SORT 43 (1) January-June 2019, 51-74 DOI: 10.2436/20.8080.02.79

Bayesian joint spatio-temporal analysis of multiple diseases

Virgilio Gómez-Rubio[1], Francisco Palmí-Perales[1], Gonzalo López-Abente[2], Rebeca Ramis-Prieto[2] and Pablo Fernández-Navarro[2]

Abstract

In this paper we propose a Bayesian hierarchical spatio-temporal model for the joint analysis of multiple diseases which includes specific and shared spatial and temporal effects. Dependence on shared terms is controlled by disease-specific weights so that their posterior distribution can be used to identify diseases with similar spatial and temporal patterns.
The model proposed here has been used to study three different causes of death (oral cavity, esophagus and stomach cancer) in Spain at the province level. Shared and specific spatial and temporal effects have been estimated and mapped in order to study similarities and differences among these causes. Furthermore, estimates using Markov chain Monte Carlo and the integrated nested Laplace approximation are compared.

MSC: 62F15, 62H11, 62M10.

Keywords: Bayesian modelling, Joint modelling, Multivariate disease mapping, Shared components. Spatio-temporal epidemiology.

1. Introduction

Bayesian hierarchical models are a popular approach to analyse public health spatio-temporal data. These data often come as counts of cases of disease at different administrative levels and time periods. Hierarchical models for these data are based on a Poisson regression model that includes different types of spatial, temporal and spatio-temporal effects in the linear predictor of the model (see, for example, Lawson, 2013, for a review). Spatial effects are often modelled using a conditionally autoregressive (CAR) specification (Besag, York and Mollié, 1991). Temporal effects often rely on smooth terms, such as random walks or splines. For non-separable space-time models, Knorr-

[1] Department of Mathematics, Universidad de Castilla-La Mancha, Albacete, Spain. Corresponding author: Virgilio Gómez-Rubio (Virgilio.Gomez@uclm.es).

[2] Environmental and Cancer Epidemiology Unit, Carlos III Institute of Health, Madrid, Spain; Consortium for Biomedical Research in Epidemiology & Public Health, CIBER Epidemiología y Salud Pública - CIBERESP, Spain.

Received: August 2018
Accepted: December 2018

Held (2000) describes different interactions for spatial and temporal effects. When studying spatio-temporal trends in disease mapping, several authors have proposed different models to detect specific patterns in particular areas. For example, Abellan, Richardson and Best (2008) and Guangquan et al. (2012) propose models that can identify areas that follow a spatio-temporal trend with similar structure or that show a specific spatio-temporal pattern.

The spatial analysis of several diseases often relies on multivariate models with shared spatial effects to capture similar patterns. For example, Knorr-Held and Best (2001) use this approach to model two diseases by considering a shared spatial term in the model with a different weight for each disease. Downing et al. (2008) propose a model with several spatial effects to model six cancers jointly. Botella-Rocamora, Martínez-Beneito and Banerjee (2015) and Martínez-Beneito, Botella-Rocamora and Banerjee (2016) propose a general approach for multivariate disease mapping that can help to identify diseases with similar spatial distributions. Marí-Dell'Olmo et al. (2014) use a smoothed analysis of the variance for the analysis of several diseases in ecological models.

We have developed a novel Bayesian spatio-temporal joint model for several diseases with specific and shared spatial and temporal effects. The shared spatial and temporal terms would account for common spatial and temporal patterns. The effect of these common patterns on specific diseases is controlled by specific-weights that measure the dependence of a given disease on these patterns. By considering specific spatial and temporal patterns we allow for departures from the shared patterns for different diseases. Finally, the posterior distribution of the weights is able to capture dependence between diseases with similar spatial or temporal patterns.

Bayesian model fitting has been tackled by using Markov chain Monte Carlo (MCMC) methods (Gilks et al., 1996), which can be slow for complex spatio-temporal models. For this reason, we have also fit the models presented using the approximation provided by the integrated nested Laplace approximation (INLA) method (Rue, Martino and Chopin, 2009). INLA is able to fit the proposed models in a fraction of the time required by MCMC and provide an approximation to the posterior marginals of the model parameters. Given that INLA focuses on approximating the posterior marginals of the model parameters, multivariate posterior inference on several parameters may be difficult to do with INLA and we will still rely on MCMC for this.

The structure of this paper is as follows. First, we give an introduction to spatio-temporal disease mapping in Section 2. In Section 3 several models for the joint analysis of several diseases are described. Next, our new spatio-temporal model is fully described in Section 4. An example on three death causes in Spain is discussed in Section 5. Finally, a summary of the paper and discussion of the main results is given in Section 6.

2. Spatio-temporal disease mapping

Disease mapping (Lawson, 2013, Elliot et al., 2000, Banerjee, Carlin and Gelfand, 2014) is commonly employed in public health and epidemiology in order to describe the spatial (and temporal) variation of disease. In the analysis of public health data, we often find the number of cases at different n administrative areas and T time periods. We will denote by $O_{i,t}$ the number of cases in area i and time period t. As studying the distribution of the cases alone is misleading, expected number of cases $E_{i,t}$ are also computed using the population structure and direct or indirect standardization (Elliot et al., 2000). In addition, area-level covariates $X_{i,t}$ may be available and these can be incorporated into the models to account for socio-economic inequalities, risk exposure and other relevant risk factors.

In order to model the observed number of cases, a Poisson distribution is often used:

$$O_{i,t}|E_{i,t},\theta_{i,t} \sim Po(E_{i,t}\theta_{i,t}) \tag{1}$$

Here, $\theta_{i,t}$ is the relative risk. Values of the relative risk higher that one indicate an area of increased risk because, in that case, the mean $\mu_{i,t} = E_{i,t}\theta_{i,t}$ is higher than the expected number of cases according to the population in the area. As stated above, it is more informative to map the relative risk $\theta_{i,t}$ than the observed cases.

The relative risk can be modelled using a Poisson log-linear model. For example, if the relative risk is thought to be dependent of area-level covariates it can be modelled as:

$$\log(\theta_{i,t}) = \alpha + \beta X_{i,t}, \tag{2}$$

with α an intercept and β a vector of coefficients of covariates $X_{i,t}$. Other fixed and random effects or smooth terms can be added on the right-hand side of the previous equation, as discussed below.

Bayesian hierarchical models for disease mapping have been widely used since the seminal paper by Besag et al. (1991) was published. In this paper, the relative risk depends on area level covariates, spatially correlated random effects v_i and independent random effects u_i. This model can be extended to the spatio-temporal case as follows:

$$\log(\theta_{i,t}) = \alpha + \beta X_{i,t} + v_i + u_i + w_t \tag{3}$$

Here, v_i is a spatial random effect, u_i is an independent random effect and w_t is a temporal effect.

Independent random effects u_i are assigned a Gaussian prior with zero mean and precision τ_u. Spatially correlated random effects follow an intrinsic conditionally autoregressive (CAR) specification. In this case, the conditional distribution of v_i given all the other spatial effects v_{-i} is Gaussian with mean $\sum_{j\neq i} w_{ij}v_j / \sum_{j\neq i} w_{ij}$ and precision $\tau_v \sum_{j\neq i} w_{ij}$. In the previous expressions, w_{ij} are spatial weights and τ_v is the precision of the spatial random effect.

Spatial weights w_{ij} are often taken as 1 if areas i and j are neighbors and 0 otherwise. In this case, if by $i \sim j$ we denote that regions i and j are neighbors and by n_i the number of neighbors of region i, the conditional distribution of v_i under an intrinsic CAR specification is

$$v_i|v_{-i} \sim N\left(\sum_{j\sim i} \frac{v_j}{n_i}, \tau_v n_i\right) \tag{4}$$

When a vector of random effects $v = (v_1, \ldots, v_n)$ has a prior that is an intrinsic CAR specification with weight matrix W and precision τ_v, we will write that as $v \sim CAR(W, \tau_v)$.

Temporal effects w_t are often assigned a random walk prior with precision τ_w or a CAR prior in one dimension, with a temporal adjacency defined so that consecutive time periods are neighbors. Knorr-Held (2000) describes a number of space-time interactions that could be added to the model in Equation (3).

Regarding models that specifically try to identify differential spatial or temporal patterns, Richardson, Abellan and Best (2006) propose a joint model for two diseases with specific space and temporal terms for the second disease which allows for the identification of disease-specific patterns. Abellan et al. (2008) propose a spatio-temporal model with a spatio-temporal term that is a mixture of terms. Each term is Normally distributed with zero mean and one has a smaller variance than the other. This allows the model to classify areas according to small or large variation. Areas with large variance indicate a strong departure from the common separable spatio-temporal pattern.

Similarly, Guangquan et al. (2012) propose a Bayesian hierarchical spatio-temporal model in which the log-relative risk is modelled on a mixture of two linear predictors with different effects. The first one is the sum of an intercept, a spatial effect and a temporal effect and, the second one is the sum of an area-specific intercept and an space-time non-separable random effect.

Both Abellan et al. (2008) and Guangquan et al. (2012) propose models that include terms to highlight areas with patterns that differ from the overall spatio-temporal pattern by using a mixture of terms. These models are aimed at targeting areas which depart from the shared spatial and temporal patterns. In Section 4 we propose a new multivariate model to identify diseases with specific spatial or temporal patterns that are different from the shared spatio-temporal pattern.

3. Joint modelling of multiple diseases

The models described in the previous section can be applied to different diseases to produce space-time risk estimates that can be mapped and analysed to identify particular patterns of high risk. Diseases with similar etiologies may show similar patterns, i.e., similar spatial or temporal variation, and a multivariate analysis could be performed to

obtain better estimates of these shared patterns. At the same time, the model must allow for specific departures from the shared pattern in certain areas.

In order to build joint models for D different diseases, we will denote by $O_{i,t}^{(d)}$ and $E_{i,t}^{(d)}$ the observed and expected cases, respectively, of disease d in area i and time period t. Hence, the distribution of the number of cases $O_{i,t}^{(d)}$ is a Poisson with mean $E_{i,t}^{(d)}\theta_{i,t}^{(d)}$, where $\theta_{i,t}^{(d)}$ is the relative risk.

In a joint model, relative risks include terms that are shared by several diseases. The effect of the shared effects may be weighted, so that these weights measure the dependence of the geographic or temporal distribution of the disease on the shared pattern. For example, Knorr-Held and Best (2001) consider a model for two diseases in which the shared spatial effect has weight δ for one disease and $1/\delta$ for the second disease. δ is a parameter that is estimated and it measures the dependence of each disease on the shared pattern.

For example, this joint model for two diseases could be written down as:

$$\begin{aligned} O_i^{(d)}|E_i^{(d)},\theta_i^{(d)} &\sim Po(E_i^{(d)}\theta_i^{(d)}),\ d=1,2 \\ \log(\theta_i^{(1)}) &= \alpha^{(1)} + \delta S_i + D_i^{(1)} \\ \log(\theta_i^{(2)}) &= \alpha^{(2)} + \frac{1}{\delta} S_i + D_i^{(2)} \end{aligned} \tag{5}$$

Here, $\alpha^{(d)}$ is a disease-specific intercept, S_i is the shared spatial pattern, and $D_i^{(d)}$ are disease-specific (spatial) patterns.

The model by Knorr-Held and Best (2001) can be extended to consider more than two diseases. For instance, Downing et al. (2008) develop a joint model for six smoking related cancers in the Yorkshire region of England. They used a Bayesian model with shared effects to explore the patterns of spatial correlation and to estimate the relative weight of some covariates like smoking and other shared risk factors.

Several authors have generalized the univariate spatial models to the multivariate case in a number of ways, such as the spatial factor modelling proposed by Wang and Wall (2003) or the smoothed analysis of variance proposed by Zhang, Hodges and Banerjee (2009). Other multivariate disease mapping proposals are based on Gaussian Markov random fields (Rue and Held, 2005) and multivariate conditional autoregressive distributions.

A multivariate conditional autoregressive distribution is a generalization of the conditional autoregressive distribution (Mardia, 1988). Gelfand and Vounatsou (2003) generalized the proper conditional autoregressive distribution to the multivariate setting. In Jin, Carlin and Banerjee (2005), the authors propose a conditional approach to the multivariate problem too. MacNab (2011) proposed a multivariate generalization of spatial structures beyond conditional autoregressive distributions, where the well-known convolution prior (Besag et al., 1991) is generalized. Martínez-Beneito (2013) proposed a novel framework that encompasses most of the models already proposed by reorga-

nizing the Kronecker products of covariance matrices as simple matrix products. This allows the combination of several different spatial structures with different multivariate dependence structures and avoid computations with Kronecker products and large covariance matrices. This last work has been reformulated in order to be more efficient in computational terms (Botella-Rocamora et al., 2015). Other recent approaches for the analysis of multivariate data in disease mapping include the smoothed analysis of variance (Marí-Dell'Olmo et al., 2014) in ecological studies.

Regarding the analysis of multivariate disease mapping models with several components in the linear predictor, different authors have already tackled this problem. Corberán-Vallet (2012) apply the shared components approach to the detection of disease outbreaks proposing a multivariate model in which spatial shared components are multiplied by indicator variables to select one of the components. Carroll et al. (2016) propose a space-time mixture model that includes in the linear predictor a purely spatial term, a spatio-temporal term or a mixture of the two. Carroll et al. (2017) apply these ideas to the spatio-temporal analysis of two types of respiratory cancers and allow for the temporal variation of the coefficients of the covariates. In Carroll et al. (2017), three different types of cancer are analysed jointly and they propose mixture models to choose among different spatial, temporal and spatio-temporal terms in the linear predictor. Finally, Lawson et al. (2017) present similar mixture models with spatially and spatio-temporally varying mixture parameters.

In the next section we develop a joint spatio-temporal model for multiple diseases. This model includes two types of spatial and temporal effects, to account for the shared pattern and allow for disease-specific patterns. In addition, the weights associated to the shared spatial and temporal effects retain the associations between different diseases with similar spatial or temporal variation. It is worth noting that our model provides a simpler and more modular specification of the different spatial and temporal effects in the model than the models discussed above and it is still able to find diseases with similar spatial and temporal patterns.

Our approach differs from previous literature in a number of ways. First of all, our goal is to detect similar spatial or temporal behaviors of different diseases in a simple way. The structure of our model is different as well, as it is not based on mixture models but on spatial and temporal shared components. The application is also different, because our aim is not to detect changes at the area level in space and time but to identify shared and specific spatial and temporal patterns that can lead to the identification of diseases with a similar aetiology.

4. Spatio-temporal joint modelling of multiple diseases

When modelling spatio-temporal data our aim is to identify shared and specific patterns of disease both in space and time. For this reason, our model will combine several ideas from the models outlined in Sections 2 and 3. In particular, our model is as follows:

$$
\begin{aligned}
O_{i,t}^{(d)} | E_{i,t}^{(d)}, \theta_{i,t}^{(d)} &\sim Po(E_{i,t}^{(d)} \theta_{i,t}^{(d)}) \\
\log(\theta_{i,t}^{(d)}) &= \alpha^{(d)} + \Phi_i^{(d)} + \Psi_t^{(d)}
\end{aligned}
\tag{6}
$$

Now, $\alpha^{(d)}$ are disease-specific intercepts, and $\Phi_i^{(d)}$ and $\Psi_t^{(d)}$ are spatial and temporal effects for disease d in area i and time period t, respectively. These two effects are defined by including disease-specific and shared patterns in the following way:

$$
\begin{aligned}
\Phi_i^{(d)} &= u_i^{(d)} + \delta_d^S U_i \\
\Psi_t^{(d)} &= v_t^{(d)} + \delta_d^T V_t
\end{aligned}
\tag{7}
$$

In the previous equation we can find shared and disease specific effects. The effect of the shared spatial effect U_i on the relative risk is modulated through weights δ_d^S. Similarly, the effect of the shared temporal pattern on the relative risk is controlled via weights δ_d^T.

The vectors of disease-specific and shared effects are defined using an intrinsic CAR specification:

$$
\begin{aligned}
u^{(d)} &\sim CAR(W, \tau_d^S) \; d = 1,2,3; \quad U \sim CAR(W, \tau_0^S) \\
v^{(d)} &\sim CAR(Q, \tau_d^T) \; d = 1,2,3; \quad V \sim CAR(Q, \tau_0^T)
\end{aligned}
\tag{8}
$$

Here, W is the spatial adjacency matrix and Q defines a temporal adjacency structure. Finally, τ_d^S, τ_d^T, τ_0^S and τ_0^T are the precisions of the different effects.

Note that the previous model does not account for space-time interactions. These could be included but additional constraints would be needed (Knorr-Held, 2000, Richardson et al., 2006, Goicoa et al., 2018), making the model more complex. By adding disease-specific spatial and temporal effects we are already allowing for departures from any shared spatial and temporal trends. This means that the diseases under study may have different spatial or temporal behavior. Furthermore, uncorrelated random effects have not been considered for the same reason.

Regarding the priors for the remainder of the parameters, several options can be considered. Disease specific intercepts $\alpha^{(d)}$ are assigned improper flat priors. Spatial and temporal weights have been assigned a log-Normal prior with zero mean and precision 1/5.9 (similarly as in Downing et al., 2008).

This assumes that weights are positive, but the prior 0.25 and 0.975 quantiles are 0.0086 and 116.8319, which allows for ample variation in the values of the weights. This will also imply that the weights can take very small values. Small weights will produce a negligible effect of the shared spatial or temporal terms in the linear predictor even if the weights are not exactly zero. Hence, it is not necessary that the diseases in the model are correlated in advance as the model can produce very small weights in this case.

Constraining the weights to be positive also means that high values of the shared effects will indicate a similar higher mortality pattern for all the diseases with non-

negligible weights. This is important in order to interpret the results and the role of the shared spatial and temporal terms.

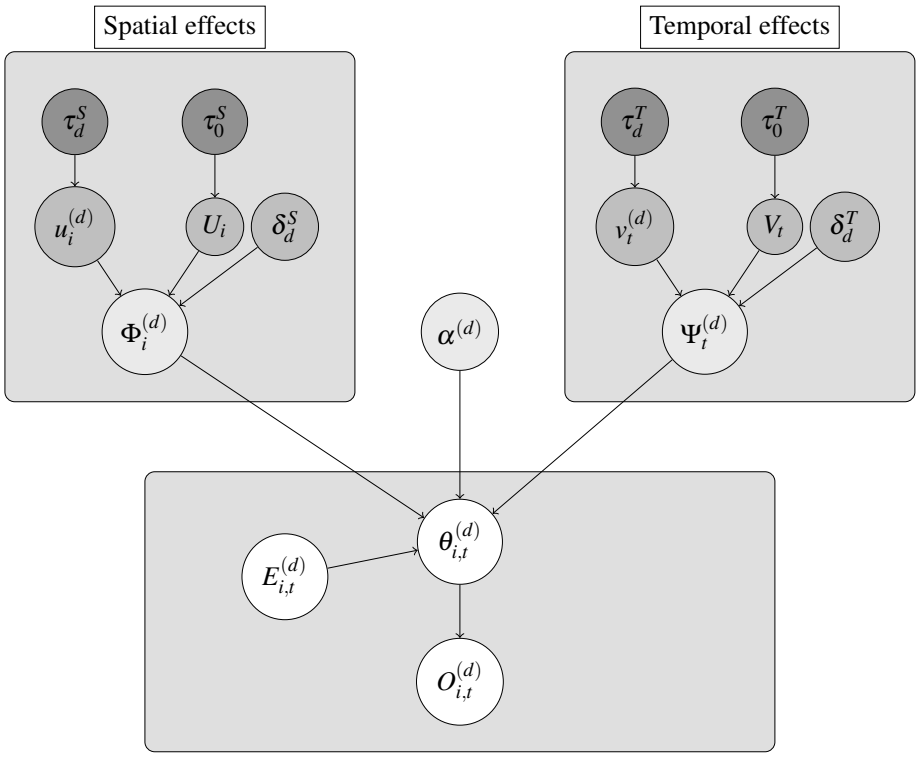

Figure 1: *Graphical representation of the joint spatio-temporal model.*

For the scale parameters of the random effects in the model we suggest trying different priors in order to conduct a sensitivity analysis on the results and investigate how different priors impact on the estimates of the relative risks and other parameters in the model. We propose fitting three different models in which all scale parameters have the same priors. First of all, we propose a uniform distribution between 0 and 10 on the standard deviations, which seems to be less informative than inverted Gammas on the precisions (Gelman, 2006). Following Gelman (2006), a half-Cauchy (with scale parameter equal to 25) as a prior for all standard deviations in the model could also be used. Finally, as an inverted Gamma is a common choice for the precision priors, a third model could be considered in which all precisions have an inverted Gamma with parameters 0.01 and 0.01 as prior.

In this model, terms U_i and V_i in the model are multiplied by disease specific weights. This may cause an identifiability problem between weights δ_d^S and δ_d^T, and the scale of the effects, i.e., precisions τ_0^S and τ_0^T. For this reason, improper priors on these parameters are not recommended. Furthermore, precisions τ_0^S and τ_0^T can be set to 1 so that the scale of spatial and temporal shared effects is incorporated into weights δ_d^S and δ_d^T.

Having said this, we have not observed any identifiability problem of effects $\Phi_i^{(d)}$ and $\Psi_t^{(d)}$ in the models fitted in the example in Section 5.

Spatial dependence between two or more diseases can be assessed by looking for correlation of weights $\{\delta_d^S\}_{d=1}^D$ in the posterior joint distribution. Similarly, temporal dependence can be assessed with the joint posterior correlation of weights $\{\delta_d^T\}_{d=1}^D$. For this reason, we will produce plots of the bivariate posterior distributions of $(\delta_k^S, \delta_l^S)_{k\neq l}$ and $(\delta_k^T, \delta_l^T)_{k\neq l}$ for all pairs of diseases to assess any posterior correlation between the weights. Furthermore, for a single disease, spatio-temporal interactions can also be inspected by considering correlations in the joint posterior distribution of (δ_d^S, δ_d^T). This analysis based on the bivariate joint posterior distributions is shown in the example developed in Section 5 using the MCMC output given that INLA focuses on marginal inference.

5. Example: Joint spatio-temporal disease mapping in Spain

In order to assess the qualities and properties of the model presented in the previous section, we develop here an example on the analysis of three causes of death in Spain. We have considered oral cavity (which includes lip, bucal cavity and pharynx), esophagus and stomach cancer. The International Classification of Disease (ICD-10) codes for the three causes that we are studying are C00-C14 for the oral cavity cancer, C15 for the esophagus cancer and C16 for the stomach cancer. All these are cancers of the upper gastrointestinal tract and are relatively frequent. Ferlay et al. (2012) has pointed out that gastric cancers were estimated to be the fourth most common cancer and the second leading cause of death in both sexes in 2008. Furthermore, oral cavity and pharyngeal cancers ranked eighth in number of new cancer cases and deaths. Also, esophageal cancer ranked sixth in terms of the number of deaths and ninth in terms of cases.

In Spain, López-Abente et al. (2007, 2014) and Aragonés et al. (2007) have studied the spatial and temporal trends of these cancers. They have provided evidence of the similarities among the spatial and temporal trends of these cancers. In particular, their analysis of oral cavity, pharynx and esophagus supported the hypothesis of shared risk factors (which could be preventable factors), such as alcohol consumption and smoking (Seoane-Mato et al., 2014). These tumors also share a South-North geographical pattern in Spain.

Population and mortality data have been obtained from the Spanish Office for National Statistics (INE). Population data contains records by age group and gender from 1996 to 2014. Mortality data comprises all deaths in Spain from 1985 to 2014, for which cause of death, age, gender and other relevant information is available.

In this analysis, the number of deaths per province in peninsular Spain in the period 1996 to 2014 has been considered. The expected number of cases per province and sex has been computed using as reference the age-specific mortality rates and the population from years 1996 to 2014. The analysis has been carried out at the province level for both

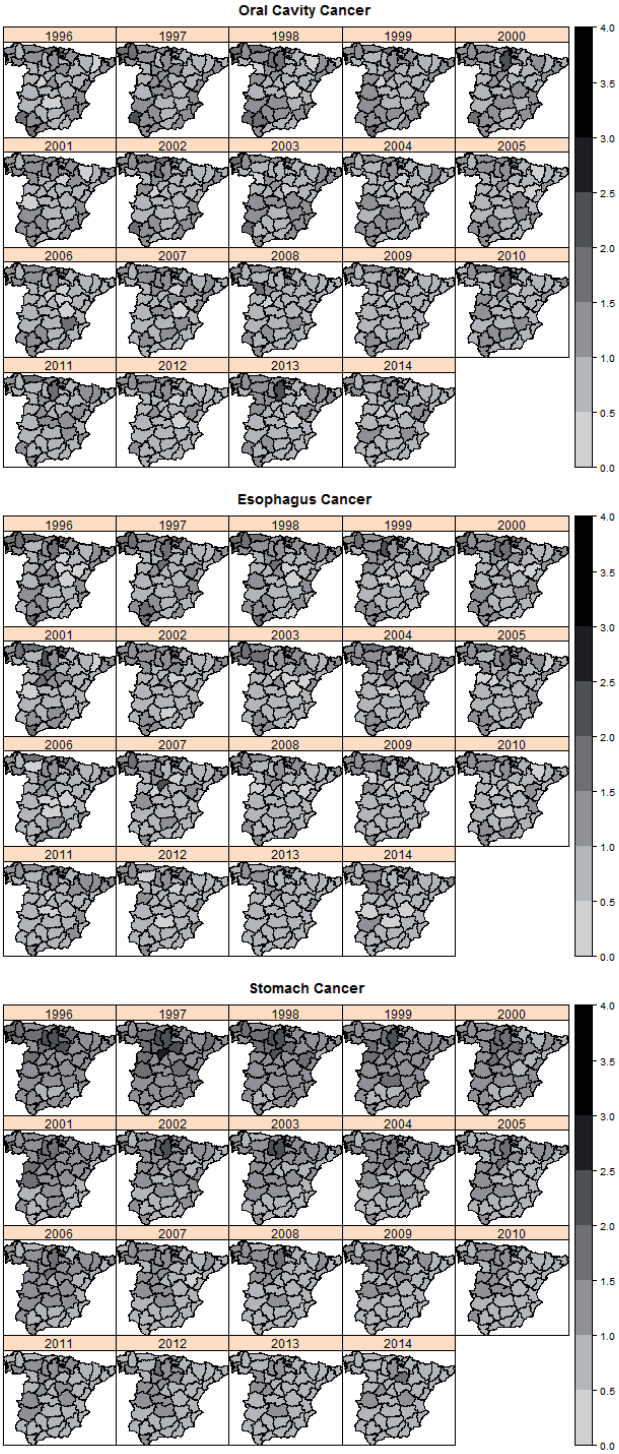

Figure 2: *Standardized mortality ratios* $O_{i,t}^{(d)}/E_{i,t}^{(d)}$.

sexes together. See López-Abente et al. (2014) for a discussion on the importance of the criteria for computing the expected number of cases in a spatio-temporal analysis.

Figure 2 shows the standardized mortality ratios for the three causes of death. Oral cavity and esophagus cancers seem to have a similar spatio-temporal pattern, whilst stomach cancer shows a different pattern. However, the three causes seem to have a region of high risk in the north of the country. These spatial patterns have already been described by Aragonés et al. (2009); López-Abente et al. (2014a,b) for stomach cancer and by Aragonés et al. (2007) for esophageal cancer for slightly different time periods to the one considered now. Furthermore, López-Abente et al. (2014); Seoane-Mato et al. (2014) also describe a decreasing temporal pattern of the risk for stomach and esophageal cancers. Finally, López-Abente et al. (2007) provide a spatial analysis of a number of types of cancer from 1989 to 1998 in Spain at the municipality level. Although our analysis has been conducted at different spatial and temporal levels, we observe a very similar pattern and we expect these patterns to show up in the analysis and to be picked up by the different effects in our model.

The model that we have fitted to the data is the one described in Section 4. The results that we show here correspond to the model with uniform priors on the standard deviations of the spatial and temporal random effects. We have also fitted the same model using half-Cauchy on the standard deviations and inverted Gamma priors on the precision parameters. A summary is provided in the sensitivity analysis in Section 5.4. Models have been fitted using the WinBUGS software (Lunn et al., 2000) using the R2WinBUGS package (Sturtz, Ligges and Gelman, 2005) for the R software (R Core Team, 2016). Regarding the MCMC simulations, we have used 4 different chains with 200,000 simulations each, of which 10% (i.e., 20,000) were used as a burn-in and we have kept one in 200 simulations to reduce autocorrelation.

In addition, INLA has been used to estimate the posterior marginals of the parameters of the models presented above. However, given the way in which INLA computes the approximations, a uniform between zero and infinity has been used instead of a uniform between 0 and 10 on the standard deviations. Details on the construction of the priors for INLA as provided in Appendix A, and computational details and R code to fit the models using MCMC and INLA are provided in the supplementary materials provided with this paper available from https://github.com/becarioprecario/joint_st_disease_mapping_INLA.

5.1. Spatial analysis

First of all, we will consider the analysis of the different spatial effects in the model. Figure 3 shows the posterior means of the total spatial effect $\Phi_i^{(d)}$ (i.e., sum of shared plus specific effects). MCMC and INLA provide very close estimates of the posterior means. Oral cavity and esophagus cancer show very similar spatial patterns, with areas of high risk in the north-west and southwest. This pattern is similar to the spatial distribution of

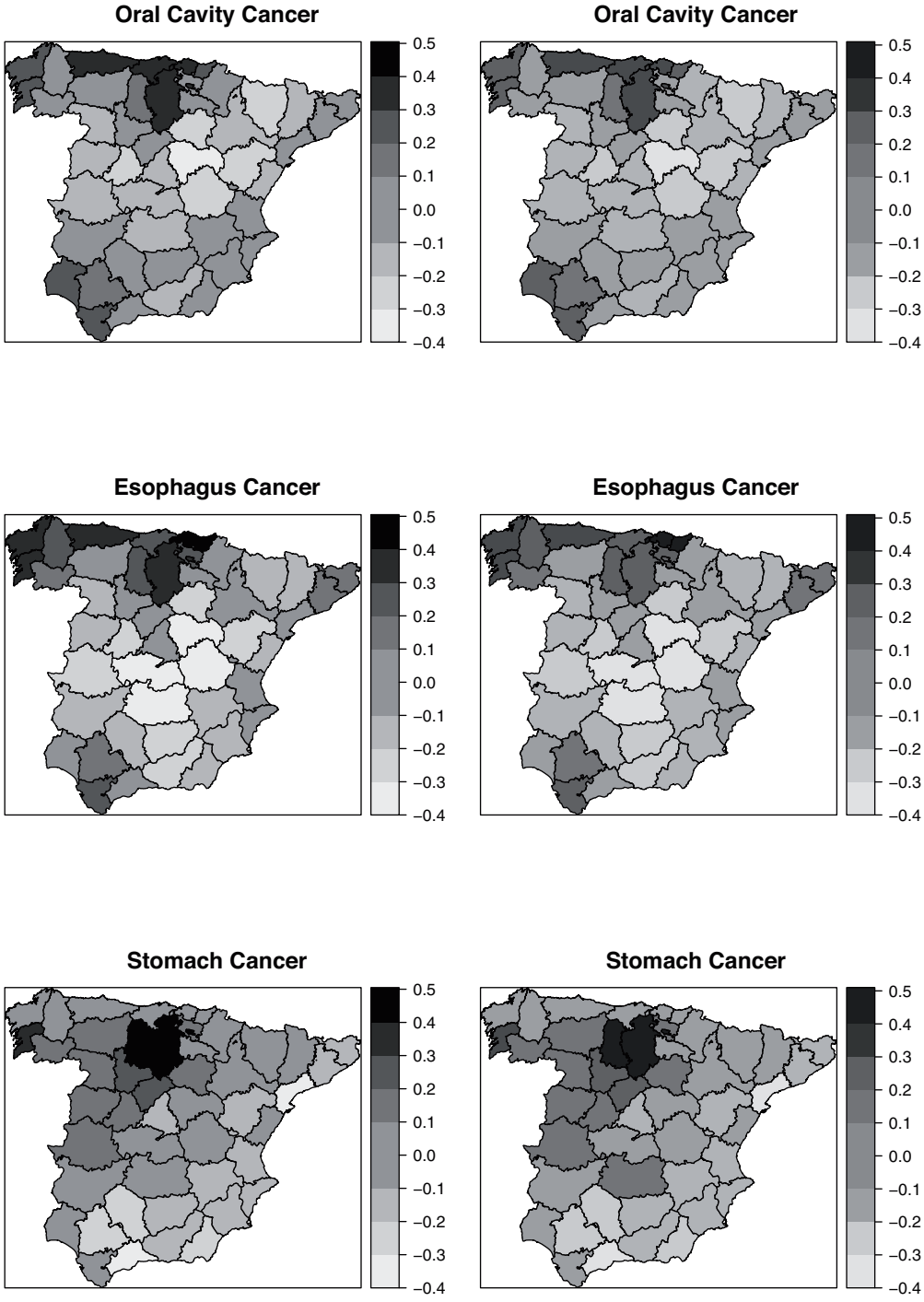

Figure 3: *Posterior means of the spatial effect* $\Phi_i^{(d)} = u_i^{(d)} + \delta_d^S U_i$ *for MCMC (left) and INLA (right).*

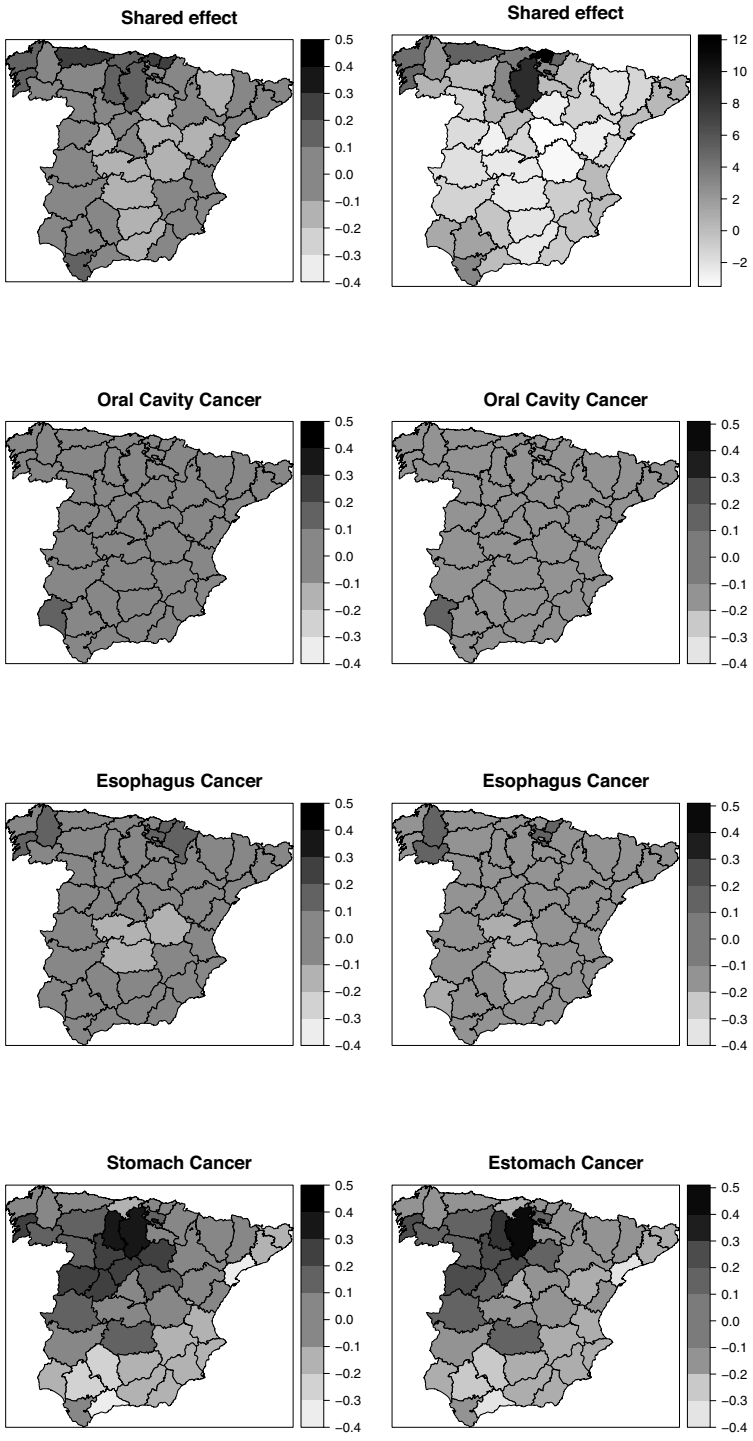

Figure 4: *Posterior means of shared spatial effect U_i (top maps) and disease specific spatial effects $u_i^{(d)}$ for MCMC (left) and INLA (right).*

Table 1: *Summary statistics of the weights for shared spatial and temporal effects for MCMC (left) and INLA (right).*

	MCMC				INLA			
Parameter	**Mean**	**Median**	**2.5% q.**	**97.5% q.**	**Mean**	**Median**	**2.5% q.**	**97.5% q.**
δ_1^S	1.455	1.425	0.854	2.263	0.716	0.371	0.022	3.501
δ_2^S	1.555	1.538	0.864	2.469	0.766	0.399	0.024	3.735
δ_3^S	0.551	0.546	0.332	0.914	0.062	0.031	0.001	0.297
δ_1^T	0.867	0.807	0.434	1.484	0.204	0.088	0.006	1.122
δ_2^T	0.943	0.892	0.515	1.483	0.270	0.109	0.007	1.530
δ_3^T	1.279	1.220	0.659	2.148	0.395	0.164	0.120	2.210

esophageal cancer between 1989 and 1998 described in Aragonés et al. (2007). Stomach cancer shows a different spatial pattern with some areas of high risk in the north. These findings are similar to the spatial patterns described by Aragonés et al. (2009) in the period 1994-2003, and López-Abente et al. (2014a,b) in the period 1989-2008.

Posterior means of the shared and disease specific spatial effects are displayed in Figure 4, and Table 1 shows summaries of the posterior distribution of the weights δ_d^S of the shared spatial effect for each disease. The estimates of the different spatial effects with MCMC and INLA are very similar but for the shared spatial term, which seems to show a very similar pattern but at different scales. This is probably due to a mild identifiability problem between the spatial weights and the precision of the shared spatial term. However, as stated above, total spatial effects are very similar between MCMC and INLA.

In all maps in Figure 4, a few areas of high risk can be found in the north part of the country. Also, the specific spatial pattern for stomach cancer seems to show more extreme values than those for oral cavity and esophagus cancer. This may be due to the lower dependence of stomach cancer on the shared spatial pattern (as seen in Table 1) which makes the specific pattern to account for most of its spatial pattern.

Table 1 shows the differences in the estimation of the weights between MCMC and INLA. This is due to the fact that INLA is not able to identify well the weights and the precision of the effects. However, as seen in Figure 3 and Figure 4 the estimates of the spatial effects are very close between MCMC and INLA. Furthermore, the results obtained with INLA also support a stronger dependence on the spatial shared term for oral cavity and esophagus cancers, and a similar dependence on the temporal shared term for all three cancers. As stated earlier, a simple way to have a better identification of the weights is by fixing the precisions of the shared terms.

Regarding the weights on the shared spatial component U_i, oral cavity and esophagus cancers seem to have a very similar weight which is significantly higher than one. Stomach cancer has a lower weight, which is significantly lower than one. This means that oral cavity and esophagus cancer have a higher dependence on the shared spatial effect, i.e., the spatial pattern is very similar to the shared pattern.

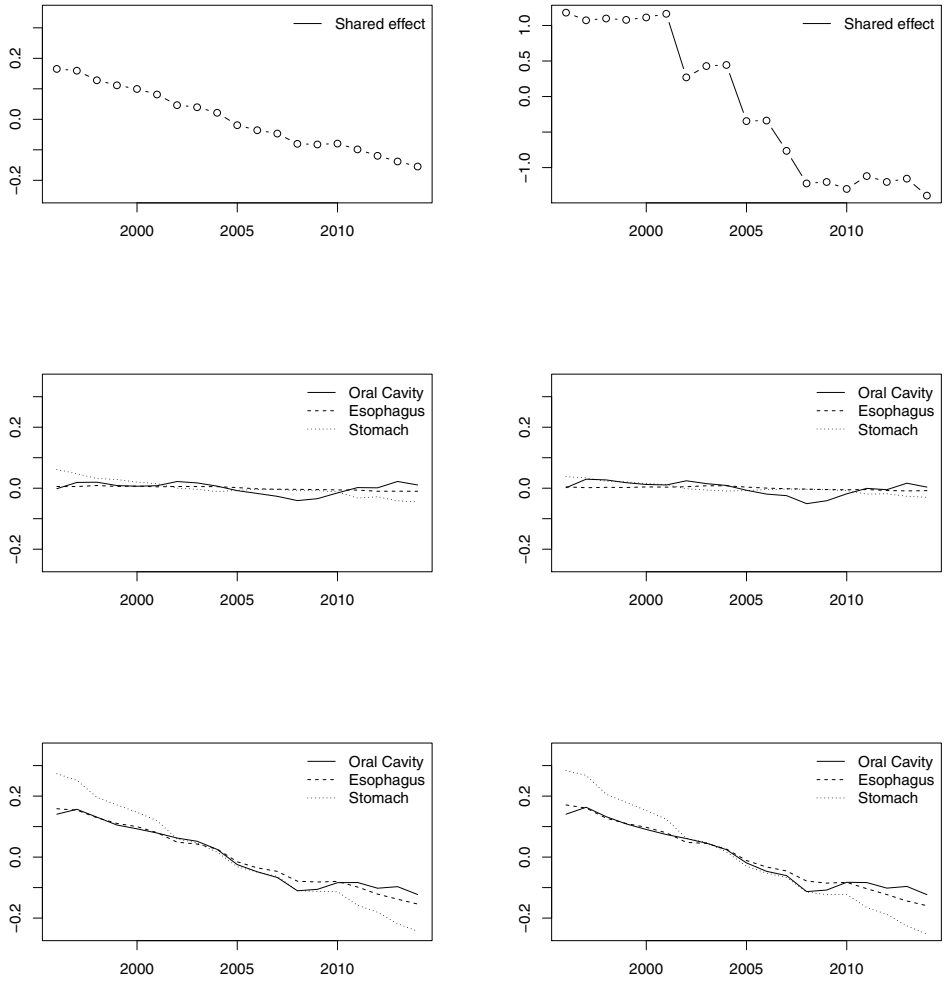

Figure 5: *Posterior means of shared temporal effect V_t (top), specific temporal effect $v_t^{(d)}$ (middle) and total temporal effect $\Psi_t^{(d)} = v_t^{(d)} + \delta_j^T V_t$ (bottom) for MCMC (left) and INLA (right).*

Finally, the dependence between oral cavity and esophagus cancers is confirmed in the analysis of weights δ_i^S on the shared spatial component shown in Section 5.3 using the MCMC output. As seen in Figure 8 (bottom row), the bivariate distribution of weights associated to oral cavity and esophagus cancers shows a strong correlation. This correlation is inexistent in the plots of each one of these causes against esophagus cancer.

5.2. Temporal analysis

Similarly, posterior means of shared and specific temporal effects are shown in Figure 5. The shared temporal effect clearly indicates a decrease in risk over time and MCMC and

INLA provide similar estimates but on different scales. Now, all three cancers show a very similar temporal pattern. The disease-specific and total temporal trends estimated by MCMC and INLA are very close.

The specific temporal effects of the three cancers do not indicate a strong departure from the shared temporal pattern and these three specific temporal patterns have an effect very close to zero for all the years. It is worth noting that the shared temporal effect captures the overall decreasing trend in time whilst the disease-specific effects are negligible, with estimates very close to zero for all years. Seoane-Mato et al. (2014) describe temporal trends for different types of tumors in the period 1952-2006. For all the causes analysed in this paper, they report a decreasing trend from 1996 to 2006, which is consistent with our findings.

Summary statistics of weights δ_d^T for the shared temporal trend are shown in Table 1. As in the spatial case, MCMC and INLA provide estimates in different scales due to the different identifiability between the temporal weights and the precision of the shared temporal term. However, the estimates of the total temporal trends are very similar between MCMC and INLA.

Oral cavity and esophagus cancers have very similar weights, with stomach cancer having a slightly higher weight. In this case, all three diseases seem to have a strong dependance on the shared temporal pattern as the weights are very close to one, which also explains the weak disease-specific temporal trends.

A joint analysis of weights δ_d^T could be done using the MCMC output to assess temporal dependence between diseases. Figure 8 (top row) shows bivariate plots of these weights. Oral cavity and esophagus cancer clearly show some correlation. Now, stomach cancer also shows a positive correlation with the other two types of cancer.

5.3. Joint spatio-temporal analysis

So far, we have analysed the results with a focus on the spatial or temporal patterns. Figure 6 shows the smoothed spatio-temporal relative risks obtained with our model. The three types of cancers considered in this study show correlation of the temporal weights. However, stomach cancer shows a different spatio-temporal pattern.

Figure 7 shows the probability of having a relative risk higher than one. Looking at the areas of high probability we can find areas of increased risk. Again, oral cavity and esophagus cancers show a very similar spatio-temporal pattern, which also seems to be persistent over time. Furthermore, the areas of high risk in our analysis are very similar to the ones reported by Aragonés et al. (2007) in the 1989-1998 period for esophageal cancer, where regions of high risk were found in the northwest and southwest of Spain.

Stomach cancer shows a persistent spatial pattern at the beginning of our study period that changes at the end, as seen in Figure 7. The areas of high risk are similar to those found by Aragonés et al. (2009) in the period from 1994 to 2003, and López-Abente et al. (2014) in the 1989-2008 period.

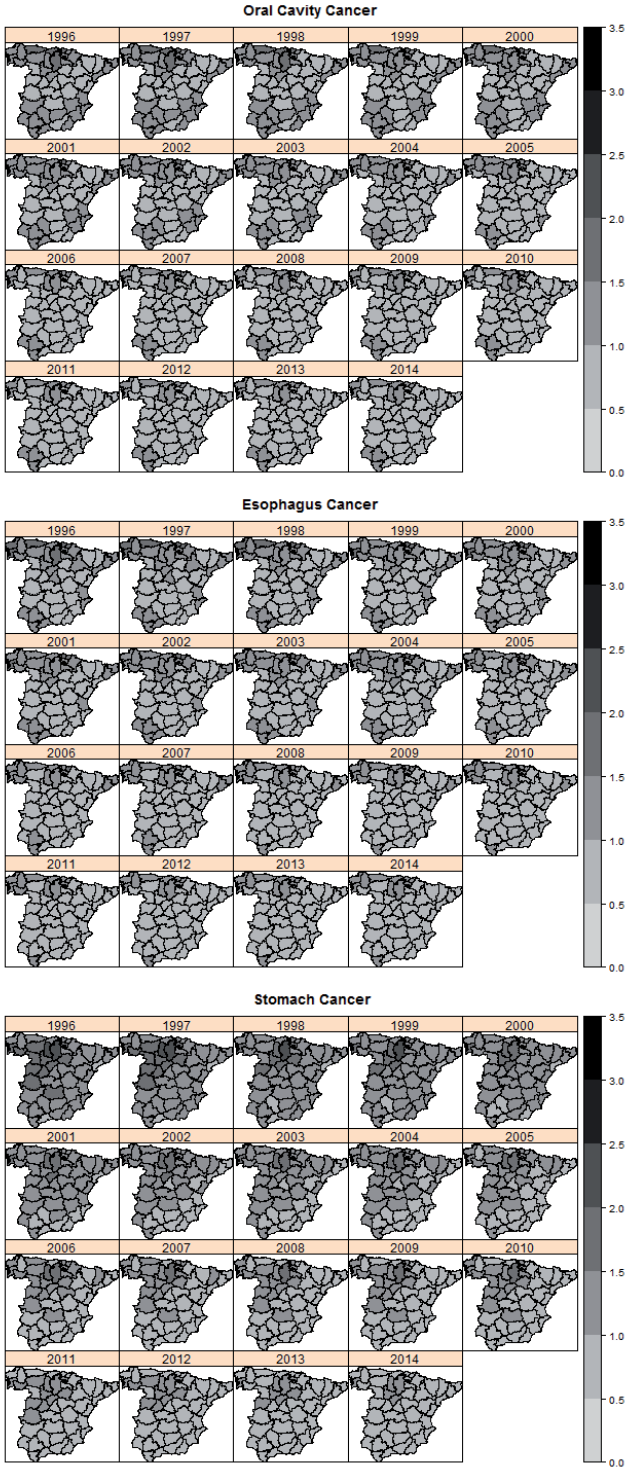

Figure 6: *Posterior means of spatio-temporal relative risks* $\theta_{i,t}^{(d)}$.

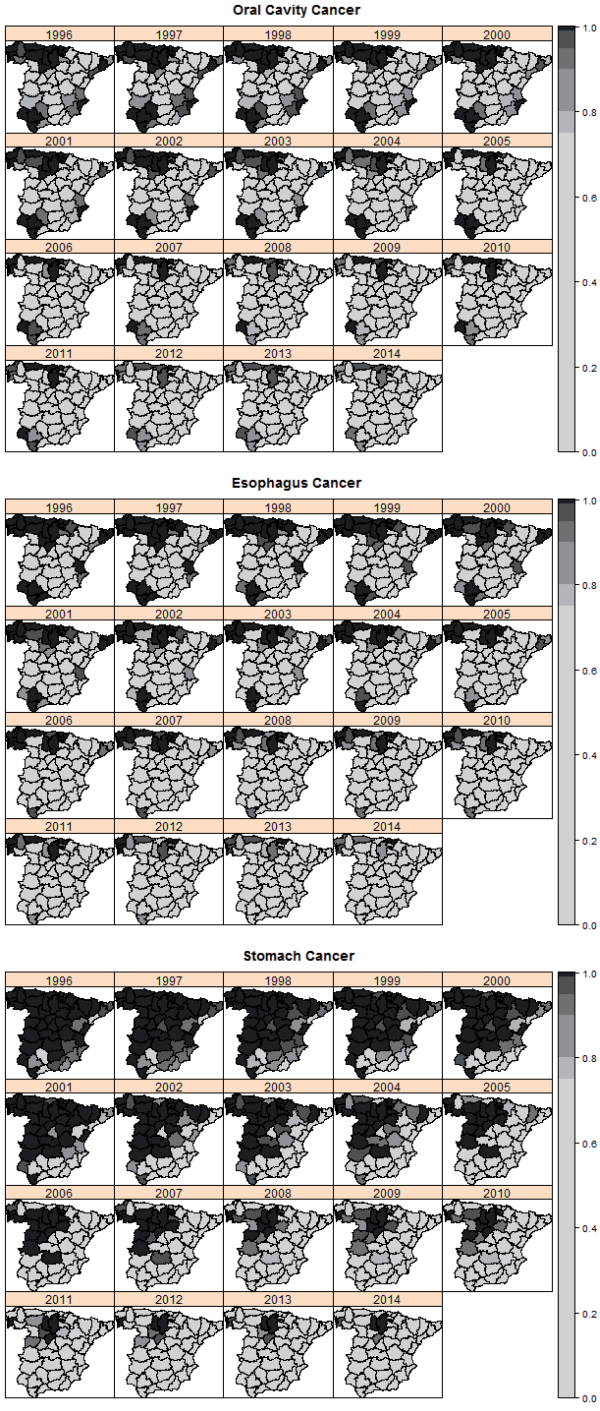

Figure 7: *Probabilities of having an estimation of the relative risk* $\theta_{i,t}^{(d)}$ *greater than 1 to identify areas of high risk.*

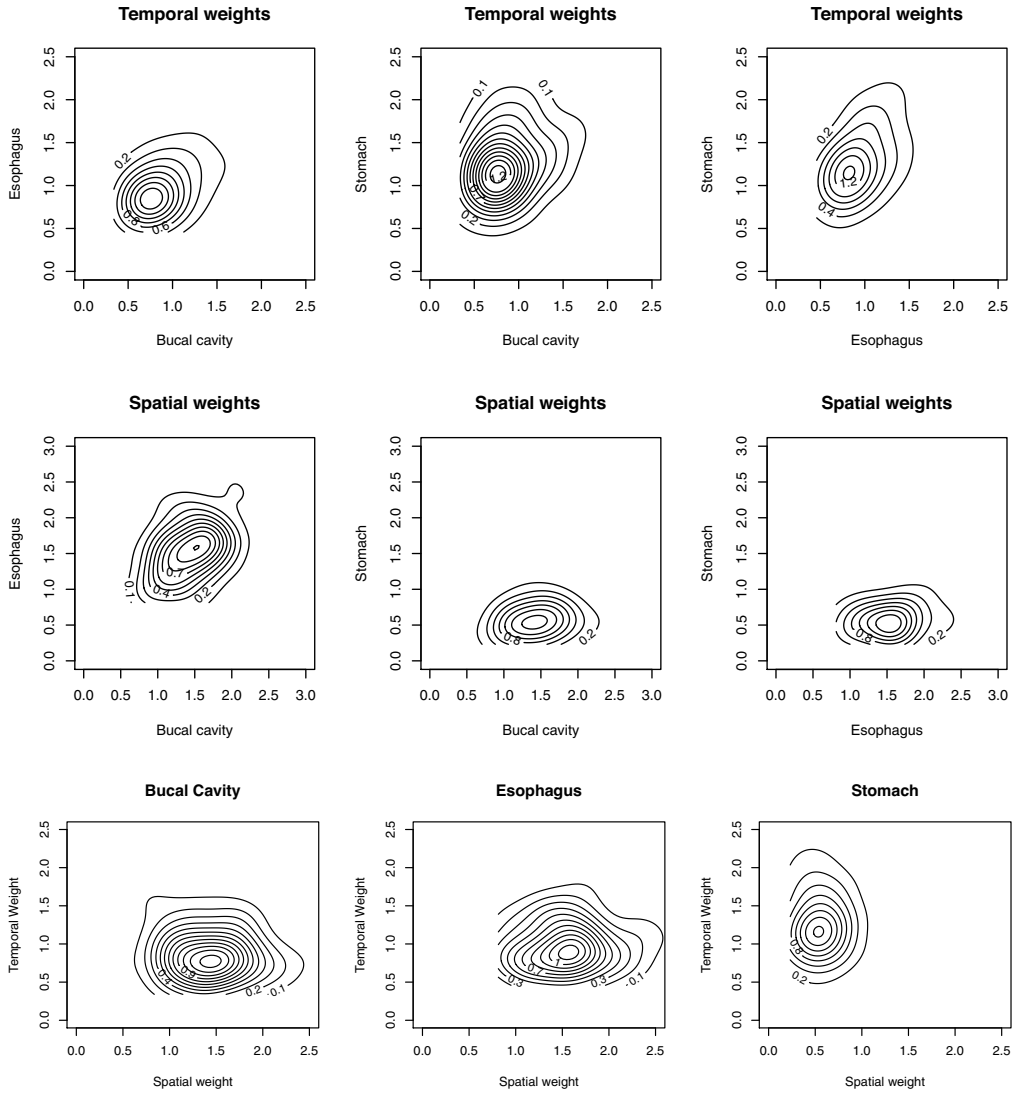

Figure 8: *Bivariate posterior distributions of weights δ_d^T (top row) and δ_d^S (middle row), and bivariate posterior distribution of the spatial and temporal weights for a given type of cancer (bottom row).*

The analysis of the posterior bivariate distribution of weights on the spatial and temporal shared effects can help to assess dependence between the different causes of death considered. Figure 8 shows the posterior bivariate distribution of each pair of weights δ_t^T (top row) and δ_i^S (middle row). In this case, spatial and temporal weights appear to be independent from each other and no correlation can be observed in the plots.

Figure 6 and Figure 7 have been produced using the MCMC output, but INLA provided similar estimates. Figure 8 has been created from the MCMC output given that it requires the bivariate joint posterior distributions of each pair of weights. These joint

distributions can be approximated with INLA, but we have preferred to use the MCMC output instead.

5.4. Sensitivity analysis

The results shown so far correspond to the model described in Section 4 with uniform priors on the standard deviations of the random effects. We have decided to use these priors because several authors (see, for example, Gelman, 2006) have questioned the use of the inverted Gamma as a prior for the precisions in the model. For this reason, we have conducted a sensitivity analysis by considering different priors for the scale parameters of the random effects in the model. We have fitted three versions of the joint spatio-temporal models where each of three options for the priors of the variances are used, as explained at the end of Section 4.

Our results show that the estimates of the relative risks do not differ when different priors for the variances of the random effects are used. Spatial and temporal effects are very similar too, as well as the estimates of the spatial and temporal weights.

6. Discussion

We have presented a Bayesian hierarchical model for the joint analysis of spatio-temporal public health data. It combines ideas from other models for spatio-temporal disease mapping (Richardson et al., 2006; Abellan et al., 2008; Guangquan et al., 2012) and the joint analysis of several diseases (Downing et al., 2008). In this way, our new model allows us to define common and specific spatial patterns of disease that are able to identify similarities and differences in the distribution of the relative risks associated to each disease. Dependence on the common spatial and temporal patterns are governed by disease-specific weights, which can help to identify diseases with shared spatial and temporal patterns. The model has been fitted using MCMC and INLA, and both methods provide similar estimates of the main effects in the model.

The analysis of the specific spatial effects can be used to detect areas with a different trend for a given disease. Similarly, by inspecting disease-specific temporal effects it is possible to highlight diseases with a different temporal variation. Furthermore, this model can help to highlight areas of high risk by looking at the posterior probabilities of the relative risk. These probabilities can also be used to detect shared patterns of high risk among several diseases.

In the example shown in this paper, we have studied oral cavity, esophagus and stomach cancers in Spain from 1996 to 2014. Our model has been able to identify a common spatial pattern between oral cavity and esophagus cancers, and a different spatial pattern for stomach cancer. It has also been able to identify that all these three types of cancer have a very similar temporal variation. All these findings are consistent with other

similar studies (Aragonés et al., 2007, Seoane-Mato et al., 2014, López-Abente et al., 2014a,b, Aragonés et al., 2009) and they support the hypothesis of a strong relationship between the spatio-temporal distribution of oral cavity and esophagus cancers.

Finding diseases with similar spatial and temporal patterns is important in public health because these patterns are often caused by similar risk factors. Hence, by identifying diseases with similar patterns it is also likely that some shared risk factors will be discovered as well. This can clearly be seen in our example as oral cavity and esophagus cancers show strong similar patterns and incidence of these cancers depends of preventable factors such as alcohol consumption and smoking (Seoane-Mato et al., 2014).

Finally, compared to other recent developments for multivariate disease mapping (see, for example, Botella-Rocamora et al., 2015, and the references therein) our model provides a simple and modular specification of different shared and specific patterns that can be explored to identify trends in the geographical and temporal distribution of disease. In the example presented in this paper we have only considered three different causes of death, but the model can be easily extended to a larger number of diseases simply by including the corresponding spatial and temporal effects.

In the future, we plan to extend this model in a number of ways. First of all, an automatic procedure could be implemented to assess for the need of the different disease-specific terms in the model. For example, in our example the disease-specific temporal trends can probably be removed given that all diseases have a very similar temporal variation. For this, being able to fit the models with INLA quickly will allow us to explore different models faster. Furthermore, model assessment criteria implemented in INLA can play an important role here to select the best model for the data.

Another way to extend this model is by clustering diseases into groups so that only diseases within the same group share spatial and temporal terms. This would involve creating a new indicator parameter for each disease to identify to which group it belongs. By computing the posterior probabilities of these indicator variables it is possible to assess what diseases have a shared spatial and temporal variation. Given that this will require exploring a large number of models, INLA will be an important asset in the implementation of this method.

A User-defined priors in INLA

INLA provides a simple way to define priors using the `muparser` library. For computational reasons, INLA works with an internal representation of the parameters and instead of dealing with the precision parameter of the random effects τ, it works with $\theta = \log(\tau)$. Hence, the prior must be specified on θ. Here, we will follow Ugarte, Adin and Goicoa (2016) to derive the two non-implemented priors on the standard deviation σ of the random effects.

First of all, note that $\sigma = (1/\tau^{1/2}) = 1/\exp(\theta/2) = \exp(-\theta/2)$. Hence, the prior on θ is defined as

$$\pi(\theta) = \pi(\sigma)|\frac{\partial\sigma}{\partial\theta}|$$

Also, note that

$$|\frac{\partial\sigma}{\partial\theta}| = \frac{1}{2}\exp(-\theta/2)$$

For the uniform prior on σ, this must be a uniform between 0 and infinity (for computational reasons), i.e., $\pi(\sigma) \propto 1$. Hence,

$$\pi(\theta) \propto 1 \cdot (\frac{1}{2}\exp(-\theta/2))$$

Similarly, the half-Cauchy prior with scale parameter γ on σ is defined as

$$\pi(\sigma|\gamma) = \frac{2}{\pi\gamma(1+(\sigma/\gamma)^2)}$$

Hence, the prior on θ is defined as

$$\pi(\theta|\gamma) = \frac{2}{\pi\gamma[1+(\exp(-\theta/2)/\gamma)^2]} \cdot (\frac{1}{2}\exp(-\theta/2))$$

Priors must be passed to INLA in the log-scale and constants can be dropped (but this will change the estimate of the marginal likelihood). Hence, the uniform prior can be set in INLA as

$$\log(\pi(\theta)) \equiv -\theta/2$$

and the half-Cauchy prior can be set using

$$\log(\pi(\theta|\gamma)) \equiv \log(1+\exp(-\theta)/\gamma^2) - \theta/2$$

Implementation details can be found at
https://github.com/becarioprecario/joint_st_disease_mapping_INLA.

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 (Castilla-La Mancha, Spain) and Fondo Europeo de Desarrollo Regional, and grant MTM2016-77501-P, funded by the Ministerio de Economía y Competitividad (Spain). F. Palmí-Perales was supported by a doctoral scholarship awarded by the University of Castilla-La Mancha (Spain). We also thank Prof. Håvard Rue for his help with the implementation of the model using INLA.

References

Abellan, J., Richardson, S. and Best, N. (2008). Use of space-time to investigate the stability of patterns of disease. *Enviromental Health Perspectives*, 116, 1111–1119.

Aragonés, N., Pérez-Gómez, B., Pollán, M., Ramis, R., Vidal, E., Lope, V., García-Pérez, J., Boldo, E. and López-Abente, G. (2009). The striking geographical pattern of gastric cancer mortality in Spain: environmental hypotheses revisited. *BMC cancer*, 9, 1.

Aragonés, N., Ramis, R., Pollán, M., Pérez-Gómez, B., Gómez-Barroso, D., Lope, V., Boldo, E., García-Pérez, J. and López-Abente, G. (2007). Oesophageal cancer mortality in Spain: a spatial analysis. *BMC cancer*,7, 1.

Banerjee, S., Carlin, B. and Gelfand, A. (2014). *Hierarchical Modeling and Analysis for Spatial Sata*. Crc Press.

Besag, J., York, J. and Mollié, A. (1991). Bayesian image restoration, with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics*, 43, 1–59.

Botella-Rocamora, P., Martínez-Beneito, M. and Banerjee, S. (2015). A unifying modeling framework for highly multivariate disease mapping. *Statistics in Medicine*, 34, 1548–1559.

Carroll, R., Lawson, A. B., Faes, C., Kirby, R. S., Aregay, M. and Watjou, K. (2016). Spatio-temporal bayesian model selection for disease mapping. *Environmetrics*, 27, 466–478.

Carroll, R., Lawson, A. B., Faes, C., Kirby, R. S., Aregay, M. and Watjou, K. (2017). Extensions to multivariate space time mixture modeling of small area cancer data. *International Journal of Environmental Research and Public Health*, 14, 503.

Carroll, R., Lawson, A. B., Kirby, R. S., Faes, C., Aregay, M. and Watjou, K. (2017). Space-time variation of respiratory cancers in South Carolina: a flexible multivariate mixture modeling approach to risk estimation. *Annals of Epidemiology*, 27, 42–51.

Corberán-Vallet, A. (2012). Prospective surveillance of multivariate spatial disease data. *Statistical Methods in Medical Research*, 21, 457–477.

Downing, A., Forman, D., Gilthorpe, M., Edwards, K. and Manda, S. (2008). Joint disease mapping using six cancers in the Yorkshire region of England. *International Journal of Health Geographics*, 7, 1.

Elliot, P., Wakefield, J., Best, N. and Briggs, D. (2000). *Spatial Epidemiology: Methods and Applications*. Oxford University Press.

Ferlay, J., Shin, H., Bray, F., Forman, D., Mathers, C. and Parkin, D. (2012). Cancer incidence and mortality worldwide: Iarc cancerbase no. 10 [internet]. international agency for research on cancer, Lyon, France. globocan 2008 v1. 2.

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

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis*, 1, 515–534.

Gilks, W., Richardson, S. and Spiegelhalter, D. (1996). *Markov Chain Monte Carlo in Practice*. Boca Raton, Florida: Chapman & Hall.

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

Guangquan, L., Best, N., Hansell, A., Ahmed, I. and Richardson, S. (2012). Baystdetect: detecting unusual temporal patterns in small area data via bayesian model choice. *Biostatistics*, 13, 695–710.

Jin, X., Carlin, B. and Banerjee, S. (2005). Generalized hierarchical multivariate car models for areal data. *Biometrics*, 61, 950–961.

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

Knorr-Held, L. and Best, N. (2001). A shared component model for detecting joint and selective clustering of two diseases. *Journal of the Royal Statistical Society, Series A*, 1, 73–85.

Lawson, A. (2013). *Bayesian Disease Mapping: Hierarchical Modeling in Spatial Epidemiology*. CRC press.

Lawson, A. B., Carroll, R., Faes, C., Kirby, R. S., Aregay, M. and Watjou, K. (2017). Spatiotemporal multivariate mixture models for bayesian model selection in disease mapping. *Environmetrics*, 28, e2465.

López-Abente, G., Aragonés, N., García-Pérez, J. and Fernández-Navarro, P. (2014). Disease mapping and spatio-temporal analysis: importance of expected-case computation criteria. *Geospatial Health*, 9, 27–35.

López-Abente, G., Aragonés, N., Pérez-Gómez, B., Pollán, M., García-Pérez, J., Ramis, R. and Fernández-Navarro, P. (2014). Time trends in municipal distribution patterns of cancer mortality in Spain. *BMC cancer*, 14, 1.

López-Abente, G., Ramis, R., Pollán, M., Aragonés, N., Pérez-Gómez, B., Gómez-Barroso, D., Carrasco, J., Lope, V., García-Pérez, J., Boldo, E. and García-Mendizabal, M. (2007). *Atlas Municipal de Mortalidad por Cáncer en España, 1989-1998*. Madrid: Instituto de Salud Carlos III.

Lunn, D., Thomas, A., Best, N. and Spiegelhalter, D. (2000). WinBUGS – a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and Computing*, 10, 325–337.

MacNab, Y. (2011). On gaussian markov random fields and bayesian disease mapping. *Statistical Methods in Medical Research*, 20, 49–68.

Mardia, K. (1988). Multi-dimensional multivariate gaussian markov random fields with application to image processing. *Journal of Multivariate Analysis*, 24, 265–284.

Marí-Dell'Olmo, M., Martínez-Beneito, M., Gotsens, M. and Palència, L. (2014). A smoothed ANOVA model for multivariate ecological regression. *Stochastic Environmental Research and Risk Assessment*, 28, 695–706.

Martínez-Beneito, M. (2013). A general modelling framework for multivarite disease mapping. *Biometrika*, 100, 539–553.

Martínez-Beneito, M., Botella-Rocamora, P. and Banerjee, S. (2016). Towards a multidimensional approach to bayesian disease mapping. *Bayesian Analysis*, 1, 239–259.

R Core Team (2016). *R: A Language and Environment for Statistical Computing*. Vienna, Austria,: R Foundation for Statistical Computing.

Richardson, S., Abellan, J. J. and Best, N. (2006). Bayesian spatio-temporal analysis of joint patterns of male and female lung cancer risks in Yorkshire (UK). *Statistical Methods in Medical Research*, 15, 385–407. PMID: 16886738.

Rue, H. and Held, L. (2005). *Gaussian Markov Random Fields: Theory and Applications*. CRC Press.

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

Seoane-Mato, D., Aragonés,N., Ferreras, E., García-Pérez, J., Cervantes-Amat, M., Fernández-Navarro, P., Pastor-Barriuso, R. and López-Abente, G. (2014). Trends in oral cavity, pharyngeal, oesophageal and gastric cancer mortality rates in Spain, 1952–2006: an age-period-cohort analysis. *BMC cancer*, 14, 1.

Sturtz, S., Ligges, U. and Gelman, A. (2005). R2WinBUGS: A package for running winbugs from R. *Journal of Statistical Software*, 12, 1–16.

Ugarte, M. D., Adin, A. and Goicoa, T. (2016). Two-level spatially structured models in spatio-temporal disease mapping. *Statistical Methods in Medical Research*, 25, 1080–1100. PMID: 27566767.

Wang, F. and Wall, M. (2003). Generalized common spatial factor model. *Biostatistics*, 4, 569–582.

Zhang, Y., Hodges, J. and Banerjee, S. (2009). Smoothed anova with spatial effects as a competitor to mcar in multivariate spatial smoothing. *The Annals of Applied Statistics*, 3, 1805.