.1966	0.2238
Theta6 for idx	0.1597	0.06079	0.0381	0.11892	0.1605	0.2009
Theta7 for idx	0.0269	0.04639	-0.0626	-0.00479	0.0261	0.0578
Theta8 for idx	0.3022	0.02667	0.2473	0.28454	0.3033	0.3206
Theta9 for idx	0.1764	0.04496	0.0924	0.14519	0.1745	0.2059
Theta10 for idx	0.1278	0.09013	-0.0398	0.06508	0.1237	0.1868
Theta11 for idx	0.2478	0.05323	0.1417	0.21195	0.2483	0.2838
Theta12 for idx	0.3725	0.06317	0.2355	0.33197	0.3783	0.4178
	quant	0.975				
Theta1 for idx		0.692				
Theta2 for idx		1.000				
Theta3 for idx		1.000				
Theta4 for idx		0.631				
Theta5 for idx		0.275				
Theta6 for idx		0.276				
Theta7 for idx		0.119				
Theta8 for idx		0.352				
Theta9 for idx		0.268				
Theta10 for idx		0.313				
Theta11 for idx		0.350				
Theta12 for idx		0.480				

Note that the first hyperparameter in the PMCAR model is the spatial autocorrelation, which is very close to one. All the other parameters are the variances and correlation parameters.

In order to recover the variance matrix, the off-diagonal entries need to be computed. Note that these depend on three parameters (i.e., correlation and marginal variances) and that computing these entries involves multivariate inference. As the joint posterior distribution of these three parameters needs to be estimated sampling will be used.

For approximate multivariate posterior inference, **INLA** can draw samples from the (approximate) joint posterior of the hyperparameters using function `inla.posterior.sample`. This sampling method is based on the internal representation of the model, which is based on values of the ensemble of the hyperparameters and their posterior log-densities (see, for example, [Gómez-Rubio 2020](#), for details).

The internal representation of the model stores different ensembles of values of the hyperparameters  $\{\gamma_g\}_{g=1}^G$  and associated values of the log-posterior density. Instead of sampling with `inla.posterior.sample` we will re-scale the log-posterior densities to obtain weights associated to  $\{\gamma_g\}_{g=1}^G$ . By avoiding sampling we obtain short computing times here. These weights are the posterior probabilities of the ensembles of values of the hyperparameters and they can be used to compute posterior quantities of interest. Furthermore, posterior estimates of transformations of the hyperparameters can be computed as well. Note that, because the ensemble of values is available, multivariate inference on the hyperparameters is also possible. We will rely on this fact to estimate some of the quantities of interest such as the posterior mean of the between-diseases variance matrix.

Model	DIC	WAIC
IMCAR	7407.59	7396.61
PMCAR	7409.99	7388.38
M-model	7384.90	7366.27

Table 4: Information criteria computed for model selection for the Comunidad Valenciana data.

For the IMCAR and PMCAR models, the posterior means of the entries of the between-diseases variance matrix are:

```
R> hyper.imcar$VAR.m

      [,1] [,2] [,3]
[1,] 0.2449 0.0811 0.129
[2,] 0.0811 0.1269 0.118
[3,] 0.1294 0.1178 0.223
```

```
R> hyper.pmcars$VAR.m

      [,1] [,2] [,3]
[1,] 0.2655 0.0874 0.140
[2,] 0.0874 0.1412 0.129
[3,] 0.1395 0.1288 0.255
```

For the M-model, the variance of the between diseases variance is given by  $M^T M$ :

```
R> hyper.mmodel$VAR.m

      [,1] [,2] [,3]
[1,] 0.349 0.106 0.179
[2,] 0.106 0.129 0.151
[3,] 0.179 0.151 0.247
```

Point estimates of the variance matrix are very similar for the PMCAR and the M-model. In addition, all models seem to point to a higher correlation between lung cancer (2nd disease in the model) and oral cancer (3rd disease) than cirrhosis (1st disease) with any of the two other diseases. This makes sense as lung and oral cancer are known to be highly correlated (Botella-Rocamora *et al.* 2015).

Finally, the different model selection criteria computed for the models (DIC and WAIC) can be compared to choose the best model. The values of these criteria are shown in Table 4. In this case, both criteria point to the M-model as the best model in this example.

## 4. Discussion

The **INLAMSM** package builds on top of the **INLA** package and implements a number of multivariate spatial latent effects. Hence, this package allows an easy and simple definition of

these multivariate effects to be used within a formula term to fit multivariate spatial models to lattice data.

Implementation of new latent effects for multivariate data is straightforward with the `rgeneric` latent effect included in the **INLA** package. This only requires the specification of the latent effect as a GMRF, which means that the mean, precision matrix and priors for the hyperparameters need to be provided. Once the model is implemented, it is easy to include it in the model formula to fit multivariate models with **INLA**.

We find important to mention that although all the models in **INLAMSM** have been defined through multivariate ensembles of spatial processes, **INLAMSM** could be also used for reproducing some multivariate spatial models defined through univariate or multivariate conditional distributions, which induce cross-correlations between the different spatial patterns (Sain and Cressie 2002). As shown, a close connection can be drawn between both approaches (Martinez-Beneito 2019) and some important classes of conditional spatial multivariate models can be reformulated as (marginal) multivariate models, as those reproduced in **INLAMSM**. Therefore, this package could be used as a starting point for multivariate models formulated as ensembles of conditional distributions

In addition, the latent models implemented in the package can be used as templates to implement new multivariate spatial models for lattice data. In the future, we hope to increase the number of multivariate spatial models in the package to include other different types of spatial cross-correlation.

In the examples provided to illustrate the use of the package we have considered a small dataset to fit all possible models. Times required to fit the models are short. The second example deals with a region with a larger number of areas which shows that our package can be used together with **INLA** to fit multivariate models.

Despite our focus on multivariate spatial models for disease mapping, it is worth mentioning that the multivariate models implemented in the **INLAMSM** package can be used in other contexts. Furthermore, these multivariate spatial models can also be used to build temporal and spatio-temporal models by using a temporal adjacency matrix.

## Acknowledgments

This work has been supported by grants PPIC-2014-001-P and SBPLY/17/180501/000491, funded by Consejería de Educación, Cultura y Deportes (Junta de Comunidades de Castilla-La Mancha, Spain) and FEDER, grant MTM2016-77501-P, funded by Ministerio de Economía y Competitividad (Spain), grant PID2019-106341GB-I00 from Ministerio de Ciencia e Innovación (Spain) and a grant to support research groups by the University of Castilla-La Mancha (Spain). F. Palmí-Perales has been supported by a Ph.D. scholarship awarded by the University of Castilla-La Mancha (Spain).

## References

- Bachl FE, Lindgren F, Borchers DL, Illian JB (2019). “**inlabru**: An R Package for Bayesian Spatial Modelling from Ecological Survey Data.” *Methods in Ecology and Evolution*, **10**, 760–766. doi:10.1111/2041-210X.13168.

- Baddeley A, Turner R (2005). “**spatstat**: An R Package for Analyzing Spatial Point Patterns.” *Journal of Statistical Software*, **12**(6), 1–42. doi:10.18637/jss.v012.i06.
- Bakka H, Rue H, Fuglstad GA, Riebler A, Bolin D, Illian J, Krainski E, Simpson D, Lindgren F (2018). “Spatial Modeling with R-INLA: A Review.” *WIREs Computational Statistics*, **10**(6), e1443. doi:10.1002/wics.1443.
- Bates D, Maechler M (2019). **Matrix**: *Sparse and Dense Matrix Classes and Methods*. R package version 1.2-18, URL <https://CRAN.R-project.org/package=Matrix>.
- Besag J, Kooperberg C (1995). “On Conditional and Intrinsic Autoregression.” *Biometrika*, **82**(4), 733–746. doi:10.2307/2337341.
- Besag J, York L, Mollié A (1991). “Bayesian Image Restoration with Two Applications in Spatial Statistics.” *Annals of the Institute of Statistical Mathematics*, **43**, 1–59. doi:10.1007/BF00116466.
- Bivand R, Keitt T, Rowlingson B (2019a). **rgdal**: *Bindings for the ‘Geospatial’ Data Abstraction Library*. R package version 1.4-4, URL <https://CRAN.R-project.org/package=rgdal>.
- Bivand R, Nowosad J, Lovelace R (2019b). **spData**: *Datasets for Spatial Analysis*. R package version 0.3.0, URL <https://CRAN.R-project.org/package=spData>.
- Bivand R, Wong DWS (2018). “Comparing Implementations of Global and Local Indicators of Spatial Association.” *TEST*, **27**(3), 716–748. doi:10.1007/s11749-018-0599-x.
- Bivand RS, Pebesma EJ, Gómez-Rubio V (2013). *Applied Spatial Data Analysis with R*. 2nd edition. Springer-Verlag, New York.
- Botella-Rocamora P, Martínez-Beneito MA, Banerjee S (2015). “A Unifying Modeling Framework for Highly Multivariate Disease Mapping.” *Statistics in Medicine*, **34**(9), 1548–1559. doi:10.1002/sim.6423.
- Carlin BP, Banerjee S (2003). “Hierarchical Multivariate CAR Models for Spatio-Temporally Correlated Survival Data.” *Bayesian Statistics*, **7**(7), 45–63.
- Carpenter B, Gelman A, Hoffman M, Lee D, Goodrich B, Betancourt M, Brubaker M, Guo J, Li P, Riddell A (2017). “**Stan**: A Probabilistic Programming Language.” *Journal of Statistical Software, Articles*, **76**(1), 1–32. doi:10.18637/jss.v076.i01.
- Corpas-Burgos F, Botella-Rocamora P, Martínez-Beneito MA (2019). “On the Convenience of Heteroscedasticity in Highly Multivariate Disease Mapping.” *TEST*, **28**(4), 1229–1250. doi:10.1007/s11749-019-00628-8.
- Cressie N, Read TRC (1985). “Do Sudden Infant Deaths Come in Clusters?” *Statistics and Decisions*, **Supplement Issue 2**, 333–349.
- Finley AO, Banerjee S, Carlin BP (2007). “**spBayes**: An R Package for Univariate and Multivariate Hierarchical Point-Referenced Spatial Models.” *Journal of Statistical Software*, **19**(4), 1. doi:10.18637/jss.v019.i04.

- Gelfand AE, Vounatsou P (2003). “Proper Multivariate Conditional Autoregressive Models for Spatial Data Analysis.” *Biostatistics*, **9**, 11–25. doi:10.1093/biostatistics/4.1.11.
- Gelman A (2006). “Prior Distributions for Variance Parameters in Hierarchical Models.” *Bayesian Analysis*, **1**(3), 515–534. doi:10.1214/06-BA117A.
- Gilks WR, Gilks WR, Richardson S, Spiegelhalter DJ (1996). *Markov Chain Monte Carlo in Practice*. Chapman & Hall/CRC, Boca Raton.
- Gómez-Rubio V (2020). *Bayesian Inference with INLA*. Chapman & Hall/CRC, Boca Raton. URL <https://becarioprecario.bitbucket.io/inla-gitbook/>.
- Gómez-Rubio V, Zheng P, Diggle P, Sterratt DC, Peng RD, Murdoch D, Rowlingson B (2017). *spatialkernel: Non-Parametric Estimation of Spatial Segregation in a Multivariate Point Process*. R package version 0.4-23, URL <https://CRAN.R-project.org/package=spatialkernel>.
- Hodges JS, Reich BJ (2010). “Adding Spatially-Correlated Errors Can Mess Up the Fixed Effect You Love.” *The American Statistician*, **64**(4), 325–334. doi:10.1198/tast.2010.10052.
- Jin X, Banerjee S, Carlin BP (2007). “Order-Free Co-Regionalized Areal Data Models with Application to Multiple-Disease Mapping.” *Journal of the Royal Statistical Society B*, **69**(5), 817–838. doi:10.1111/j.1467-9868.2007.00612.x.
- Krainski ET, Gómez-Rubio V, Bakka H, Lenzi A, Castro-Camilo D, Simpson D, Lindgren F, Rue H (2018). *Advanced Spatial Modeling with Stochastic Partial Differential Equations Using R and INLA*. Chapman & Hall/CRC, Boca Raton.
- Lee D (2013). “CARBayes: An R Package for Bayesian Spatial Modeling with Conditional Autoregressive Priors.” *Journal of Statistical Software*, **55**(13), 1–24. doi:10.18637/jss.v055.i13.
- Leroux BG, Lei X, Breslow N (2000). “Estimation of Disease Rates in Small Areas: A New Mixed Model for Spatial Dependence.” In *Statistical Models in Epidemiology, the Environment, and Clinical Trials*, pp. 179–191. Springer-Verlag, New York. doi:10.1007/978-1-4612-1284-3\_4.
- Lindgren F, Rue H (2015). “Bayesian Spatial Modelling with R-INLA.” *Journal of Statistical Software*, **63**(19), 1–25. doi:10.18637/jss.v063.i19.
- Lindgren F, Rue H, Lindström J (2011). “An Explicit Link between Gaussian Fields and Gaussian Markov Random Fields: The Stochastic Partial Differential Equation Approach.” *Journal of the Royal Statistical Society B*, **73**(4), 423–498. doi:10.1111/j.1467-9868.2011.00777.x.
- Lunn DJ, Spiegelhalter DJ, Thomas A, Best N (2009). “The BUGS Project: Evolution, Critique and Future Directions.” *Statistics in Medicine*, **28**(25), 3049–3067. doi:10.1002/sim.3680.

- Lunn DJ, Thomas A, Best N, Spiegelhalter DJ (2000). “WinBUGS – A Bayesian Modelling Framework: Concepts, Structure, and Extensibility.” *Statistics and computing*, **10**(4), 325–337. doi:10.1023/A:1008929526011.
- MacNab YC (2018). “Some Recent Work on Multivariate Gaussian Markov Random Fields.” *TEST*, **27**(3), 497–541. doi:10.1007/s11749-018-0605-3.
- Martinez-Beneito MA (2013). “A General Modelling Framework for Multivariate Disease Mapping.” *Biometrika*, **100**(3), 539–553. doi:10.1093/biomet/ast023.
- Martinez-Beneito MA (2019). “Some Links between Conditional and Coregionalized Multivariate Gaussian Markov Random Fields.” *Spatial Statistics*, p. 100383. doi:10.1016/j.spasta.2019.100383.
- Martinez-Beneito MA, Botella-Rocamora P (2019). *Disease Mapping: From Foundations to Multidimensional Modeling*. Chapman & Hall/CRC, Boca Raton.
- Morris M, Wheeler-Martin K, Simpson D, Mooney SJ, Gelman A, DiMaggio C (2019). “Bayesian Hierarchical Spatial Models: Implementing the Besag York Mollié Model in Stan.” *Spatial and Spatio-temporal Epidemiology*, **31**, 100301. doi:10.1016/j.sste.2019.100301.
- Neuwirth E (2014). **RColorBrewer: ColorBrewer Palettes**. R package version 1.1-2, URL <https://CRAN.R-project.org/package=RColorBrewer>.
- Palmí-Perales F, Gómez-Rubio V, Martínez-Beneito MA (2021). **INLAMSM: Multivariate Spatial Models with INLA**. R package version 0.2-3, URL <https://CRAN.R-project.org/package=INLAMSM>.
- Pebesma EJ (2004). “Multivariable Geostatistics in S: The **gstat** Package.” *Computers & Geosciences*, **30**(7), 683–691. doi:10.1016/j.cageo.2004.03.012.
- Pebesma EJ, Wesseling CG (1998). “**gstat**: A Program for Geostatistical Modelling, Prediction and Simulation.” *Computers & Geosciences*, **24**(1), 17–31. doi:10.1016/S0098-3004(97)00082-4.
- Rue H, Martino S, Chopin N (2009). “Approximate Bayesian Inference for Latent Gaussian Models by Using Integrated Nested Laplace Approximations.” *Journal of the Royal Statistical Society B*, **71**(2), 319–392. doi:10.1111/j.1467-9868.2008.00700.x.
- Rue H, Riebler AI, Sørbye SH, Illian JB, Simpson DP, Lindgren FK (2017). “Bayesian Computing with INLA: A Review.” *Annual Reviews of Statistics and Its Applications*, **4**(March), 395–421. doi:10.1146/annurev-statistics-060116-054045.
- Sain SR, Cressie N (2002). “Multivariate Lattice Models for Spatial Environmental Data.” In *Proceedings of the Joint Statistical Meetings, Section on Statistics and the Environment*, pp. 2820–2825. American Statistical Association.
- Simpson D, Illian JB, Sørbye SH, Rue H (2016). “Going Off Grid: Computationally Efficient Inference for Log-Gaussian Cox Processes.” *Biometrika*, **1**(103), 49–70. doi:10.1093/biomet/asv064.

- Spiegelhalter DJ, Best NG, Carlin BP, Van der Linde A (2002). “Bayesian Measures of Model Complexity and Fit.” *Journal of the Royal Statistical Society B*, **64**(4), 583–616. doi:10.1111/1467-9868.00353.
- Stan Development Team (2019). “**RStan**: R interface to Stan.” R package version 2.19.2, URL <http://mc-stan.org/>.
- Sturtz S, Ligges U, Gelman A (2005). “**R2WinBUGS**: A Package for Running WinBUGS from R.” *Journal of Statistical Software*, **12**(3), 1–16. doi:10.18637/jss.v012.i03.
- Sun D, Tsutakawa RK, Speckman PL (1999). “Posterior Distribution of Hierarchical Models Using CAR(1) Distributions.” *Biometrika*, **86**(2), 341–350. doi:10.1093/biomet/86.2.341.
- Wall MM (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. doi:10.1016/S0378-3758(03)00111-3.
- Watanabe S (2013). “A Widely Applicable Bayesian Information Criterion.” *Journal of Machine Learning Research*, **14**, 867–897. URL <https://jmlr.csail.mit.edu/papers/v14/watanabe13a.html>.

### Affiliation:

Francisco Palmí-Perales, Virgilio Gómez-Rubio  
Departamento de Matemáticas  
Escuela Técnica Superior de Ingenieros Industriales de Albacete  
Universidad de Castilla-La Mancha  
Avda. España, s/n  
02002, Albacete, Spain  
E-mail: [Francisco.Palmi@uclm.es](mailto:Francisco.Palmi@uclm.es), [Virgilio.Gomez@uclm.es](mailto:Virgilio.Gomez@uclm.es)  
URL: <https://becarioprecario.github.io/>

Miguel A. Martínez-Beneito  
Departamento de Estadística i Investigació Operativa  
Universitat de València  
Doctor Moliner, 50  
46100, Burjassot (Valencia), Spain  
E-mail: [miguel.a.martinez@uv.es](mailto:miguel.a.martinez@uv.es)  
URL: [https://www.uv.es/~mamtnez/inicio\\_es.html](https://www.uv.es/~mamtnez/inicio_es.html)