



Published in final edited form as:

*Bayesian Anal.* 2017 March ; 12(1): 239–259. doi:10.1214/16-BA995.

## Towards a Multidimensional Approach to Bayesian Disease Mapping

Miguel A. Martinez-Beneito<sup>\*</sup>, Paloma Botella-Rocamora<sup>†</sup>, and Sudipto Banerjee<sup>‡</sup>

<sup>\*</sup>Fundación para el Fomento de la Investigación Sanitaria y Biomédica (FISABIO-Salud Pública), Valencia, Spain and CIBER de Epidemiología y Salud Pública, Madrid, Spain

<sup>†</sup>Subdirección Gral.de Epidemiología y Vigilancia de la Salud – Conselleria de Sanitat, Valencia, Spain and Fundación para el Fomento de la Investigación Sanitaria y Biomédica (FISABIO-Salud-Pública), Valencia, Spain

<sup>‡</sup>University of California, Los Angeles (UCLA), USA

### Abstract

Multivariate disease mapping enriches traditional disease mapping studies by analysing several diseases jointly. This yields improved estimates of the geographical distribution of risk from the diseases by enabling borrowing of information across diseases. Beyond multivariate smoothing for several diseases, several other variables, such as sex, age group, race, time period, and so on, could also be jointly considered to derive multivariate estimates. The resulting multivariate structures should induce an appropriate covariance model for the data. In this paper, we introduce a formal framework for the analysis of multivariate data arising from the combination of more than two variables (geographical units and at least two more variables), what we have called Multidimensional Disease Mapping. We develop a theoretical framework containing both separable and non-separable dependence structures and illustrate its performance on the study of real mortality data in Comunitat Valenciana (Spain).

### 1 Introduction

Areally-referenced spatial data arise frequently in epidemiological studies seeking to describe the geographical distribution of diseases over a region of study. Disease maps describe the geographic variation of disease and generate etiological hypotheses about the possible causes for apparent differences in disease risk. They can also be used to detect spatial clusters attributable to common environmental, demographical, or cultural effects shared by neighbouring regions. However, mapping crude rates can be misleading when the population sizes for some of the geographical units are small and result in excessive variation in the estimated rates, which make the traditional epidemiological risk estimates unreliable. Statistical models built specifically for analysing datasets over small areas are required for exhibiting clearer patterns in the geographical distribution of the diseases. They

Supplementary Material

Supplementary material of “Towards a Multidimensional Approach to Bayesian Disease Mapping” (DOI: 10.1214/16-BA995SUPP; .zip).

allow us to borrow strength across regions by using not only the data from a given region, but also the data from neighbouring regions, thereby increasing the amount of information used for estimating the risks in each unit. Univariate models account for information on a single disease, while multivariate models enable us to reliably estimate the geographical distribution of the risks corresponding to several diseases over a region of study; see, e.g. Jin et al. (2007). Nevertheless, correlations across diseases may arise, for example, from common sets of (spatially distributed) risk factors. Multivariate models can permit modelling of dependence among diseases while capturing spatial dependence between regions. Estimating the joint spatial distribution for multiple diseases will reflect better the underlying risks than would be available from the analysis of any single disease separately using univariate models. See, for example, Dobra et al. (2011); Macnab (2011); Marí Dell'Olmo et al. (2014) for recent contributions in this field.

Recently, Martinez-Beneito (2013) proposed a versatile framework colligating a variety of multivariate disease mapping models arising from Gaussian Markov Random Field (GMRF) models with separable and non-separable covariance structures. Furthermore, Martinez-Beneito (2013) meld different spatial dependence patterns with different covariance models between diseases, producing a large number of models as special cases. A further modification by Botella-Rocamora et al. (2015) accrues substantial computational benefits. This enables joint modelling for a larger collection of diseases (tens of them) and integrates information from the spatial patterns associated with each disease.

The multivariate disease mapping literature has presented models with just two factors—disease types and geographical units. Hereafter, we will use the classical terminology *factor* and *levels* to denote, respectively, categorical variables and the different values that they can take. We are aware of only two articles dealing with more than two factors. Zhang et al. (2006) considered a separable dependence structure with four factors: (i) time period, (ii) sex, (iii) age group, and (iv) geographical unit for studying the incidence of colorectal cancer. Tzala and Best (2008) studied three factors: (i) disease, (ii) time period, and (iii) geographical unit in studying gastric cancers in Greece. Although these two papers are examples of multivariate studies with more than two factors, they deploy specific models suitable for the data in their papers, i.e. they do not set any theoretical framework for the joint study of geographical patterns defined by the combination of three or more factors. The current article seeks to generalize the framework of Martinez-Beneito (2013) and Botella-Rocamora et al. (2015) to more than two factors (geographical units and at least two more factors). We refer to this as *multidimensional modelling*, in contrast to the more common two-factor *multivariate modelling*. We will introduce some general guidelines intended to be useful in multidimensional studies instead of introducing a particular model to be used in some specific dataset.

This paper is organized as follows: Section 2 introduces some basic tensor algebra that will be later used for building up the models in the rest of sections. Section 3 shows how to generalize the separable multivariate modelling proposal to the multidimensional case. Section 4 introduces non-separability into the multidimensional context and describes the high number of models that arise when separability is no longer assumed. Section 5 shows two examples illustrating multidimensional modelling in a real setting. First, we show on a

trivariate study example how separability can be a restrictive assumption in some cases and how we could use the theory introduced in the former sections to overcome the separability assumption. Second, we undertake a four-dimensional study considering two unstructured factors (Disease and Sex) and two structured ones (Geographical unit and Period). Finally, Section 5 contains some conclusions about the results and models developed in the previous sections.

## 2 Basic tensor algebra

Let  $\mathcal{X}$  an  $N$ th order *tensor*, or array, of dimensions  $(L_1, L_2, \dots, L_N)$ . The *vector unfolding* of  $\mathcal{X}$  into the column vector  $\mathbf{x} = (x_{1\dots 1}, x_{2\dots 1}, \dots, x_{L_1\dots 1}, \dots, x_{L_1\dots L_N})'$  is denoted by  $\text{vec}(\mathcal{X})$ . Similarly, reordering the elements of a tensor into a matrix will be called *matrix unfolding*. Thus, the *n-dimensional matrix unfolding* of the tensor  $\mathcal{X}$ , denoted by  $\mathbf{X}^{(n)}$ , is just the  $L_n \times \prod_{i=1, i \neq n}^N L_i$  matrix  $[\mathbf{x}_{1\dots 1 \cdot 1\dots 1} : \mathbf{x}_{2\dots 1 \cdot 1\dots 1} : \dots : \mathbf{x}_{L_1\dots 1 \cdot 1\dots 1} : \dots : \mathbf{x}_{L_1\dots L_{n-1} \cdot L_{n+1}\dots L_N}]$ , where  $\mathbf{x}_{i_1\dots i_{n-1} \cdot i_{n+1}\dots i_N}$  with “ $\cdot$ ” in the  $n$ th position is the  $L_n \times 1$  vector with entries  $x_{i_1\dots i_{n-1} \cdot j \cdot i_{n+1}\dots i_N}$  for  $j = 1, 2, \dots, L_n$ .

More generally, let  $\mathbf{a} = \{1, \dots, N_{|\mathbf{a}|}\}$  be a subset of  $\{1, \dots, N\}$  for some integer  $N_{|\mathbf{a}|} < N$ . The  *$\mathbf{a}$ -matrix unfolding* of the tensor  $\mathcal{X}$ , denoted by  $\mathbf{X}^{(\mathbf{a})}$ , is the  $(\prod_{i \in \mathbf{a}} L_i) \times (\prod_{j \in \{1, \dots, N\} \setminus \mathbf{a}} L_j)$  matrix formed by stacking the column-vectors

$$\{\text{vec}(\mathbf{x}_{\dots i_{\alpha+1} \dots i_N}); i_{\alpha+1} = 1, \dots, L_{\alpha+1}; \dots; i_N = 1, \dots, L_N\}.$$

The  *$\mathbf{a}$ -matrix unfolding* can further be generalized to any set of indices by simply performing a permutation of the indices before applying the above definition. We will refer to the inverse process of unfolding a tensor as *folding*.

The *n-dimensional product* of a tensor  $\mathcal{X}$  with an  $L_n \times L_n$  matrix  $\mathbf{A}$  is defined as the tensor  $\mathcal{Y}$  resulting from folding  $\mathbf{A}\mathbf{X}^{(n)}$  into a tensor of the same dimension as  $\mathcal{X}$ . This is denoted using  $\circ_n$  as follows:

$$\mathcal{Y} = \mathbf{A} \circ_n \mathcal{X} \leftrightarrow \mathbf{Y}^{(n)} = \mathbf{A}\mathbf{X}^{(n)}. \quad (1)$$

We also generalize the *n-dimensional product* analogous to how we defined the  *$\mathbf{a}$ -matrix unfolding* as follows. Given an index set  $\mathbf{a}$ , if  $\mathbf{A}$  is a  $(\prod_{i \in \mathbf{a}} L_i) \times (\prod_{j \in \mathbf{a}} L_j)$  matrix, then the tensor obtained from folding  $\mathbf{A}\mathbf{X}^{(\mathbf{a})}$  into a tensor with the same dimensions as of  $\mathcal{X}$  will be referred to as the  *$\mathbf{a}$ -product* of a tensor  $\mathcal{X}$  with  $\mathbf{A}$ . We denote this product as  $\mathbf{A} \circ_{\mathbf{a}} \mathcal{X}$ .

## 3 A fully-separable multidimensional proposal

Let us elucidate further with the example of a trivariate setting, which presents all the challenges in the multidimensional approach. Therefore, for easier exposition, we restrict our attention to the trivariate setting and, when required, point out any specific complexities

of models with more than three factors. Here, we are interested in modelling the spatial distribution of risks for several combinations of two factors. The first factor in this setting will always be the geographical unit, while one of the other two factors will usually be the disease (from a set of diseases) and the third factor may either be unstructured, such as Sex or Race, or structured in some way such as Time period or Age group. The spatial term may also be considered as a special case of a structured factor. Let  $O_{ijk}$  and  $E_{ijk}$  denote, respectively, the number of observed and expected outcomes for the  $i$ th geographical unit of study and for the specific combination of the other two factors in the study, indexed with subindexes  $j$  and  $k$ . In disease mapping, one customarily assumes that  $O_{ijk} \sim Po(E_{ijk}R_{ijk})$  for  $i = 1, \dots, I, j = 1, \dots, J$  and  $k = 1, \dots, K$ , where  $R_{ijk}$ , the relative risk for the  $i$ th geographical unit and  $(j, k)$  values for the second and third factors in the study, satisfies  $\log(R_{ijk}) = \mu_{jk} + \theta_{ijk}$ . We let the  $\mu_{jk}$ 's represent a set of intercepts for the different combinations of the second and third factor (not geographic unit) of study. If so desired, these intercepts could also be modelled as a function of one (or more) covariate(s) in order to explain the geographical variability associated with it (them). Modelling  $\theta_{ijk}$ , the term inducing dependence between relative risks, in a rich, flexible and computationally efficient manner constitutes the main goal of multidimensional modelling and, hence, of the current article.

### 3.1 The fully-separable model

Let  $\mathbf{X}$  and  $\mathcal{X}$  be a matrix and an array of independent Gaussian random variables, respectively, which will be used as support for defining the  $\theta_{ijk}$ s as some specific transformations of them. Those transformation will induce suitable dependence relationships on the relative risks.

Consider the expression in (1). This expression, when applied to the first dimension of  $\mathbf{X}$ , yields  $\mathbf{A} \circ_1 \mathbf{X} = \mathbf{AX}^{(1)} = \mathbf{AX}$ . Similarly, for its second dimension, we obtain  $\mathbf{B} \circ_2 \mathbf{X} = \mathbf{BX}^{(2)} = \mathbf{XB}'$ . Consequently, the matrix expression  $\mathbf{AXB}'$  can be expressed as

$$\mathbf{AXB}' = (\mathbf{AX})\mathbf{B}' = \mathbf{B} \circ_2 \mathbf{A} \circ_1 \mathbf{X}. \quad (2)$$

Alternatively, the associative property of matrix products further yields

$$\mathbf{AXB}' = \mathbf{A}(\mathbf{XB}') = \mathbf{A} \circ_1 \mathbf{B} \circ_2 \mathbf{X}. \quad (3)$$

This extends easily to produce the following for a general  $n$ -dimensional array:

$$\text{vec}(\mathbf{A} \circ_n \mathcal{X}) = (\mathbf{I}_{L_N} \otimes \dots \otimes \mathbf{I}_{L_{n+1}} \otimes \mathbf{A} \otimes \mathbf{I}_{L_{n-1}} \otimes \dots \otimes \mathbf{I}_{L_1})\text{vec}(\mathcal{X}). \quad (4)$$

The expression  $\mathbf{AXB}'$  was the key starting point in the multivariate disease mapping work of Martinez-Beneito (2013). It also forms the basis for our multidimensional proposal. The matrices  $\mathbf{A}$  and  $\mathbf{B}$ , when applied to a Gaussian random noise matrix  $\mathbf{X}$ , induce, respectively,

dependence between diseases and within diseases into the model. Since each row of  $\mathbf{AX}$  is a linear combination of the rows of  $\mathbf{X}$ ,  $\mathbf{A}$  introduces spatial dependence among geographical units. Similarly,  $\mathbf{B}$  combines information across the different columns of  $\mathbf{X}$  and introduces dependence between the different diseases. This interpretation is apparent from expression (2)—both  $\mathbf{A}$  and  $\mathbf{B}$  represent identical operations on the different dimensions of the random matrix  $\mathbf{X}$ .

Expression (2) also tells us how to generalize the approach in Martinez-Beneito (2013) to the multidimensional context. Adapting to the trivariate case, we obtain

$$\boldsymbol{\theta} = \mathbf{M}_3 \circ_3 \mathbf{M}_2 \circ_2 \mathbf{M}_1 \circ_1 \mathcal{X}, \quad (5)$$

where  $\mathbf{M}_1$ ,  $\mathbf{M}_2$  and  $\mathbf{M}_3$  are matrices of suitable dimensions inducing dependence on  $\boldsymbol{\theta}$  along each of the three factors considered in the model and  $\mathcal{X}_{ijk} \sim \mathcal{N}(0, 1) \forall i, j, k$ . Applying (4) successively to the previous expression, we easily obtain

$$\text{vec}(\boldsymbol{\theta}) = (\mathbf{M}_3 \otimes \mathbf{M}_2 \otimes \mathbf{M}_1) \text{vec}(\mathcal{X})$$

and, therefore,

$$\text{vec}(\boldsymbol{\theta}) \sim \mathcal{N}(\mathbf{0}, (\mathbf{M}_3 \mathbf{M}_3') \otimes (\mathbf{M}_2 \mathbf{M}_2') \otimes (\mathbf{M}_1 \mathbf{M}_1')).$$

In this manner, we can easily build a fully separable dependence structure for all the factors considered, by means of successive  $\circ_i$  operations, with different values of  $i$ . Applying (2) and (3) to the trivariate case in (5) reveals that the order in which dependence on the factors are introduced in the separable case is irrelevant; they all yield identical models. Thus, for the separable case, the introduction of dependence on the different factors is a commutative operation.

Let us now turn to the definition of the  $\mathbf{M}_i$ 's. For any of the factors considered in the analysis, we will distinguish between those factors having a completely unstructured covariance matrices (Sex, Race, Disease, ...), i.e. a general symmetric positive definite matrix, as opposed to those having some kind of structure (time, ordered factor, ...). For example, we may want to account for an ordinal structure for Age group or Time period and for the obvious spatial arrangement (neighbourhood structure) for the set of geographical units. For any unstructured factor, such as a set of causes for mortality, Botella-Rocamora et al. (2015) suggests that a reasonable modelling choice for  $\mathbf{M}$  is to assume that each of its elements follow  $\mathcal{N}(0, \sigma^2)$  for a suitable value of  $\sigma^2$ . On the other hand, for a structured factor, the dependence arising from the structure should be incorporated within the corresponding  $\mathbf{M}_i$ . Here,  $\mathbf{M}_i \mathbf{M}_i'$  is the covariance matrix for the elements of that factor and  $\mathbf{M}_i$  is chosen so as to yield the desired covariance matrix. For example, in a time-structured

factor, such as the years within the period of study, we could be interested in considering a first-order autoregressive structure. The entries in the covariance matrix among years is given by

$$\Sigma_{i,j} = \sigma^2 \frac{\rho^{|i-j|}}{1-\rho^2},$$

so the corresponding  $\mathbf{M}_j$  matrix inducing temporal dependence on the observations could be, for example, the lower-triangular Cholesky square-root of  $\Sigma$ ,

$$\sigma \begin{pmatrix} (1-\rho^2)^{-1/2} & 0 & 0 & \dots & 0 \\ \rho(1-\rho^2)^{-1/2} & 1 & 0 & \dots & 0 \\ \rho^2(1-\rho^2)^{-1/2} & \rho & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{J-1}(1-\rho^2)^{-1/2} & \rho^{J-2} & \rho^{J-3} & \dots & 1 \end{pmatrix}. \quad (6)$$

Any other decomposition of the form  $\Sigma = \mathbf{M}_i \mathbf{M}_i'$  can, in principle, be used, although the Cholesky decomposition is numerically more stable and relatively efficient (Banerjee and Roy, 2014).

Spatial dependence is introduced in a similar manner with  $\Sigma$  being a suitable spatial covariance matrix. Some attention to the size of  $\Sigma$  is usually warranted; otherwise, model fitting becomes infeasible with larger covariance matrices. As argued by Martinez-Beneito (2013), we can conveniently recast (5) as

$$\boldsymbol{\theta} = \mathbf{M}_3 \circ_3 \mathbf{M}_2 \circ_2 \mathcal{Y}, \quad (7)$$

where the spatial dependence on the first dimension of  $\mathcal{Y} = \mathbf{M}_1 \circ_1 \mathcal{X}$  has already been incorporated, i.e.

$$\text{vec}(\mathcal{Y}) \sim N(\mathbf{0}, \mathbf{I}_{L_3} \otimes \mathbf{I}_{L_2} \otimes (\mathbf{M}_1 \mathbf{M}_1')).$$

Unfortunately, covariance structures arising from tensor products may not be identifiable. For example, in the trivariate setting,  $\mathbf{M}_1$ ,  $\lambda \mathbf{M}_2$  and  $\lambda^{-1} \mathbf{M}_3$  would lead to the same covariance matrix for every nonzero scalar  $\lambda \in \mathbb{R} \setminus \{0\}$ . We need to impose constraints on the  $\mathbf{M}_i$ 's to ensure identifiability, once again, distinguishing between structured and unstructured factors. Structured factors usually have a scalar variance term that scales the correlations. Fixing this scalar variance term to 1, for example, will resolve the

aforementioned identifiability issue. To ensure identifiability in separable models, we cannot have separate variance terms associated with each factor. Instead, a single global variance term will absorb the variability due to all structured factors.

Matters are somewhat more lenient with unstructured factors and one can restrict the unstructured  $\mathbf{M}_j$ 's matrices in several different ways to ensure identifiability. We point out one such model, which is especially convenient from a computational standpoint. Botella-Rocamora et al. (2015) propose to model the elements on the unstructured  $\mathbf{M}_j$  matrices, say  $\mathbf{M}_2$  and  $\mathbf{M}_3$ , as Gaussian random effects. Identifiability is guaranteed by simply setting the variances of these random effects to be equal, whence the entries in  $\mathbf{M}_2$  and  $\mathbf{M}_3$  will have similar scale parameters. Now matrices such as  $\lambda\mathbf{M}_2$  and  $\lambda^{-1}\mathbf{M}_3$ , for any arbitrary nonzero scalar  $\lambda \neq 1$ , would be discarded because they would introduce different scales.

Imposing one of these restrictions on every factor will remove the identifiability issues during the inference. This comment also applies to the non-separable settings we describe below.

## 4 A non-separable multidimensional proposal

The separable model discussed in the previous section is a straightforward extension of the separable model in Martinez-Beneito (2013) subjected to the  $\mathcal{M}$ -based reparameterization proposed in Botella-Rocamora et al. (2015). However, it is not difficult to envision situations where separable disease mapping models are inappropriate. For example, seeking different between-diseases covariance matrices for males and females would lead to non-separable models. Spatio-temporal situations where every disease is likely to have its own temporal auto-regressive processes (with disease-specific parameters) is another example where separable models may be too restrictive. After all, why should all the disease risks share a common temporal dependence structure? Here, we show how tensor algebra can be exploited to construct non-separable multidimensional disease mapping models. We first extract some tools from the separable proposal and then use them to construct non-separable models. As earlier, we elucidate with the three-factor case.

### 4.1 Going beyond separability

Under separability, dependence is separately introduced for every factor by means of an  $n$ -dimensional product of a tensor with the corresponding structured matrix. For non-separable models, the dependence structure for one factor will change according to the different levels of the other(s). For example, if space and disease are two factors, then each disease (i.e. each "level") will have its own spatial covariance matrix. To fix matters, let us assume that we want to introduce non-separable covariance structures between the second and third factors in the trivariate setting.

In the separable case, dependence of these two factors was induced using (7). Following (4), the separable case yields

$$\text{vec}(\boldsymbol{\theta}) = (\mathbf{M}_3 \otimes \mathbf{I}_{L_2} \otimes \mathbf{I}_{L_1})(\mathbf{I}_{L_3} \otimes \mathbf{M}_2 \otimes \mathbf{I}_{L_3})\text{vec}(\mathcal{Y}).$$

We introduce non-separable structures by *nesting* the dependence structure for one factor within the levels of the other. For example, we can nest level 2 into level 3 by generalizing the preceding expression to

$$\text{vec}(\boldsymbol{\theta}) = (\mathbf{M}_3 \otimes \mathbf{I}_{L_2} \otimes \mathbf{I}_{L_1})(\text{Bdiag}(\{\mathbf{M}_2(i): i = 1, \dots, L_3\}) \otimes \mathbf{I}_{L_1})\text{vec}(\mathcal{Y}), \quad (8)$$

where  $\{\mathbf{M}_2(i): i = 1, \dots, L_3\}$  induce, for every level of the third factor, a different correlation structure for the second factor, in contrast to the separable case which considers the same correlation structure for all the levels of the third factor. Indeed, the separable case can be viewed as a particular case of (8), with  $\mathbf{M}_2(1) = \dots = \mathbf{M}_2(L_3)$ . Alternatively, factor 3 could be nested into factor 2 as

$$\text{vec}(\boldsymbol{\theta}) = (\text{Bdiag}(\{\mathbf{M}_3(i): i = 1, \dots, L_2\}) \otimes \mathbf{I}_{L_1})(\mathbf{I}_{L_3} \otimes \mathbf{M}_2 \otimes \mathbf{I}_{L_1})\text{vec}(\mathcal{Y}) \quad (9)$$

yielding a different covariance structure and, therefore, a different model.

The complexity in non-separable dependence structures depends upon the number of variables defining each of them and will vary as long as the numbers of levels in factors 2 and 3 are different. Unlike in separable models, where the order of the factors is irrelevant because the matrix products in (5) are commutative, for non-separable models the order matters because the matrix products are not commutative in (8) and (9). Therefore, we can further generate two new models from (8) and (9) based upon the order in which the factors appear. Consequently, the nesting of two variables will generate four different models.

Nesting factors, as above, is not the only way to generate non-separable models. Consider the expressions in (8) and (9) and the models obtained by reversing the order of the variables in the model. These can be viewed as alternative expressions for

$$\text{vec}(\boldsymbol{\theta}) = (\mathbf{M}_{(2,3)} \otimes \mathbf{I}_{L_1})\text{vec}(\mathcal{Y}),$$

for some  $\mathbf{M}_{(2,3)}$  with a specific structure. This expression could also be alternatively formulated as

$$\boldsymbol{\theta} = \mathbf{M}_{(2,3)} \circ_{(2,3)} \mathcal{Y}. \quad (10)$$



An obvious way to generalize these expressions is to consider  $\mathbf{M}_{(2,3)}$  as an unstructured  $(L_2L_3) \times (L_2L_3)$  matrix as opposed to the nested design. Therefore, non-separability can also be induced between factors as a kind of *factorial* design since expression (10) does not consider a factor to be put into the other, as in the nested design. Instead, it models any possible combination of the two factors at hand. In this case, the two corresponding factors are jointly modelled as a single factor with  $L_2L_3$  levels to model their interaction flexibly. Therefore, factorial interactions between two factors may be considered the most flexible way to induce dependence on them.

Hitherto, we have only considered interactions between two factors. Interactions between three or more factors is treated analogously. Factorial non-separability for higher orders is fairly straightforward to achieve by considering the  $\mathbf{a}$ -product of a matrix with a tensor, where  $\mathbf{a}$  is a vector of length greater than two. To introduce dependence among more than two factors, we can nest one factor within a combination of others. One example of this interaction is to allow the parameter(s) controlling the spatial structure to vary, for example, for every combination of disease and sex. However, for three or more factors the number of different interactions that could be defined is much higher than for only two factors. This is not dissimilar to the explosion in the number of models arising in ANOVA when considering high-order interactions. This problem is further exacerbated because under non-separability the order in which the dependence structures are included into the model also matter. Hence, we advise caution when introducing high-order interactions within multidimensional settings in order to avoid a large number of models.

To summarize, multidimensional disease mapping models can be treated as a series of operations on an array  $\mathcal{X}$

$$\mathbf{M}_{\alpha_1} \circ_{\alpha_1} \mathbf{M}_{\alpha_2} \circ_{\alpha_2} \dots \mathbf{M}_{\alpha_n} \circ_{\alpha_n} \mathcal{X}, \quad (11)$$

where  $\mathbf{a}_i (i = 1, \dots, n)$  are subsets of  $\{1, \dots, N\}$  with one or more elements and  $\mathbf{M}_{\mathbf{a}_i} (i = 1, \dots, n)$  are  $(\prod_{j \in \mathbf{a}_i} L_j) \times (\prod_{j \notin \mathbf{a}_i} L_j)$  matrices. If  $\mathbf{a}_i = (j)$  for any  $j \in \{1, \dots, N\}$  and  $j \notin \mathbf{a}_{i'}, \forall i' \neq i$ , then factor  $j$  will be separable with respect to the rest of factors in the model. In contrast, if one factor belongs to just one of the  $\mathbf{a}_i$ 's, whose length is greater than one, then it will have a non-separable covariance structure with regard to the rest of factors included in  $\mathbf{a}_i$  but a separable covariance structure with regard to the rest of factors in the model. Moreover, whenever non-separable dependence structures are present, i.e.  $\mathbf{a}_i$  is of length greater than 1, some of the operations in (11) will cease to be commutative. More precisely, if  $\mathbf{a}_i$  and  $\mathbf{a}_j$  satisfy  $\mathbf{a}_i \cap \mathbf{a}_j = \emptyset$  then both operations  $\circ_{\mathbf{a}_i}$  and  $\circ_{\mathbf{a}_j}$  will commute; otherwise they will not. Besides, since we have at least two different tools (nesting and factoring) for defining  $\mathbf{M}_{\mathbf{a}_i}$  where  $\mathbf{a}_i$  has length greater than one, the number of non-separable multidimensional models grows rapidly with growing numbers of factors.

## 4.2 The trivariate case

We now illustrate, in greater detail, the three-dimensional setting. First, we are going to introduce the following nomenclature to name the different models that can be built with the above-described tools. We use  $i$  to denote the model where a separable dependence structure is induced for the  $i$ th factor. We let  $i(j)$  represent when factor  $i$  is nested within factor  $j$ , i.e. the covariance matrix for factor  $i$  varies by the levels of factor  $j$ , while  $ij$  denotes a factorial covariance structure for factors  $i$  and  $j$  in the model. Thus,  $1\cdot 2\cdot 3$  denotes the trivariate model, while  $1(2)\cdot 23$  represents the model where the covariance matrix for the levels of the first factor varies across the levels of the second factor, and the second and third factors are modelled using a factorial covariance structure. Finally, for any two dependence structures (separable or non-separable),  $x\cdot y$  depicts the model where the dependence structure on  $x$  precedes that on  $y$ . Hence, in general, the  $x\cdot y$  model is different from  $y\cdot x$ .

Multidimensional modelling can be seen as a combination of mathematical operations on an unfolded Gaussian array. These elemental operations for the trivariate case are shown in Table 1. For instance, the fully separable model is a combination of operations corresponding to a separable structure for the first (1 $\cdot$ ), second (2 $\cdot$ ) and third (3 $\cdot$ ) factors in the model. Table 1 also shows, in the unstructured cases, the matrix (matrices) involved and the number of variables in it (them). The different number of variables in every model indicates the different levels of complexity. It is also instructive to note that if, for example,  $L_2 > L_3$ , then nesting factor two within factor three (i.e. 2(3) $\cdot$ ) yields a more complex model than nesting factor three into factor two. Consequently, these two nesting operations produce different models. Moreover, the number of parameters implied by any nesting design is lower than that in the corresponding factorial design. This, again, shows the added complexity in factorial models.

Each row of Table 1 corresponds to a different  $\mathbf{M}_{\mathbf{a}} \circ \mathbf{a}$  operation for different values of  $\mathbf{a}$  and  $\mathbf{M}_{\mathbf{a}}$ . The last two columns of Table 1 reveal these two values, for every elemental operation shown. These operations can involve one, two or three factors altogether and they can induce separable, nested, factorial or mixed (nested/factorial) covariance structures. These operations impart structure to the variance in different ways. They range from the simplest model, which assigns structure for all three factors (the fully-separable model), to the most complex, which, by separate, assigns a factorial dependence structure for all three factors (the 123 $\cdot$  model).

The fifth column in Table 1 depicts  $\mathbf{M}_{\mathbf{a}}$  corresponding to every elemental operation in that table. Note that each of these operations lead to mathematically different expressions, thereby yielding different covariance structures. Furthermore, for any of the rows in Table 1 and any  $k \notin \mathbf{a}$ , the corresponding  $\mathbf{M}_{\mathbf{a}}$  is either a Kronecker product or the sum of Kronecker products of matrices with their  $k$ th component being equal to  $\mathbf{I}_{L_k}$ . The product of a square matrix with a suitable identity matrix is commutative. This implies that  $\mathbf{M}_{\mathbf{a}} \circ \mathbf{a}$  and  $\mathbf{M}_{\mathbf{a}'} \circ \mathbf{a}'$  commute for any two mutually disjoint index sets  $\mathbf{a}$  and  $\mathbf{a}'$ .

We also remark that certain combinations of elemental operations, while mathematically legitimate, may lack statistical interpretability. For example, the combination of the 12 $\cdot$  and 23 $\cdot$  operations is difficult to interpret because they assume that factors one and two on one

side and two and three on the other side are combined with as much flexibility as possible for every one of these pairs. In that case, it would seem much more natural to consider instead a 12(3)·, a 23(1)· or a 123· relationship.

Table 1 considers that factors one, two and three in multidimensional modelling are unstructured. However, in our context there will always be a structured factor among these—the spatial component. We will henceforth assume that the spatial units correspond to the levels of the first factor. As is customarily assumed in spatial models, we specify certain spatial process or distribution such as Intrinsic CAR or proper CAR to induce spatial dependence on the geographical units. Now the operations 12·, 13·, 123·, 12(3)· and 13(2)· in Table 1 will no longer be sensible because they produce unstructured relationships for both the spatial units and the other factor, which contradicts the assumed spatial structure for factor one. Therefore, for structured factors in general (and not just for spatial), nesting appears to be the only practical way of building non-separable relationships with other terms in the model as it preserves the original dependence structure of factors.

However, models incorporating any factor(s) nested within the spatial factor do not seem reasonable either. These models would allow some covariance matrix (for any of the factor(s) in the model) to vary by every spatial unit. This would surely yield overparameterized models since the number of geographical units is typically much higher than the number of levels in the rest of the factors in the model. Hence, although the combination of operations in Table 1 could generate a large number of models, all these considerations will limit that quantity to some extent and, as we will see in the next example, that quantity will be (at least for the trivariate case) very reasonable in practice.

We conclude this section with some remarks on the practical implementation of the proposed models. Although  $\mathbf{M}_\alpha$  in Table 1 can appear to be very intricate, they are usually much easier to implement in practice. For example, to induce a separable covariance structure we simply consider the product between one of the dimensions of a Gaussian array and a matrix endowed with the appropriate structure. Second, if we want to nest that structure within another factor, for example  $j$ , the only change that we need to do is to allow that matrix to change for every level of the  $j$ th factor. Finally, if we wanted to introduce a factorial interaction between two factors, we would just have to consider every combination of both of them as a single factor and would have to introduce a separable covariance structure, as described above, for that combination of factors. These mathematical operations are conceptually straightforward and computationally inexpensive. In fact, for the illustrations in the next section, all the models (three and four-dimensional) were easily implemented using the Bayesian software WinBUGS (Lunn et al., 2000).

## 5 Two multidimensional studies on Comunitat Valenciana's mortality data

We have carried out two separate multidimensional studies with Comunitat Valenciana's mortality data. The dataset corresponds to the Spatio-temporal Mortality Atlas of Comunitat Valenciana (Zurriaga et al., 2010), which contain the deaths occurred in that region (540 municipalities) during the period 1987–2006. The first of the examples shows the effect of different kinds of non-separable trivariate structures, illustrating the models introduced in

Section 3.2. We demonstrate the inappropriateness of the separable hypothesis and illustrate how non-separable models, constructed using our approach, offer vastly improved fits. Finally, we analyse a second dataset with four factors: two unstructured factors (Sex, Disease) and two structured factors (Period and geographical unit).

All the models we describe below have been implemented in WinBUGS 1.4.3 (Lunn et al., 2000) and the programs are available at <http://www.uv.es/mamtnez/MultiDim.html>. For each model, we have considered proper CAR distributions to model the spatial dependence among geographical units. We ran three chains for each model with a total of 15,000 iterations per chain. The first 5,000 of these, were discarded as burn-in and only one of every 30 iterations was retained for subsequent posterior analysis. Thus, a total of 1,002 iterations (334 per chain) were finally saved. Improper flat prior distributions were used for the  $\mu$  parameters (intercepts) for all the models implemented. Uniform(0,100) prior distributions were used for all the standard deviations of the random effects in the model. The upper value of these uniform distributions was intended to yield vague prior distributions since the random effects in the model are used in a logarithmic scale.

The different chains used for every model were run in parallel in order to speed up computations. That is, instead of sending all three chains in a single call to WinBUGS, we made three different calls (one per chain) by means of an R (R Development Core Team, 2009) function developed for this purpose. Instead of running all three chains in a single core (as is the default in WinBUGS), each chain was run in a different core of the processor(s). This accrues considerable computational savings. Convergence was assessed by means of visual inspection of the history of the Deviance and a sample of variables in the models (models will typically contain thousands of variables) during the MCMC process.

An additional simulation study has been carried out in order to assess the performance of DIC for model selection in our context and the ability of some of the entertained models to retrieve the original variance-covariance matrix between geographical patterns. For lack of space the results of this study are included as supplementary material (Martinez-Beneito et al., 2016) to the paper which can be download also from <http://www.uv.es/~mamtnez/MultiDim.html>.

### 5.1 A non-separable trivariate mortality study

We next consider two trivariate scenarios with factors: Geographical Unit (540 levels), Disease (2 levels) and Sex (2 levels). We will refer to them as factors 1 to 3, respectively. We embark upon two separate studies. First, we consider the joint study of Colon and Rectum Cancer for both sexes and, second, the study of Lung Cancer and Diabetes also for both sexes. For these two studies we have ran all those models arising from the combination of the elemental operations in Table 1. Nevertheless, some of those combinations were not implemented because they produce trivial or uninterpretable models. Thus, for all models leading to identical fits because they just permute commutative operations (such as the  $1 \cdot 2 \cdot 3$ ,  $1 \cdot 3 \cdot 2$ ,  $3 \cdot 1 \cdot 2$ , . . . models), we ran just one of these equivalent choices. Moreover, we did not consider the nesting of the factors disease or sex into the geographical component since, as alluded to earlier, this does not make much sense, neither did we consider any factorial design involving a geographical factor because the corresponding model will miss

the spatial structure inherent to the geographical component. Finally, for computational convenience and to exploit (7), we restricted ourselves to models with the spatial factor entering first.

Table 2 shows the DIC, Deviance Information Criterion (Spiegelhalter et al., 2002), and the effective number of parameters (pD) for the implemented models and for both datasets. Model 1 in Table 2 corresponds to a fully separable dependence structure, Models 2–8 modify Model 1 by nesting one factor inside other(s) and Model 9 corresponds to a factorial relationship for Disease and Sex. Other models are also possible combining two or more of the elemental operations implemented in models of rows 2–9 in Table 2. Model 10 is the only such model we implemented as it was expected to yield some improvement, for the lung cancer/disease study, as mentioned below.

For the Colon/Rectum study, the model with the lowest DIC is Model 3. This model accommodates spatial dependence parameters for the CAR models to vary across sexes. We point out that none of the models accounting for non-separability between Disease and Sex (Models 5–9) show notable improvements with respect to the fully separable model. Model 10 was not run for this study because it too considered non-separability between Disease and Sex and was not expected to yield any improvement.

In contrast, non-separability between Disease and Sex seems to improve the fit for the Lung/Diabetes study. One such model with the factorial structure delivers the lowest DIC. Nesting of the geographical component within the other factors may also yield some improvement in some occasions, mainly the nesting of the geographical component within diseases. Therefore, we have run the model incorporating a factorial interaction between Disease and Sex and nesting the geographical structure within diseases. However, this model does not perform better than that incorporating only the factorial relationship between Disease and Sex. The fifth column of Table 2 shows, for illustrative purposes, the computing time needed to run every model for the Lung/Diabetes study (for the Colon/Rectum study those times were basically the same). These times are very reasonable although the factorial relationship substantially increases computational time.

Besides model selection with DIC, we have also assessed the fit of the models implemented for both datasets. For this goal, we have used Posterior Predictive  $p$ -values (PP  $p$ -values) as introduced by Gelman et al. (1996). There, for models with Poisson data likelihoods, they introduced a chi-square statistic based on the assumed normality of  $(O_{ijk} - (E_{ijk}RR_{ijk}))^2 / (E_{ijk}RR_{ijk})$ . Since in our case the observed cases per unit are in general low, often zero, the normal approximation of the Poisson distribution would not work necessarily well, so we decided to make our assessment based on the deviance of the model. Specifically, let  $D(O_{ij}, RR_{ij})$  denote the deviance ( $-2$  times the log-likelihood) for the relative risks  $RR_{ij}$  and the available data  $O_{ij}$  where  $i$  and  $j$  denote, respectively, the corresponding disease and sex. Similarly, let  $O_{ij}^{rep}$  denote a replicate of the observed data sampled from the posterior predictive distribution  $P(O_{ij}^{rep} | RR_{ij}, O_{ij})$ . Then we used  $q_{ij} = P(D(O_{ij}, RR_{ij}) > D(O_{ij}^{rep}, RR_{ij}))$  as a deviance-based PP  $p$ -value to assess the fit of the models run.

Few differences were found among models in terms of the mentioned  $q_{ij}$ s. For example, for the Colon/Rectum dataset, the separable model  $1 \cdot 2 \cdot 3$  yielded the following PP  $p$ -values: (0.25, 0.19, 0.02, 0.78) meanwhile the PP  $p$ -values for the  $1 \cdot 23$  model, the model with the most complex interaction, were (0.23, 0.16, 0.02, 0.68). The PP  $p$ -values for the rest of models ranged on similar quantities. For the Lung/Diabetes data set, the PP  $p$ -values for the separable  $1 \cdot 2 \cdot 3$  model were (0.36, 0.18, 0.12, 0.71). Thus, no extreme PP  $p$ -value was observed for any of the combinations of disease and sex considered. Although the PP  $p$ -value for Colon Cancer in Women is a bit low, maybe pointing out a slight lack of fit, this value is not very worrisome, mostly bearing in mind that it corresponds to the most extreme value out of a set of 8 PP  $p$ -values.

Table 3 shows the estimated correlations (posterior means and 80% Credible Intervals) between the different maps for the  $1 \cdot 23$  model in both studies. Results in that table correspond to the model considering a factorial non-separable relationship between Disease and Sex. The upper and lower rows of every cell correspond respectively to the Colon/Rectum and Lung/Diabetes studies. For the Colon/Rectum study, the fully-separable model yielded a posterior mean of the correlation between diseases of 0.79 (80% Credible Interval, [0.62, 0.93]) and a correlation between sexes of 0.92 (80% Credible Interval, [0.83, 0.99]). Therefore, the improvement of jointly considering both sexes in this study is higher than that of considering both diseases altogether.

Table 3 reveals that for the factorial relationship in the Colon/Rectum study, correlations between maps are mainly driven by the product of those correlations in the fully separable model. This explains why models imposing non-separable relationships on these two factors in Table 2 were performing worse. On the other hand, for the Lung/Diabetes study, correlation between diseases was estimated as 0.35 (80% Credible Interval, [0.23, 0.46]) and that between sexes was 0.62 (80% Credible Interval, [0.53, 0.71]) for the fully-separable model. In this case, the combination of these values does not reproduce the correlation matrix shown for these diseases in Table 3, which even yields a somewhat counter-intuitive negative correlation for Lung cancer and Diabetes in women. For these two diseases, a separable relationship between Disease and Sex is clearly an excessively simplistic assumption.

Figure 1 shows all four maps for the Colon/Rectum (upper row of the plot) and Lung/Diabetes (lower row of the plot) studies. Results for the Colon/Rectum study correspond to the  $1(3) \cdot 2 \cdot 3$  model, while results for the Lung/Diabetes study correspond to the  $1 \cdot 23$  model; these produce the best fit in their respective cases. For the Colon/Rectum study, all four maps appear to be very similar. The more appreciable discrepancies in them correspond to both different diseases and sexes. Relationships among maps are somewhat more intricate for the Lung/Diabetes study. Here, both maps for men share common features but the two maps for women are very different, perhaps the most different among all pair-wise comparison of maps. This suggests, again, the non-separability of Disease and Sex in this study.

## 5.2 A four-dimensional mortality study

We now present a four-dimensional version of the Lung/Diabetes study from the previous Section. We consider the same dataset, dividing the whole period of study (1987–2006) into five different four-year periods. Hence, we have a new factor, the Time period, to include in the multidimensional study. This factor, unlike Disease and Sex, has a specific structure reflecting temporal dependence that should, ideally, be accounted for. We assume a first-order autoregressive structure to model this factor and specify the resulting dependence structure using the matrix in (6).

As in our earlier experiments, we have again fitted several models and compared their performances using the DIC. Results are shown in Table 4. When a model engenders an alternative by permuting the order in which dependence is introduced to the factors, the alternative model is shown in the right-hand-side of the table. Since Time period has a specific (temporal) structure, no factorial interaction with any other factor of the model has been considered. These models incorporate only non-separable modifications of the fully-separable model involving two factors. If more than one model performed better than the fully-separable model, such as the  $1(4) \cdot 2 \cdot 3 \cdot 4\cdot$  and  $1 \cdot 23 \cdot 4\cdot$  models, we combined them into a single one in a second step of the analysis, e.g. into a  $1(4) \cdot 23 \cdot 4$  model.

Results in Table 4 are shown in the following way. Row 1 shows the DIC for the fully-separable model, rows 2–3 show the results for models imposing non-separability for the new factor in the study (Time period), and rows 4–11 show the results for those models assuming separability for Time period. As can be appreciated, models in rows 2–3 perform worse than the fully-separable model. This suggests that a non-separable dependence structure for Time period is not appropriate. Put differently, temporal evolutions for every combination of Disease and Sex can be considered as first-order autoregressive processes of a common parameter. When time is considered a separable factor, results are quite similar to the trivariate case. A non-separable relationship between Disease and Sex enjoys credence and the spatial parameters do not seem to vary for any of the factors considered in the model. Moreover, no non-separable relationship between Time period and either Disease or Sex seems appropriate. The only two modifications substantially improving the fully-separable models (models  $1 \cdot 3(2) \cdot 2 \cdot 4\cdot$  and  $1 \cdot 23 \cdot 4\cdot$ ) propose a non-separable relationship between Disease and Sex, with the factorial model being more general than the nested. It makes little sense to combine these two models because the factorial model is the most general proposal incorporating non-separability between these factors.

Regarding computing times for the models run in this study, the fully-separable model took 780 minutes to run. This time is about 40 times higher than the corresponding trivariate model. We have also run the four-dimensional model without considering any particular temporal structure for Time period and the computing time decreased to 351 minutes. Therefore, the temporal structure seems to considerably slow down the MCMC sampling. For the remaining models, the increase from the three to the four-dimensional case is similar. The best-performing model, the factorial  $1 \cdot 23 \cdot 4\cdot$  model took 2,223 minutes to run. All models revealed excellent convergence and could surely have been run with less iterations than those simulated in our study.

Models  $1 \cdot 3(2) \cdot 2 \cdot 4$  and  $1 \cdot 23 \cdot 4$  have been selected as the most appropriate models based on Table 4. These two models, obviously without modelling Time period, also produced excellent results in the trivariate case. Hence, the results from both analyses clearly agree. Regarding the  $1 \cdot 23 \cdot 4$  model, which has the best DIC score in the trivariate case, the parameter controlling the temporal correlation of the auto-regressive process has a posterior mean of 0.92 (80% Credible Interval, [0.90, 0.94]), which is strongly indicative of high temporal dependence for all four combinations of Disease and Sex. The correlation matrix between every combination of Disease and Sex for the four-dimensional study (results not shown) is very similar to that shown in Table 3 for the Lung/Diabetes study. Again, the results clearly agree for both studies.

Finally, we have also included, as supplementary material to the paper, the maps of the first, third and fifth period of study for every combination of Disease and Sex. These maps clearly show temporal dependence, although they also show temporal variability for each of these combinations. Such a temporal coherence of maps between periods is very rewarding because all of them are based upon a very limited amount of information that is compensated for by the sharing of information between maps.

## 6 Conclusions

This paper has tried to set forth some theoretical bases for the development of multivariate disease mapping analyses involving more than one factor besides the geographical factor, what we have called multidimensional disease mapping studies. Very clear links can be drawn between the multidimensional disease mapping problem and tensor algebra-calculus therefore the latter offers a clear contextual framework where multidimensional methods can be developed, formalized and studied. In our opinion the establishment of new links between these two areas of research may yield new tools and very valuable ideas for the development of multidimensional models.

Most of the models compared in the examples produce quite similar risk estimates with hardly any practical difference, at least in terms of their posterior means. Maybe, as pointed out by a reviewer, performing quite an extensive model selection as that performed in our examples does not make much sense, nevertheless, we considered it convenient to implement and compare such a large number of models in order to illustrate the variety of models introduced along the paper. In practical terms, we advise users to fit fewer models than those considered in our examples. For example, from an epidemiological point of view, we do not see any relevant difference between the  $1 \cdot 2 \cdot 3 \cdot 4(2)$  and the  $1 \cdot 3 \cdot 4(2) \cdot 2$  models in Example 5.2. Since both models produce similar estimates we would advise users to fit just one of them, i.e. we advise to compare just models which have relevant epidemiological differences in their interpretations. This will keep simpler the analysis made and neither their interpretations nor their conclusions should be very different.

Some models have already been formulated which may be competitive alternatives to the framework proposed in this paper. Thus SANOVA (Zhang et al., 2009; Marí Dell'Olmo et al., 2014) is a method for multivariate modelling which allows to structure in some specific ways the covariance structure between geographical patterns. This particular feature of



SANOVA makes it also suitable for modelling complex dependence relationships like in multidimensional settings, even for structured factors (Torres-Avilés and Martínez-Beneito, 2015). Nevertheless, SANOVA has several drawbacks in comparison to multidimensional modelling. First, in contrast to SANOVA, multidimensional modelling does not rely on any specific choice of contrasts to make specific comparisons between geographical patterns. This makes SANOVA results to be contrast-dependent when usually the choice of contrasts may be rather arbitrary. Second, to achieve such a flexibility as in multidimensional modelling, SANOVA models would have to include several interaction terms between the contrasts used. This would make their results even more contrasts-dependent and to be as parameterized as multidimensional models. Nevertheless, for modelling in multidimensional settings we acknowledge that SANOVA may be an interesting alternative to the approach introduced in this paper.

The models developed within this framework, despite their high complexity due to the difficulty of incorporating several factors within a unique dependence structure, are reasonably affordable from an applied point of view. All of them can be run within `winBUGS` what makes this methodology available for a very large community of users. Moreover, computing times are also reasonable what makes this methodology available in practice for the joint study of several factors altogether. Multidimensional modelling makes also possible to decompose the data in smaller geographical or temporal pieces since other diseases, sexes, races, . . . will provide complementary information making it possible to yield reliable estimates in such a small units. This paper introduces some guidelines that will make possible some new studies on that direction and allowing to work with smaller units than those currently used. Finally, we find convenient to mention that the methodology introduced in this paper can be used beyond areal data through Gaussian Markov random fields. This framework could be also used for the analysis of spatial data in a continuous domain through Gaussian random fields, for the modelling of time or for simply structuring the covariance of multivariate Gaussian random effects in general. The possibilities of this framework in these settings have not yet been explored.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

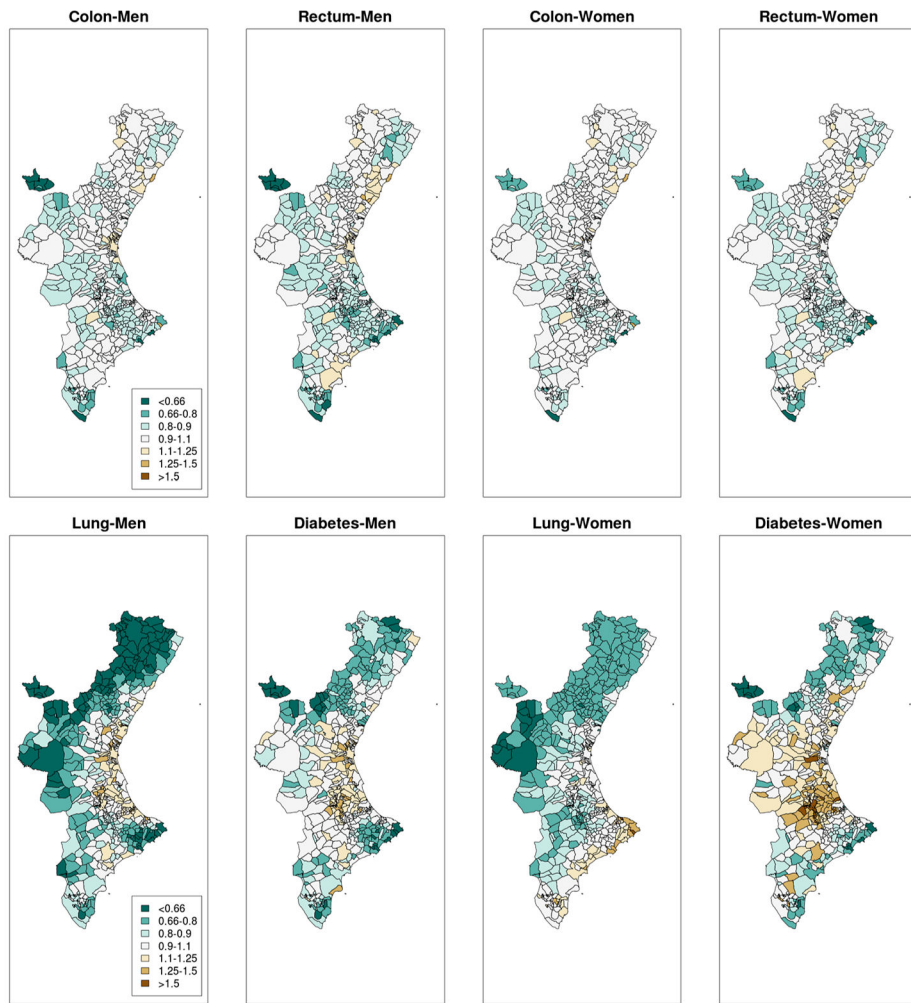
The authors thank the associate editor and two reviewers for their constructive comments. We also thank Professor Ying Macnab for her helpful comments and interesting conversations on this work.

Dr. Botella-Rocamora acknowledges the financial support of Ministerio de Educación, Cultura y Deporte, via the Programa Nacional de Movilidad de Recursos Humanos del Plan Nacional de I-D+i 2008–2011, prorogued by agreement of the Consejo de Ministros of 2011/10/7. Dr. Martínez-Beneito acknowledges the financial support of the research grants MTM2013-42323-P from the Spanish Ministry of Economy and Competitiveness and ACOMP/2015/202 from the Generalitat Valenciana. Dr. Banerjee's work was supported in part by grants NIH/NIGMS 1-RC1-GM092400-01, NIH/NCI 1R03CA179555-01A1 and NSF/DMS-1513654.

## References

Banerjee, S., Roy, A. Linear Algebra and Matrix Analysis for Statistics. Chapman & Hall/CRC; 2014.

- Botella-Rocamora P, Martinez-Beneito MA, Banerjee S. A unifying modeling framework for highly multivariate disease mapping. *Statistics in Medicine*. 2015; 34(9):1548–1559. [PubMed: 25645551]
- Dobra A, Lenkoski A, Rodriguez A. Bayesian inference for general Gaussian graphical models with application to multivariate lattice data. *Journal of the American Statistical Association*. 2011; 106(496):1418–1433. [PubMed: 26924867]
- Gelman A, Meng XL, Stern H. Posterior predictive assessment of model fitness via realized discrepancies. *Statistica Sinica*. 1996; 6:733–807.
- Jin X, Banerjee S, Carlin BP. Order-free co-regionalized areal data models with application to multiple-disease mapping. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. 2007; 69(5):817–838. doi: <http://dx.doi.org/10.1111/j.1467-9868.2007.00612.x>. [PubMed: 20981244]
- Lunn D, Thomas A, Best N, Spiegelhalter D. WinBUGS – a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and Computing*. 2000; 10:325–337.
- Macnab YC. On Gaussian Markov random fields and Bayesian disease mapping. *Statistical Methods in Medical Research*. 2011; 20:49–68. doi: <http://dx.doi.org/10.1177/0962280210371561>. [PubMed: 20547586]
- Marí Dell’Olmo M, Martinez-Beneito MA, Gotséns M, Palència L. A Smoothed ANOVA model for multivariate ecological regression. *Stochastic Environmental Research and Risk Assessment*. 2014; 28(3):695–706.
- Martinez-Beneito MA. A general modelling framework for multivariate disease mapping. *Biometrika*. 2013; 100(3):539–553. doi: <http://dx.doi.org/10.1093/biomet/ast023>.
- Martinez-Beneito, MA., Botella-Rocamora, P., Banerjee, S. Supplementary material of “Towards a multidimensional approach to Bayesian disease mapping”. *Bayesian Analysis*. 2016. doi: <http://dx.doi.org/10.1214/16-BA995SUPP>
- R Development Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing; Vienna, Austria: 2009. URL <http://www.R-project.com>
- Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A. Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. 2002; 64:583–641. doi: <http://dx.doi.org/10.1111/1467-9868.00353>.
- Torres-Avilés F, Martinez-Beneito MA. STANOVA: a smooth-ANOVA-based model for spatio-temporal disease mapping. *Stochastic Environmental Research and Risk Assessment*. 2015; 29(1): 131–141. Advance online publication. DOI: 10.1007/s00477-014-0888-1
- Tzala E, Best N. Bayesian latent variable modelling of multivariate spatio-temporal variation in cancer mortality. *Statistical Methods in Medical Research*. 2008; 17(1):97–118. doi: <http://dx.doi.org/10.1177/0962280207081243>. [PubMed: 17855747]
- Zhang S, Sun D, He CZ, Schootman M. A Bayesian semi-parametric model for colorectal cancer incidences. *Statistics in Medicine*. 2006; 25:285–309. doi: <http://dx.doi.org/10.1002/sim.2221>. [PubMed: 16381075]
- Zhang Y, Hodges JS, Banerjee S. Smoothed ANOVA with spatial effects as a competitor to MCAR in multivariate spatial smoothing. *Annals of Applied Statistics*. 2009; 3(4):1805–1830. doi: <http://dx.doi.org/10.1214/09-AOAS267>. [PubMed: 20596299]
- Zurriaga, O., Martínez-Beneito, MA., Botella-Rocamora, P., López-Quílez, A., Mel-chor, I., Amador, A., Vanaclocha, H., Nolasco, A. [Accessed October 13th, 2014] Spatio-temporal Mortality Atlas of Comunitat Valenciana. 2010. URL <http://www.geeitema.org/AtlasET/index.jsp?idioma=I>



**Figure 1.** Posterior mean of the Relative Risk for every municipality. Results in the first row correspond to the Colon/Rectum study and those in the second row correspond to the Lung/ Diabetes study.

**Table 1**

Elemental operations for building up three-dimensional models.

Operation	Matrices involved	# parameters	$\alpha$	$M_a$
<i>Separable covariance structures</i>				
1.	$M_1$	$L_1 L_1$	(1)	$(\mathbf{I}_{L_3} \otimes \mathbf{I}_{L_2} \otimes M_1)$
2.	$M_2$	$L_2 L_2$	(2)	$(\mathbf{I}_{L_3} \otimes M_2 \otimes \mathbf{I}_{L_1})$
3.	$M_3$	$L_3 L_3$	(3)	$(M_3 \otimes \mathbf{I}_{L_2} \otimes \mathbf{I}_{L_1})$
<i>Non-separable nested structures</i>				
1(2).	$M_1^{L_2}, \dots, M_1^{L_2}$	$L_1 L_1 L_2$	(1,2)	$\sum_{j=1}^{L_2} (\mathbf{I}_{L_3} \otimes \mathbf{e}_j^{L_2}) \otimes M_1^j$
1(3).	$M_1^{L_3}, \dots, M_1^{L_3}$	$L_1 L_1 L_3$	(1,3)	$\sum_{k=1}^{L_3} (\mathbf{e}_k^{L_3}) \otimes \mathbf{I}_{L_2} \otimes M_1^k$
1(23).	$M_1^{(1,1)}, \dots, M_1^{(L_2, L_3)}$	$L_1 L_1 L_2 L_3$	(1,2,3)	$\sum_{j=1}^{L_2} \sum_{k=1}^{L_3} (\mathbf{e}_k^{L_3}) \otimes \mathbf{e}_j^{L_2} \otimes M_1^{(j,k)}$
2(1).	$M_2^{L_1}, \dots, M_2^{L_1}$	$L_2 L_2 L_1$	(1,2)	$\sum_{i=1}^{L_1} (\mathbf{I}_{L_3} \otimes M_2^i \otimes \mathbf{e}_i^{L_1})$
2(3).	$M_2^{L_3}, \dots, M_2^{L_3}$	$L_2 L_2 L_3$	(2,3)	$\sum_{k=1}^{L_3} (\mathbf{e}_k^{L_3}) \otimes M_2^k \otimes \mathbf{I}_{L_1}$
2(13).	$M_2^{(1)}, \dots, M_2^{(L_3)}$	$L_2 L_2 L_1 L_3$	(1,2,3)	$\sum_{i=1}^{L_1} \sum_{k=1}^{L_3} (\mathbf{e}_i^{L_1}) \otimes M_2^{(i,k)} \otimes \mathbf{e}_k^{L_3}$
3(1).	$M_3^{L_1}, \dots, M_3^{L_1}$	$L_3 L_3 L_1$	(1,2,3)	$\sum_{i=1}^{L_1} (M_3^i \otimes \mathbf{I}_{L_2} \otimes \mathbf{e}_i^{L_1})$
3(2).	$M_3^{L_2}, \dots, M_3^{L_2}$	$L_3 L_3 L_2$	(2,3)	$\sum_{j=1}^{L_2} (M_3^j \otimes \mathbf{e}_j^{L_2}) \otimes \mathbf{I}_{L_1}$

Operation	Matrices involved	# parameters	$\alpha$	$M_a$
3(12)	$M_3^{(1)}, \dots, M_3^{(L_2)}$	$L_3 L_2 L_1 L_2$	(1,2,3)	$\sum_{i=1}^{L_1} \sum_{j=1}^{L_2} \sum_{k=1}^{L_3} \mathbf{M}_3^{(i,j)} \otimes \mathbf{e}_j^{L_2} \otimes \mathbf{e}_i^{L_1}$
<i>Non-separable factorial structures</i>				
12	$M_{12}$	$(L_1 L_2)^2$	(1,2)	$(\mathbf{I}_{L_3} \otimes \mathbf{M}_{12})$
23	$M_{23}$	$(L_2 L_3)^2$	(2,3)	$(\mathbf{M}_{23} \otimes \mathbf{I}_{L_1})$
13	$M_{13}$	$(L_1 L_3)^2$	(1,3)	$\sum_{i,i'=1}^{L_1} \sum_{k,k'=1}^{L_3} \mathbf{e}_k^{L_3} \otimes \mathbf{e}_i^{L_1} \otimes \mathbf{e}_{i'}^{L_1} \otimes \mathbf{I}_{L_2} (\mathbf{M}_{13})^{((k-1)L_1 + i, (k'-1)L_1 + i')}$
123	$M_{123}$	$(L_1 L_2 L_3)^2$	(1,2,3)	$M_{123}$
<i>Mixed nested/factorial structures</i>				
12(3)	$M_{12}^{L_3}, \dots, M_{12}^{L_3}$	$(L_1 L_2)^2 L_3$	(1,2,3)	$\sum_{k=1}^{L_3} \mathbf{e}_k^{L_3} \otimes \mathbf{M}_{12}^k$
23(1)	$M_{23}^{L_1}, \dots, M_{23}^{L_1}$	$(L_2 L_3)^2 L_1$	(1,2,3)	$\sum_{i=1}^{L_1} \mathbf{M}_{23}^i \otimes \mathbf{e}_i^{L_1}$
13(2)	$M_{13}^{L_2}, \dots, M_{13}^{L_2}$	$(L_1 L_3)^2 L_2$	(1,2,3)	$\sum_{i,i'=1}^{L_1} \sum_{j=1}^{L_2} \sum_{k,k'=1}^{L_3} \mathbf{e}_k^{L_3} \otimes \mathbf{e}_j^{L_2} \otimes \mathbf{e}_{i'}^{L_1} \otimes \mathbf{e}_j^{L_2} \otimes \mathbf{e}_i^{L_1} \otimes \mathbf{M}_{13}^{(j)}$

**Table 2**

DIC (and pD, within brackets) for the Colon cancer/Rectum cancer and the Lung cancer/Diabetes studies. The fifth column shows the computing times (in minutes) for every model implemented for the Lung cancer/Diabetes study.

Model	Dependence structure	DIC (pD) (Colon/Rectum)	DIC (pD) (Lung/Diabetes)	Computing time
1	1 · 2 · 3·	7546.9 (171.5) 9674.9 (489.6)	18.5	
2	1(2) · 2 · 3·	7553.4 (165.5)	9669.6 (499.0)	20.8
3	1(3) · 2 · 3·	<b>7539.1</b> (177.2)	9677.4 (493.2)	21.5
4	1(23) · 2 · 3·	7545.3 (173.6)	9670.4 (506.2)	18.0
5	1 · 2(3) · 3·	7547.2 (170.1)	9672.5 (474.6)	20.4
6	1 · 3 · 2(3)·	7556.0 (173.4)	9689.6 (465.9)	16.5
7	1 · 2 · 3(2)·	7551.0 (170.6)	9671.0 (473.0)	17.4
8	1 · 3(2) · 2·	7549.7 (175.0)	9670.9 (475.7)	14.5
9	1 · 23·	7552.0 (193.0)	<b>9666.8</b> (458.5)	67.8
10	1(2) · 23·	–	9670.5 (476.3)	66.9

**Table 3**

Estimated correlations matrix (posterior means and 80% Credible Intervals) between the different maps for the 1 · 23· model in both studies. Upper/lower row of every cell corresponds respectively to the Colon/Rectum and Lung/Diabetes studies.

	<b>Disease 1 Men</b>	<b>Disease 2 Men</b>	<b>Disease 1 Women</b>	<b>Disease 2 Women</b>
Disease 1 Men	1 1	0.73 [0.53, 0.89] 0.62 [0.49, 0.74]	0.84 [0.70, 0.96] 0.40 [0.22, 0.58]	0.63 [0.39, 0.85] 0.40 [0.29, 0.51]
Disease 2 Men		1 1	0.79 [0.60, 0.95] 0.00 [-0.22, 0.24]	0.77 [0.57, 0.94] 0.77 [0.69, 0.86]
Disease 1 Women			1 1	0.63 [0.33, 0.90] -0.24 [-0.45, -0.02]
Disease 2 Women				1 1

**Table 4**

DIC (and pD) for the models run. Rows two and three consider Time period as a non-separable factor while rows 4 to 11 consider it as a separable factor. Models on the right-hand-side of the table correspond to the left-hand-side models changing the order in which dependence is induced into the factors of the models.

Model	DIC (pD)	Model	DIC (pD)
1 · 2 · 3 · 4 ·	29008.7 (737.5)	–	–
1 · 2 · 3 · 4(2) ·	29014.9 (718.7)	1 · 3 · 4(2) · 2 ·	29021.3 (748.9)
1 · 2 · 3 · 4(3) ·	29015.7 (719.4)	1 · 2 · 4(3) · 3 ·	29015.4 (743.1)
1(2) · 2 · 3 · 4 ·	29013.5 (747.4)	–	–
1(3) · 2 · 3 · 4 ·	29008.3 (754.6)	–	–
1(4) · 2 · 3 · 4 ·	29020.3 (750.2)	–	–
1 · 2(3) · 3 · 4 ·	29003.6 (732.8)	1 · 3 · 2(3) · 4 ·	29007.9 (715.1)
1 · 2 · 3(2) · 4 ·	29019.9 (702.9)	1 · 3(2) · 2 · 4 ·	29002.5 (735.5)
1 · 23 · 4 ·	29002.5 (704.4)	–	–
1 · 2(4) · 3 · 4 ·	29018.6 (708.4)	1 · 4 · 2(4) · 3 ·	29018.8 (719.7)
1 · 2 · 3(4) · 4 ·	29028.9 (730.7)	1 · 2 · 4 · 3(4) ·	29019.3 (723.9)