2013•2014
FACULTY OF SCIENCES
Master of Statistics

Master's thesis

Space-time latent component modeling of TSH in newborns

Promotor :
Prof.dr. Christel FAES

Kevin Watjou
Thesis presented in fulfillment of the requirements for the degree of Master of Statistics

Transnational University Limburg is a unique collaboration of two universities in two countries: the University of Hasselt and Maastricht University.

universiteit hasselt
KNOWLEDGE IN ACTION

Universiteit Hasselt | Campus Hasselt | Martelarenlaan 42 | BE-3500 Hasselt
Universiteit Hasselt | Campus Diepenbeek | Agoralaan Gebouw D | BE-3590 Diepenbeek

universiteit hasselt | Maastricht University

2013•2014
FACULTY OF SCIENCES
Master of Statistics

Master's thesis

Space-time latent component modeling of TSH in newborns

Promotor :
Prof.dr. Christel FAES

Kevin Watjou
Thesis presented in fulfillment of the requirements for the degree of Master of Statistics

universiteit hasselt | Maastricht University

Acknowledgments

First and foremost I would like to thank my parents for giving me the opportunity to commence this programme and their listing ear whenever I needed to vent my thoughts. Their unrelenting support and love helped me through the dark and tough times.

Next I want to thank my supervisor Professor Christel Faes for being a beacon of wisdom whenever the path that had to be followed was obscured and clouded. Being the ideal guide on this journey, her indispensable advice kept me on the right track and greatly improved my understanding in the matter.

Additionally I would like to thank Professor Andrew Lawson, providing me with 6 desperately needed characters (mv.car), and Professor Francisco Gude Sampedro (Santiago University Clinical Hospital) for providing the data for this project.

Last but definitely not least, I would like to thank the two people without whom I would have descended into the abyss of insanity and despair a long time ago, Ewoud and Marijke. These two individuals kept me on my feet throughout the last couple of years, providing laughter, sharing stress and delivering brain farts. They were my instigation to aim higher and keep on improving.

Abstract

Iodine deficiency is a serious threat to public health throughout the world. The lack of iodine could result in major detrimental diseases and underdevelopments, especially for newborn babies. In Spain, screening of iodine deficiency was done by means of a heel prick in newborns, which allows measuring the concentration of thyroid-stimulating hormones (TSH), a hormone which triggers the thyroid gland. This study concentrates on the area of Galicia, situated in the northwestern part of Spain. The gathered information on the TSH levels served as a guide to highlight areas that could be considered as iodine deficient, in the sense that in iodine deficient areas the mean TSH levels in newborns is higher.

In this paper, a space-time latent component model was constructed in order to study how TSH levels changed in time. Results indicated that there were two distinct components which followed a different temporal trend between 2004 and 2009. Furthermore the characteristics Gender and Feeding Type were investigated as well and found to have an impact on the TSH levels overall.

Contents

1 Introduction

1.1 Background

Iodine has an important function in the human body as it is the keystone of the thyroid hormones triiodothyronine (T_3) and thyroxine (T_4). The iodine elements are actively absorbed from the blood by the thyroid gland to make the T_3 and T_4 hormones. The production of these molecules is regulated by the thyroid-stimulating hormones (TSH) (Irizarry 2014).

However when the dietary intake of iodine falters, the production of TSH increases in order to maintain equilibrium. This might cause the phenomenon called goiter, an enlargement of the thyroid, increasing the area to absorb more iodine. It was estimated that around 2 billion people suffered from this disease, of which approximately 300 million were between the age of 6 and 12 (WHO 2004).

For newborn babies, the TSH values could go up to 39 mU/L (Thureen 2011). Due to the fact that the thyroid hormones of the mother may still be active as well, the first outcomes may not be conclusive and thus another series of tests may be advisable after a couple of weeks in order to confirm the readings. At this stage, a normal TSH reading is situated between 1.7 mU/L and 9.1 mU/L (Roth 2013). As the infant grows older, the maximal normal TSH value will further decrease.

Iodine deficiency can have different consequences across ages. The most severe adverse effects of increased TSH occur during the fetal state and the critical period of brain development, and can result in functional, developmental and irreversible mental abnormalities (WHO 2004; Vitti 2001). Indicators to assess the iodine status are the thyroid size, urinary iodine and the blood constituent TSH. In this study the latter will be used in order to investigate areas with possible worrisome iodine levels.

In most European countries, neurological and neuropsychological impairments are prevalent, caused by mild to moderate iodine deficiency. More specifically in Spain, different regions exhibit prevalence percentages up to 25% of endemic goiter in school-going children. This indicates that there are local areas where people are severely lacking a iodine-rich diet.

1.2 Objective

In this study, information on the TSH levels of all newborns between 2004 and 2009 was collected in each municipality of Galicia, a region in the northwest of Spain. This region is comprised of four provinces: A Coruña, Lugo, Ourense and Pontevedra (see Figure 1). This information was obtained by means of a heel prick. Since Gender and Feeding Type can affect the TSH levels, these effects should be accounted for in the analysis.

The objective of this project is to model the TSH levels over the aforementioned period of time, taking into account the heterogeneous nature of this evolution across the municipalities in order to check whether or not TSH levels change over time across subsets of municipalities. These temporal components will then in turn provide information on the presence of iodine deficiency within the municipalities.

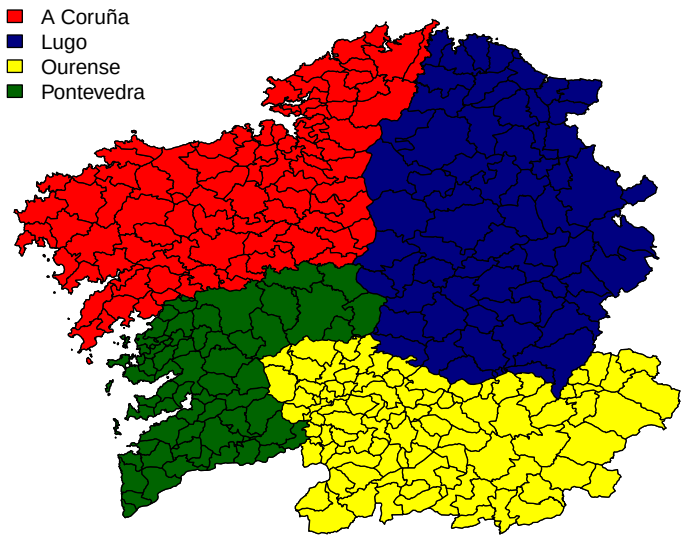


Figure 1: *Province Plot for Galicia*

1.3 Data

The dataset for this project was provided by the Santiago University Clinical Hospital. It contained 97216 observations, providing information on the 315 municipalities of Galicia over a period of 6 years. For each newborn the following characteristics were given:

- *Municipality:* Contains information on the residence of the newborn;
- *Province:* The province in which the baby is living;
- *Year:* Year at which the heel prick was taken (2004 - 2009);
- *TSH:* The TSH level of the newborn (in micro units per liter);
- *Gender:* Categorical variable: 1 for males, 2 for females;
- *Feeding type:* Categorical variable for the method of feeding: 1 for breastfed (= maternal), 2 for a mixture of breastfed and artificial (= mixta), 3 for artificial.

From this dataset several observations were deleted, in order to meet the criteria below:

- TSH values which are not higher than 20 mU/L
- Gestational ages within the range of 38-42 weeks
- Newborns of which the gender, feeding type or birth weight were reported
- Newborns with a birth weight above 500 grams
- Observations coming from the time frame 2004 - 2009

The reduced dataset, which contained 73774 observations, was used for the remainder of the analysis. It should be noted that we did not work with the individual data. The averaged TSH levels for each municipality

2 Methodology

In this project, we wish to look into the temporal variation of TSH across regions, and investigate whether there are groups of areas where the trend is different than in other areas. The use of a latent structure model, as proposed by Lawson et al. (2010), will be investigated for this.

2.1 Model Specification

Define the response of interest, TSH_{fgij} be the average TSH level in municipality i ($i = 1,...,315$), in year j ($j = 1,...,6$), for males/females ($g = 1,2$ for Gender) getting either breast-feeding, mixed or artificial feeding ($f = 1,2,3$ for the different levels of Feeding Type). When specifying the first level of the hierarchy, the assumption is made that the logarithmic transformation of the TSH levels follows a Normal distribution (See Figure 3), namely:

$$log(\text{TSH}_{fgij}) \sim N(\mu_{fgij}, \sigma^2)$$

where μ_{fgij} is the expected mean of $\log(\text{TSH}_{fgij})$ and σ^2 the variance. The mean structure μ_{fgij} consists out of two major components: (1) a predictor which is a function of the fixed effect covariates Gender and Feeding Type, and (2) a mixture component that models the temporal trend in municipalities. This can be defined as:

$$\mu_{fgij} = \beta_{0,j} + \beta_{1,j} \cdot \text{Gender}_g + \beta_{2,j} \cdot \text{Feedtype}_f + \beta_{3,j} \cdot \text{Gender}_g \cdot \text{Feedtype}_f + \Gamma_{ij}$$

The $\beta_{.,j}$ parameters reflect the time-dependent covariate effects and are globally defined from a spatial perspective.

2.1.1 Specification of the Mixture Component

The Γ_{ij} term is a mixture component, indexed for the i^{th} municipality and j^{th} time point, and will be the main focus throughout this paper. This term disaggregates the various temporal

components across different spatial subsets and will thus account for the spatio-temporal variation in the model. The idea behind this proposal of a mixture model is that the time trend corresponding to a particular component could be attached to particular areas, and we are interested in knowing how these areas form groups with similar time trends. We assume the following mixture model:

$$\Gamma_{ij} = \boldsymbol{w}_i' \boldsymbol{\chi}_j = \sum_{l=1}^{L} w_{il} \cdot \chi_{lj}$$

with L latent components χ_{lj} which describe the temporal trends: $\chi_{1j}, ..., \chi_{Lj}$. The parameters w_{il} are the weights given to component l in area i. In this general specification the weights can be spatially varying, allowing to estimate which of the components is better attached to the areas. The choice of the temporal components and weights are further described in the next sections.

2.1.2 The Temporal Component

A model for the unobserved temporal components, $\boldsymbol{\chi}_j = (\chi_{1j}, ..., \chi_{Lj})'$, should be considered. Since it is not clear how the temporal trends behave, we will assume a flexible model for this time trend. We assume that the temporal trend can be described by a autoregressive-dependent model as follows:

$$\begin{aligned} &\chi_{l0} \sim N(0, \sigma^2_{\chi_l}), && l = 1, .., L \\ &\chi_{lj} \sim N(\rho_l \cdot \chi_{l,j-1}, \sigma^2_{\chi_l}), && l = 1, .., L, j = 1, ..., 6 \end{aligned}$$

The mean structure of this distribution consists of two elements: (1) a single-lag autoregressive-dependent term $\chi_{l,j-1}$, smoothing the transition between time points and (2) a parameter ρ_l which controls the dependency on the previous time point. Note that for the starting component χ_{l0} a zero-centered Normal distribution is assumed. Since these temporal components were not observed, assumptions had to be made on their form. This assumption allows for a separate variance parameter $\sigma^2_{\chi_l}$ for each component l.

2.1.3 Specification of Spatial Weights

Moreover the region-specific weights w_{il} are to be specified. These weights reflect the relative contribution of the temporal components to each region and ought to satisfy the obvious restrictions: $w_{il} > 0$ and $\sum_{l=1}^{L} w_{il} = 1$. Hence, these weights follow a probabilistic distribution across the different components for a given municipality. In order to insure that these spatial weights follow the aforementioned criteria, a normalization is in order: $w_{il} = w_{il}^* / \sum_{l=1}^{L} w_{il}^*$. When modeling w_{il}^*, one has to keep in mind that these unnormalized weights have to be positive in order to make sense.

Independent Weights

One possible distribution which guarantees the positivity of the unnormalized weights is the Gamma distribution. This approach implies that the original, normalized spatial weights follow a Dirichlet prior distribution. Also note that in this paper this more straightforward model will not assume any spatial dependence structure.

$$w_{il}^* \sim Gamma(1,1)$$

Spatially Distributed Weights

Another suitable prior distribution for w_{il}^* is the Log-Normal distribution:

$$w_{il}^* \sim LN(\alpha_{1il}, \sigma_w^2)$$

Notice that since we do not expect the unnormalized weights to behave differently across the components, the same variance is applied. When specifying the prior distribution of the weight mean values α_{1il}, a univariate CAR model is applied for each separate component. This has the following specification, as proposed by Besag et al. (1991):

$$\alpha_{1il} | \alpha_{1i'l, i \neq i'} \sim N\left(\frac{1}{n_i} \sum_{i \neq i'} B_{ii'} \alpha_{1i'l}, \frac{\sigma_{\alpha_{1il}}^2}{n_i} \right)$$

Here $B_{ii'}$ contains the neighborhood information: if municipality i is adjacent to municipality i', $B_{ii'}$ is assigned the value 1, and $B_{ii'} = 0$ otherwise. This strategy is also known as the *Queen's Case*. The number of neighboring areas for municipality i is represented by $n_i = \sum_{i \neq i'} B_{ii'}$. The variance term $\sigma^2_{\alpha_{1il}}$ is allowed to vary across the different components.

2.1.4 Number of Components

The only thing left to determine is the number of components that have to be used, since in general the number L is unknown and thus needs to be estimated. Two approaches are possible to select the number of components. The traditional approach is to compare models with a number of components using e.g. DIC. An alternative approach which could be used is the introduction of an entry parameter to each component of the mixture term. This entry parameter will in turn indicate the absence or presence of the respective components, based on their posterior average. This is related to Bayesian model selection. However this method needs to start from a so-called "full" model: a model with a large number of components, in order to make sure no components would be missed. When implementing the entry parameters, ψ_l, the mixture component has the following notation:

$$\Gamma_{ij} = \sum_{l=1}^{L} \psi_l \cdot w_{il} \cdot \chi_{lj}$$

As a prior distribution for the ψ_l-parameters, a Bernoulli distribution is considered:

$$\psi_l \sim Bern(p_l)$$

Thus the components will enter the model with probability p_l. In order to produce a non-informed guess, p_l is assumed to be equal to 0.5 here.

2.1.5 Hyperprior Specification

Since we are working in a Bayesian setting, a prior distribution ought to be given to each parameter. The parameters for the standard deviations used throughout this section assume the

same prior Uniform distribution:

$$\sigma, \sigma_{\chi_l}, \sigma_w, \sigma_{\alpha_{1il}} \sim U(0, 100)$$

The fixed effect parameters $\beta_{0,j}, \beta_{1,j}, \beta_{2,j}, \beta_{3,j}$ are assumed to follow a widely-dispersed Normal distribution:

$$\beta_{.,j} \sim N(0, 10^6)$$

Lastly the temporal dependency parameter ρ_l will be assigned an uninformative Beta prior distribution:

$$\rho_l \sim Beta(1, 1)$$

In short, the complete likelihood of the model which is used for the remainder of the project can be written in the following general form (in vector notation):

$$L(log(\boldsymbol{TSH})|.) = \prod_{i=1}^{315} \prod_{j=1}^{6} N(log(TSH_{fgij})|\boldsymbol{\beta}_j, \boldsymbol{w}, \boldsymbol{\chi}; \sigma^2)$$

Combining this with the aforementioned prior distributions, leads to the following definition of the posterior definition:

$$\begin{aligned} p(\boldsymbol{\beta}, \boldsymbol{w}, \boldsymbol{\chi}, \boldsymbol{p}|log(\boldsymbol{TSH})) =& L(log(\boldsymbol{TSH})|.) \cdot p(\boldsymbol{w}|\boldsymbol{\alpha_1}, \sigma_w) \cdot p(\boldsymbol{\alpha_1}|\boldsymbol{\sigma_{\alpha_1}}) \cdot p(\boldsymbol{\chi}|\boldsymbol{\sigma_\chi}, \boldsymbol{\rho}) \\ & \cdot p(\boldsymbol{\beta}) \cdot p(\boldsymbol{\rho}) \cdot p(\boldsymbol{p}) \cdot p(\sigma_w) \cdot p(\boldsymbol{\sigma_\chi}) \cdot p(\boldsymbol{\sigma_{\alpha_1}}) \end{aligned}$$

Obviously this formula will simplify drastically if no spatial dependence is assumed.

2.2 Model Comparison

A standard measure in Bayesian model selection is the Deviance Information Criterion (DIC). This goodness-of-fit measure can be decomposed into two parts: the deviance $D(\boldsymbol{\theta})$ and the

number of effective parameters pD:

$$D(\boldsymbol{\theta}) = -2 \cdot log\ L(log(\boldsymbol{TSH})|\boldsymbol{\theta})$$
$$pD = \overline{D(\boldsymbol{\theta})} - D(\overline{\boldsymbol{\theta}})$$

Here $\boldsymbol{\theta}$ denotes the total collection of variables $\{\boldsymbol{\beta}, \boldsymbol{w}, \boldsymbol{\chi}, \boldsymbol{p}, \boldsymbol{\rho}, \boldsymbol{\alpha_1}, \sigma_w, \boldsymbol{\sigma_\chi}, \boldsymbol{\sigma_{\alpha_1}}\}$. Thus the DIC can also be expressed in function of the mean posterior deviance $\overline{D(\boldsymbol{\theta})}$ and the deviance of the posterior mean $D(\overline{\boldsymbol{\theta}})$, as proposed by Spiegelhalter et al. (2002):

$$DIC = 2 \cdot \overline{D(\boldsymbol{\theta})} - D(\overline{\boldsymbol{\theta}})$$

When using mixture models, there is a general concern that the correct effective number of parameters are not properly accounted for. In order to provide for alternative measures, the Mean Squared Prediction Error (MSPE) and the Marginal Predictive Likelihood (MPL) are also applied. The additional benefit of using the latter measures is that they also evaluate the prediction performance, an aspect which the DIC overlooks (See Celeux et al. (2006)). The used definitions for MSPE and MPL are as follows:

$$MSPE = \frac{1}{N} \sum_{f,g,i,j} (TSH_{fgij} - \widehat{TSH}_{fgij})^2$$
$$MPL = \sum_{f,g,i,j} log(CPO_{fgij})$$

where N reflects the number of observations, TSH_{fgij} is the observed value and $\widehat{\text{TSH}}_{fgij}$ represents the predicted value of TSH_{fgij}, sampled from the posterior predictive distribution. The CPO_{fgij} represents the Conditional Predictive Ordinate of TSH_{fgij}, given the data without observation TSH_{fgij}. This could be interpreted as a cross-validation check for a future observation versus the actual value of the observed TSH value.

In addition to the model with the entry parameters, these measures can also be used to obtain a model with an appropriate amount of temporal components. Models with lower values for the DIC and MSPE, and higher values for the MPL are preferred.

3 Results

3.1 Exploratory Data

In this section the variables, described in Section 1.3, will be briefly discussed. In addition, the variable of interest TSH will be further explored, both in terms of distributional assumptions, temporal evolution and spatial variation.

A first remark should be given on the number of observations at each year. The numbers, which can be consulted in Table 1, tell us that the amount of observations in the year 2004 are more sparse compared to the other time points. This is due to incomplete reporting of all newborns in the database. This might introduce some additional variability for this time point, possibly adding difficulties when modeling the temporal components.

Year	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**
# Newborns	6736	13146	13301	13735	14978	11878

Table 1: *Number of Observations at every Time Point*

Figure 2 shows the histogram of the average TSH values, from which it is clear that a Normal distribution is insufficient to describe the data at hand. A solution was found by using a logarithmic transformation, as mentioned in the methodology section. Figure 3 depicts the TSH values that have undergone this transformation and shows the Gaussian curve that is needed for the model.

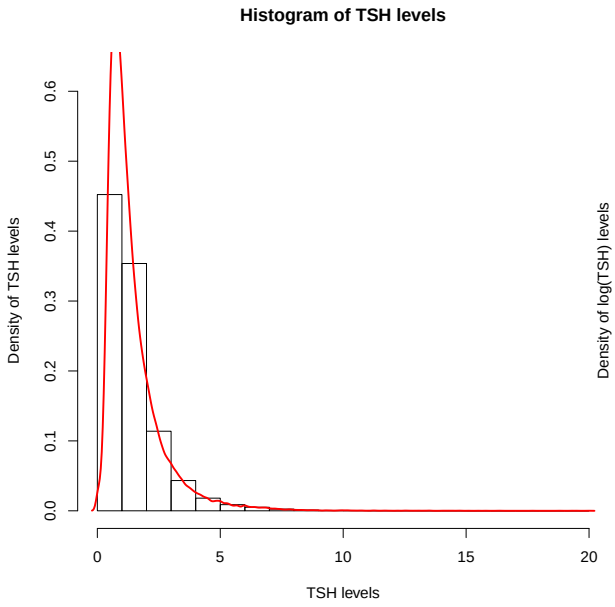


Figure 2: *Histogram of TSH Values*

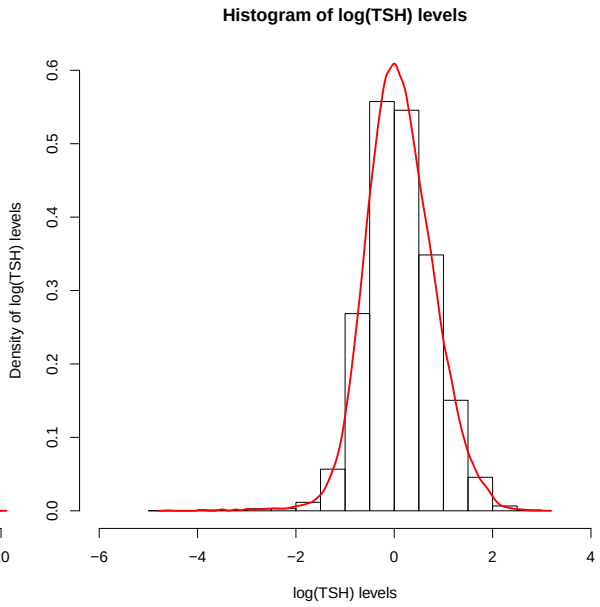


Figure 3: *Histogram of log(TSH) Values*

Information on 73774 newborn babies was used in the analysis, 51.2% of which were male whereas 48.8% were female. Furthermore, as 22.9% of the infants were fed artificially, 58.8% were breastfed and the remaining 18.3% received a combination of the former feeding types. When plotting the different combinations of Feeding Type and Gender against the averaged TSH levels at each time point (Figure 4), a couple of remarks could be made. First the average TSH levels seem to decrease over the years, though an increase is again seen in the year 2009. Second, on average the male infants seem to have higher TSH values as compared to their female counterparts. Third, babies that were exclusively breastfed exhibited a lower TSH level as compared to the other two feeding types. The difference between the artificial and the mixed approach seems a bit less obvious although the former seems to yield higher values for the response of interest in more recent years.

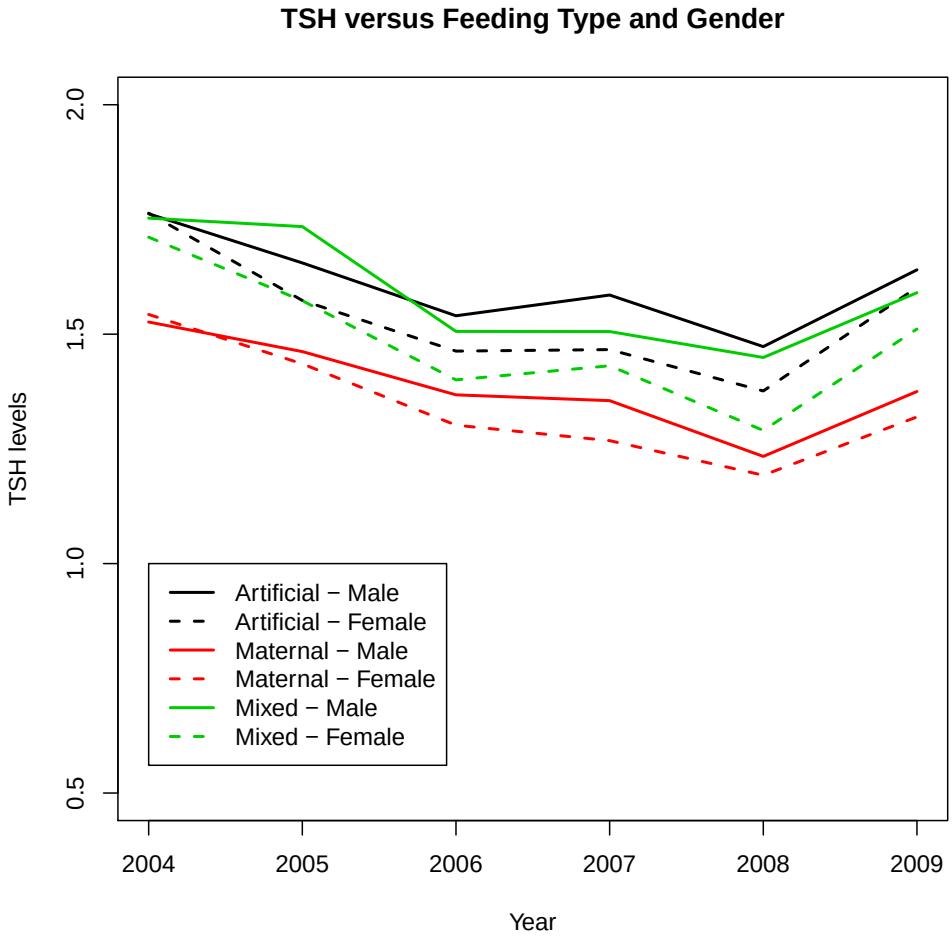


Figure 4: *Average Evolution of TSH Values over the Feeding Types per Gender*

The spatial distribution of the TSH values, averaged over all time points, was calculated and can be consulted in Figure 5. It is clear that higher TSH values are recorded in the eastern part of Galicia. The attentive reader will also notice the two municipalities which got assigned a black color. These municipalities, Negueira de Muñiz and Teixeira, contained only two observations in the original dataset. However these observations did not meet the earlier defined criteria, which explains them being absent in the reduced dataset.

When comparing Figure 5 with the plotted TSH levels of the different time points (Figure 6), something odd could be observed. Where in 2005 till 2009 the color scheme resembles the one in Figure 5, apart from the increase in non-observable areas, the pattern seems a bit more random for the year 2004. One has to keep in mind that the number of observations for the year 2004 was lower than those for the other years, causing a larger amount of variability to be

introduced.

Furthermore areal maps were provided for the different combinations of Gender and Feeding Type in Figure 7. Again the province of Lugo seems to have enclosed the majority of the high TSH-leveled municipalities. Also the information retrieved from Figure 4 could be recycled. In general males had higher TSH level compared to females across all feeding types. Also breastfed newborns produced much lower TSH values compared to the artificially and mixed feeding type for both males and females.

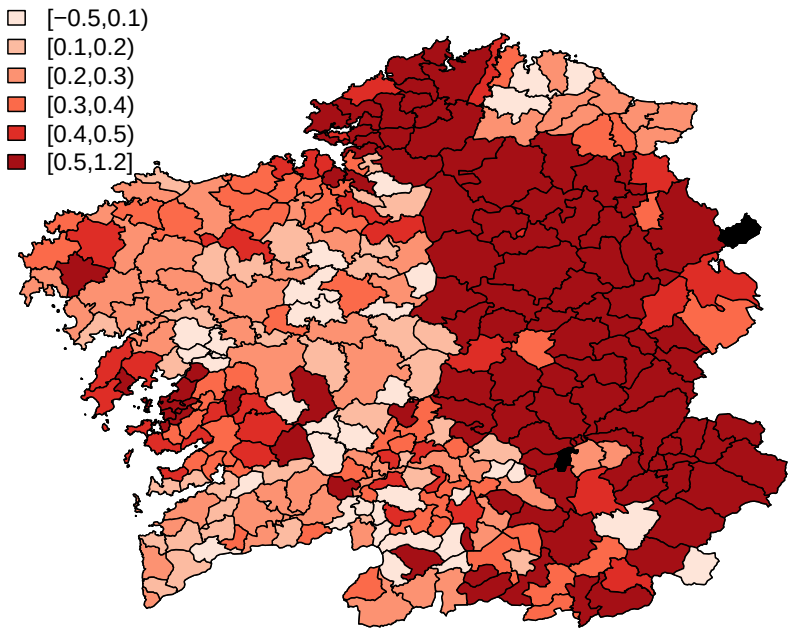
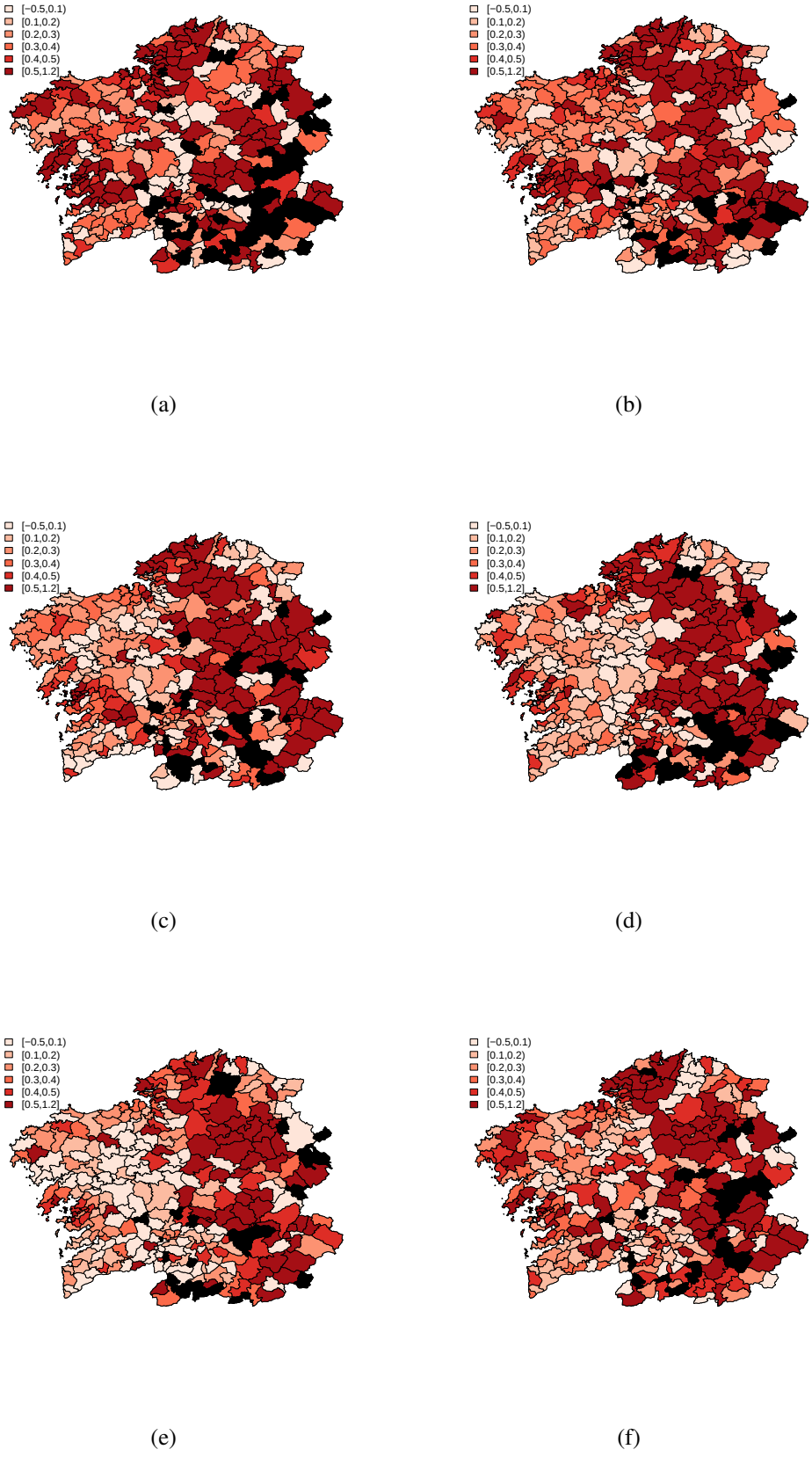


Figure 5: *Average log(TSH) Values per Municipality*

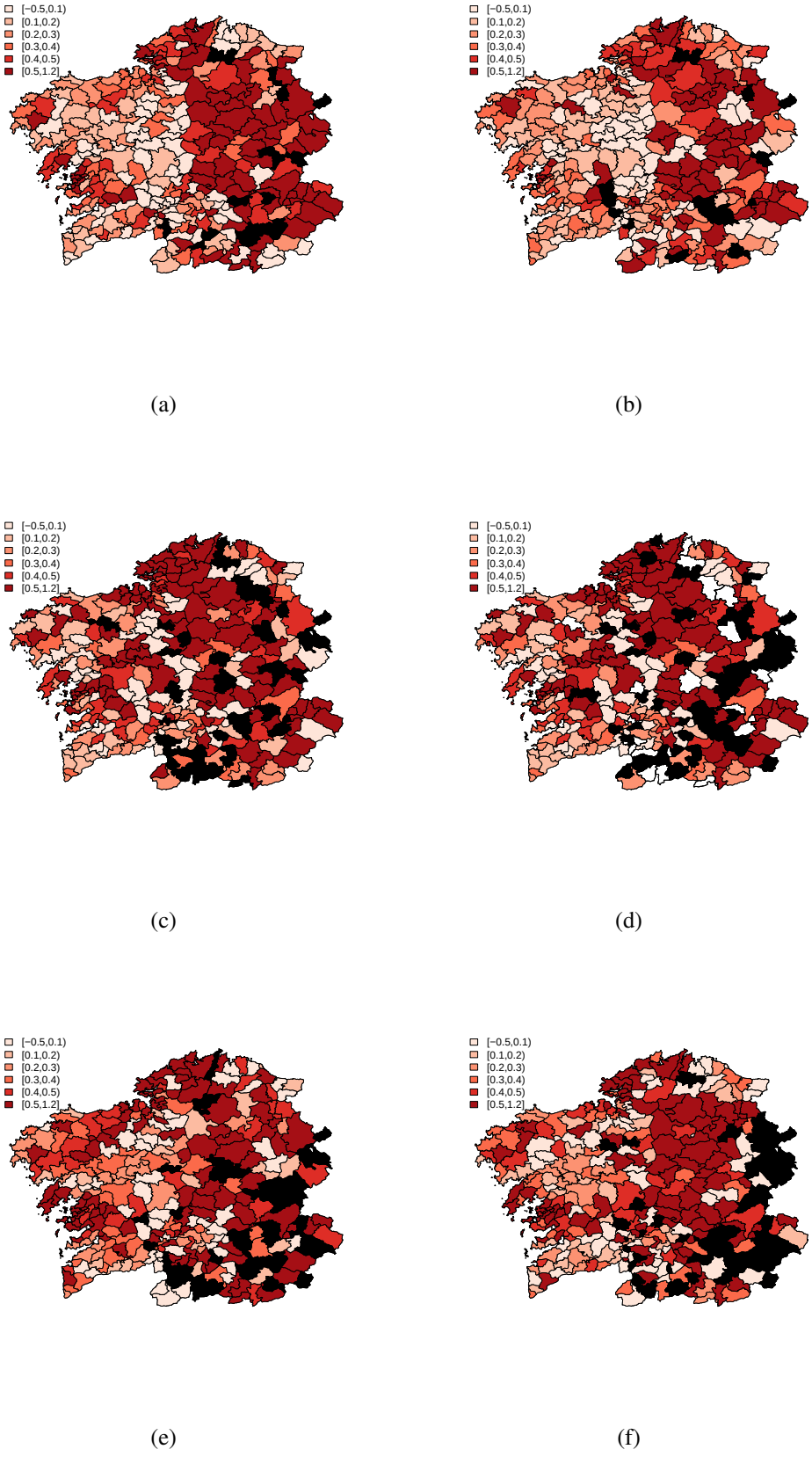


(a) (b)

(c) (d)

(e) (f)

Figure 6: *Average log(TSH) Values for (a) 2004, (b) 2005, (c) 2006, (d) 2007, (e) 2008, (f) 2009*



(a) (b)

(c) (d)

(e) (f)

Figure 7: *Average log(TSH) Values for (a) Males-Breastfed, (b) Females-Breastfed, (c) Males-Mixed, (d) Females-Mixed, (e) Males-Artificial, (f) Females-Artificial*

3.2 Analysis

In this section, we explored the models as presented in previous sections.

3.2.1 Model Choice

When fitting the model stated in the methodology with two components, it became clear that the convergence of some parameters could not be assured. The core problem lied within the unwanted synergy between the temporal components and the time-dependent intercept. As these two elements tried to explain the same thing, the convergence of these variables became problematic. A solution could be found in the form of a time-independent intercept, forcing the temporal components to clarify all the information not yet explained by the variables Gender and Feeding Type.

Another, more practical adaptation to the model could be found in the evaluation of the interaction term as it had no significant effect at any of the time points. This was already hinted by Figure 4 in the exploratory analysis. These findings were translated into the following model:

$$\mu_{fgij} = \beta_0 + \beta_{1,j} \cdot \text{Gender}_g + \beta_{2,j} \cdot \text{Feedtype}_f + \sum_{l=1}^{2} w_{il} \cdot \chi_{lj}$$

These adaptations secured convergence of the model in the non-spatial setting, however for the spatial model convergence could not be guaranteed. Thus it was opted to remove the global intercept β_0 in order to remove any interaction whatsoever between the intercept and the temporal components. By following this road, convergence problems in general were resolved, as was validated using trace plots, the MC Error and the Geweke diagnostic when in doubt. These models were run using a single chain for 50000 iterations, applying a thinning factor of five and inducing overrelaxation, in order to break the autocorrelation and improve convergence.

3.2.2 Selection Number of Components

When setting up the model, the number of components had to be identified first. This was done by including the entry parameters, in combination with the calculation of the MPL and MSPE.

The traditional DIC was also calculated to check how well it would hold up, despite being the lesser of the four methods.

		$\overline{\mathbf{D}(\theta)}$	**pD**	**DIC**	**MSPE**	**MPL**
Non-Spatial Weights	*2 Components*	16500	7811.37	24311.37	0.2127	1.193
	3 Components	16580	8033.48	24613.48	0.2103	1.199
	4 Components	16620	8147.97	24767.97	0.2090	1.201
Spatial Weights (CAR)	*2 Components*	16490	7729.34	24219.34	0.2136	1.191
	3 Components	16560	7915.34	24475.34	0.2111	1.197
	4 Components	16580	7969.32	24549.32	0.2104	1.198

Table 2: *Model Diagnostics for Spatial and Non-Spatial Models*

In Table 2, models with two, three and four components were compared. By looking at the MSPE and MPL values between the different sets of components, one readily sees that the subsequent differences for both the spatial and the non-spatial models on the weights are negligible. This implies that introducing another component does not improve the fit of the model, discouraging unwanted model complexity. Even though the MSPE and MPL criteria are preferred here, due to their predictive nature, the DIC values also confirm the general idea that the usage of more than two components is unnecessary in this setting.

This finding was backed up by the "full" model, which included the entry parameters. The model which was set up could be described as follows:

$$\mu_{fgij} = \beta_0 + \beta_{1,j} \cdot \text{Gender}_g + \beta_{2,j} \cdot \text{Feedtype}_f + \sum_{l=1}^{10} \psi_l \cdot w_{il} \cdot \chi_{lj}$$

The "full" model contains ten components, whose eligibility was assessed using the posterior means of the entry parameters. After running a single-chain of 30000 iterations, the results suggested that two components were sufficient to describe the different temporal trends.

3.2.3 Fixed Effects Results

When looking at the posterior means for the main effects, Table 3 could be consulted. After a brief inspection, one could observe that the differences between the two approaches are minimal. Furthermore the results learn us that none of the variables are significant at the first year. This does not come as a surprise, given the low number of observations at the given time point, causing an increase in variability compared to the other years. Since these variables are non-significant at some time points, one might question their overall effect. Using a model with time-independent covariates, further investigation teaches us that both Gender and Feeding Type had a significant effect of -0.0361 (-0.05795, -0.01417) and 0.0393 (0.02627, 0.0523) respectively for the non-spatial model. The estimates of the spatial model yielded similar results. Hence it was decided that these covariates were to remain in the model in order to give the interpretation of the results an additional dimension.

Translated, these numbers exemplify the fact that male newborns on average have slightly more elevated TSH values compared to their female counterparts. When looking at the feeding type of the infants, the period between 2005 and 2009 provided significant values in the range of (0.0279, 0.0598) for the non-spatial model, and (0.0346, 0.0607) for the spatial model. This empowers the belief that breastfeeding counters iodine deficiency compared to bottle feeding and the mixed approach, by the the transferral of iodine from mother to child.

		Year 1	**Year 2**	**Year 3**	**Year 4**	**Year 5**	**Year 6**
Non-Spatial	*Gender*	0.0143*	-0.0086*	-0.0522	-0.0695	-0.0761	-0.0121*
		(-0.0378, 0.0666)	(-0.0517, 0.0334)	(-0.0959, -0.0105)	(-0.1114, -0.0283)	(-0.1170, -0.0334)	(-0.0575, 0.0354)
	Feeding Type	0.0195*	0.0279	0.0345	0.0598	0.0339	0.0585
		(-0.0129, 0.0511)	(0.0002, 0.0556)	(0.0065, 0.0620)	(0.0323, 0.0875)	(0.0066, 0.0616)	(0.0292, 0.0885)
Spatial	*Gender*	0.0140*	-0.0085*	-0.0528	-0.0725	-0.0803	-0.0180*
		(-0.0382, 0.0662)	(-0.0504, 0.0331)	(-0.0952, -0.0118)	(-0.1129, -0.0321)	(-0.1207, -0.0387)	(-0.0618, 0.0292)
	Feeding Type	0.0215*	0.0301	0.0359	0.0607	0.0346	0.0572
		(-0.0110, 0.0536)	(0.0022, 0.0573)	(0.0086, 0.0635)	(0.0334, 0.0880)	(0.0076, 0.0617)	(0.0285, 0.0868)

* *Main effect is not significant at the given time point*

Table 3: *Posterior Means for given Main Effects and their respective Credible Intervals*

3.2.4 Temporal Components

The next step is to explore the two components which originated from the model, starting with the temporal aspect. The fitted results for both the non-spatial and spatial model, along with their pointwise 95% credible intervals, were displayed in Figure 8. The temporal evolutions seem to be similar for the two models: one component which stabilizes around zero as the other one starts a bit higher, but decreases over time. This indicates that the regions that belong to the second component had a high rate of iodine deficiency in the beginning of the study, which reduced steadily in more recent years.

The attentive observer might notice two additional things: the variability at the first year is slightly larger compared to the other time points across all components, and the overall variability is smaller for the spatial model than for the non-spatial model. It is already known that the former is due to the lower amount of information available for that time point. The latter observation on the other hand could be explained by the nature of the models. As the non-spatial model neglects the modeling of the correlation between neighboring areas, this induces an inflation of the posterior standard deviation. The CAR model on the other hand does take this relationship into account, which is translated in this narrow credible interval.

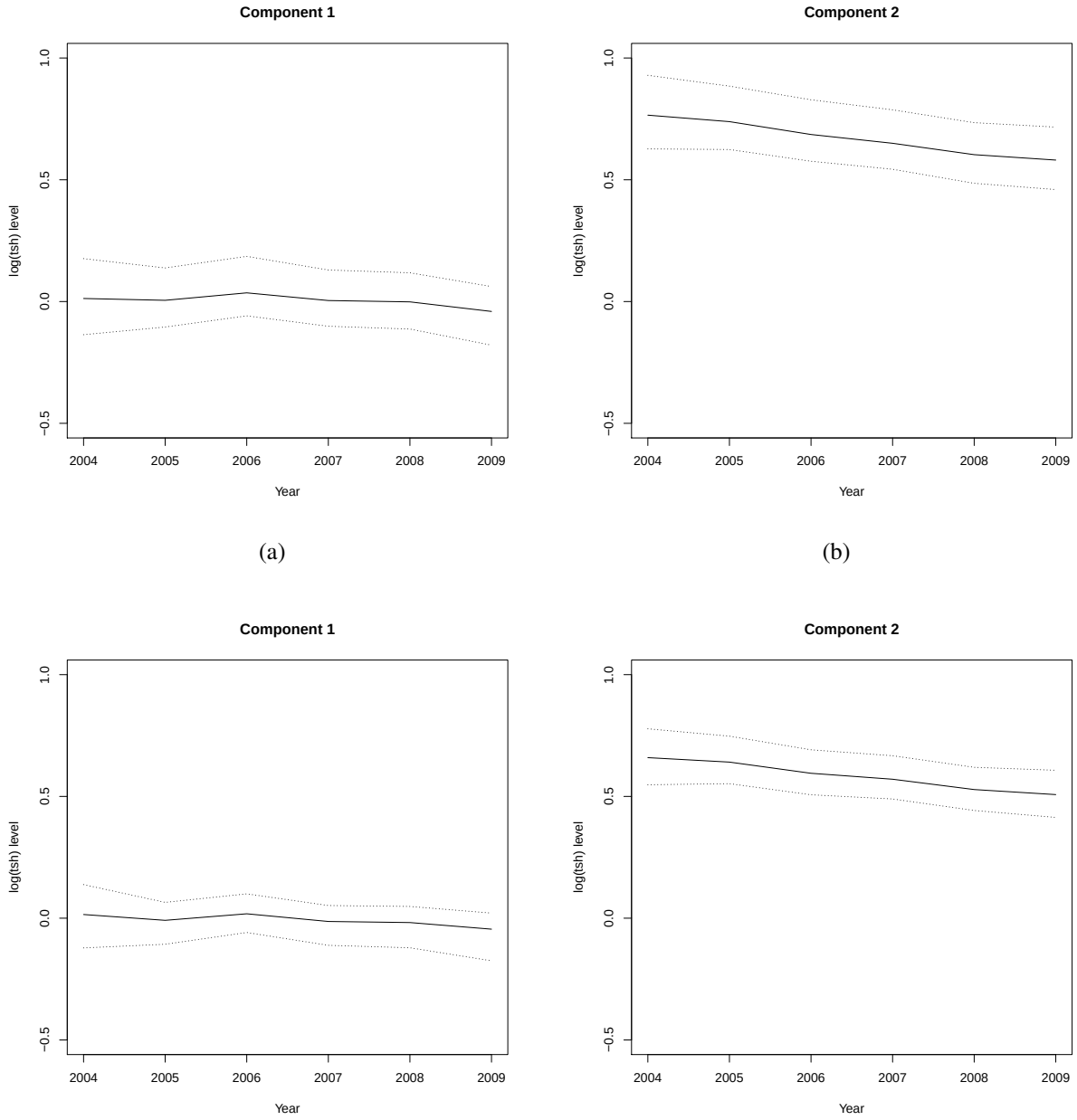
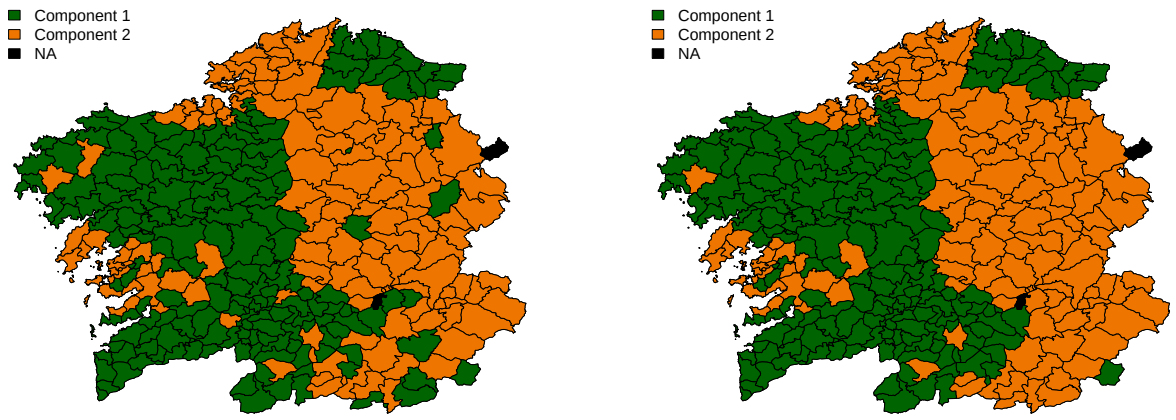


Figure 8: *Posterior Expected Temporal Effects for (a)-(b) Non-Spatial Weights and (c)-(d) Spatial Weights*

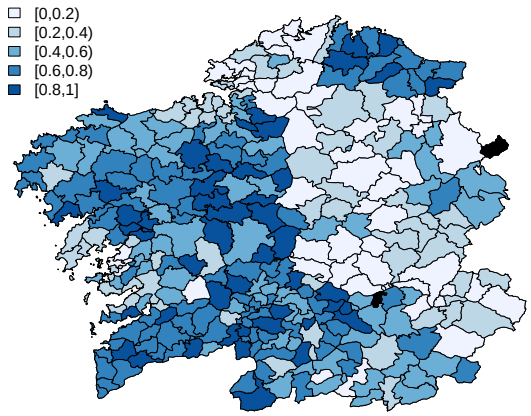
3.2.5 Spatial Weights

One also ought to take a look at the posterior spatial weights. The maps, providing information on the allocation, are presented in Figure 9 for the non-spatial and spatial model. The maps depicting the estimated values for the fitted weights could be consulted in Figure 10. The municipalities which showed larger TSH values are apparently inclined to be situated in the second component, which is a logical result since this component produces the higher TSH values of the two. Therefore these regions, which are depicted in orange, could be considered iodine deficient. Furthermore these maps nicely show the spatial smoothing caused by the CAR model. The plots of the spatial weights also display the symbiosis between the magnitude of the TSH values and the posterior weights: the higher the TSH level for a particular municipality, the higher the corresponding weight for the second temporal component will be (and vice versa).

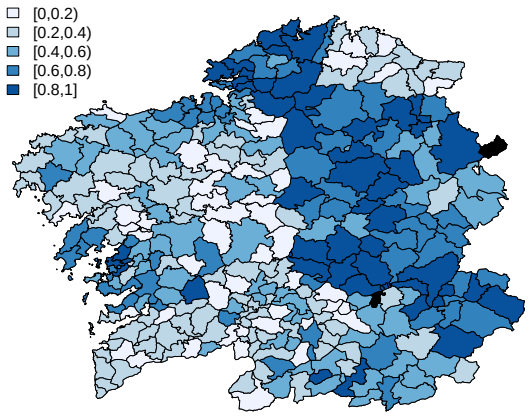


(a) (b)

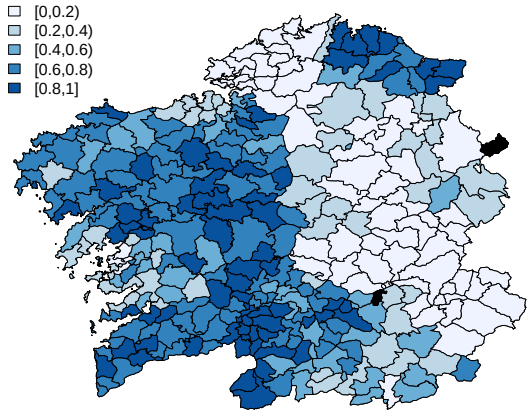
Figure 9: *Maximum Posterior Expected Weight Map for (a) Non-Spatial Weights and (b) Spatial Weights*



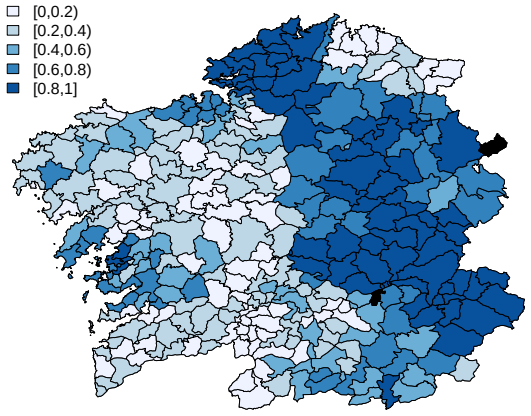
(a)



(b)



(c)



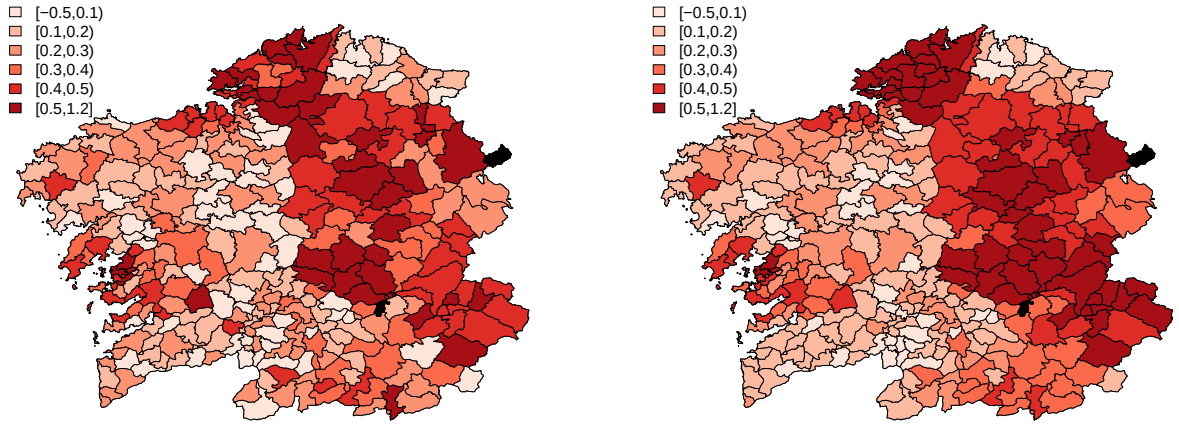
(d)

Figure 10: *Posterior Expected Weight Maps for (a)-(b) Component 1 and 2 of the Non-Spatial Model and (c)-(d) Component 1 and 2 of the Spatial Model*

3.2.6 Model Fitting

Lastly, an important feature of any appropriate model is the ability to fit the data in an adequate manner, which will be investigated now. Figure 11 provide us with the fitted posterior means, averaged over the six time points, of the non-spatial and spatial model respectively. While the non-spatial model only aims at approximating the data, the spatial model also considers information from neighboring municipalities, taking into account the possible correlation between municipalities. The spatial smoothing which results from this methods can be easily observed in the corresponding areal map.

For both these models the eastern cluster, containing a lot of municipalities with relatively high TSH levels, immediately catches the attention. Geographically, the utmost southeasterly part of this cluster coincides with the western part of the Cantabrian Mountains, a mountainous area which stretches across the north of Spain. The central and northwestern part is surrounded by a series of mountain ranges. These areas are thus secluded from the sea, making it more difficult to obtain seafood, a prominent source of sea salt and thus iodine. This lack of iodine triggers the thyroid to work harder, by increasing the production of TSH and thus elevating the overall TSH level for those municipalities.



(a) (b)

Figure 11: *Posterior Fitted Means for (a) Non-Spatial Weights and (b) Spatial Weights*

3.3 Sensitivity

In order to check the consistency of the results, a handful of models with small changes were applied to the dataset as well. One of them is already discussed briefly, namely the model with time-independent covariates. The second model is defined by an autoregressive dependence structure of order two for the temporal components, allowing for a more flexible assumption:

$$\chi_{lj} \sim N(\rho_{1l} \cdot \chi_{l,j-1} + \rho_{2l} \cdot \chi_{l,j-2}, \sigma^2_{\chi_l})$$

Finally the last model assumes a linear mean structure for the temporal components, allowing for a different evolution within each component. Their model diagnostics can be found in Table 4 and were fit solely for the non-spatial model.

For the spatial models, the sensitivity of the neighborhood structure was evaluated. Up until now, only the Queen's Case was looked into, however a variety of neighborhood structures could be applied. Instead of applying weights equal to 1, the inverse distances between the cen-

troids of the municipalities could be used as weights for the neighborhood structure. This was done using the nearest neighbors method, where only the association between municipalities whose centroids lie within a certain prespecified distance is evaluated. In this situation, larger weights are granted for areas which lie close within the prespecified range. This user-defined distance was chosen such that the municipality with the largest amount of neighbors in the Queen's Case still retained all its "neighbors" within the distance band. This also guaranteed that each area was connected with at least one other area, producing no "islands".

Even though the differences in MPL and MSPE between the models are minimal, one can notice that the better model comes in the form where Gender and Feeding Type are time-independent, indicating that the proposed model could be simplified. Globally, looking at the diagnostic features and the recovery of the components (not shown here), the deviations from the proposed model are kept to a minimum, inferring that the model is robust against small alterations.

		$\overline{\mathbf{D}(\theta)}$	**pD**	**DIC**	**MSPE**	**MPL**
Non-Spatial	*Time-Independent Covariates*	16490	7784.80	24274.80	0.2130	1.192
	AR(2) Structure	16520	7947.37	24467.37	0.2123	1.194
	Linear Temporal Structure	16500	7806.56	24306.56	0.2128	1.193
Spatial	*Inverse Distance*	16490	7740.33	24230.33	0.2130	1.192

Table 4: *Model Diagnostics for Sensitivity Analysis*

4 Discussion and Conclusion

In this paper the TSH values of 315 municipalities in Galicia, a region in the north-west of Spain, were examined over a period of six years as an indicator for the widespread problem of iodine deficiency. When investigating the possible spatial and temporal effects, a space-time latent component model was constructed in order to catch these evolutions on the logarithmic scale of the TSH response. It ought to be stressed that this model takes into account spatial heterogeneity, meaning that not a global trend was sought after but rather several local trends. In addition to the mixture component, our term of interest, the predictor variables Gender and

Feeding Type were also looked into.

Componentwise, only two were needed to explain the spatio-temporal heterogeneity which was apparent in this study. This was confirmed by means of two approaches: a "full" model containing entry parameters and several model diagnostic features. It was found that the inclusion of another component would not be deemed necessary.

The investigation of the two temporal components elucidated that the major difference lies in the magnitude of the areal TSH levels. The discrepancy between the two components is larger for the non-spatial model as compared to the spatial model. The reasoning behind this could be found in the fact that the information exchange between neighboring areas for the spatial model causes the more extreme values to be moderated downwards, reducing the dissimilarities between areas. Another fact which should be highlighted is the use of the intercept term in the non-spatial model, whereas this was not needed in the CAR model. The moderation of the spatial model could provide an explanation for this phenomenon. Given that the value for the intercept was -0.0584 for the non-spatial model, the overall trend of the components would be situated lower if the intercept would be left out. The function of the intercept in the non-spatial could be questioned in general as it only sets the first component as a reference around zero for the period of six years. Also one could remark that the use of a time-independent intercept is only viable here since the temporal trends are relatively linear and stable to begin with. If this were not the case, the model could be severely underspecified, creating the possibility that it would miss some critical information.

The weight maps reflect the findings and suspicions of the temporal components. The weights of those municipalities which exhibit a large TSH value are more elevated for the second component as would be expected, since this component portraits relative high TSH levels overall. In this project it was assumed that these weights have spatial dependence only, whereas the temporal components only had an AR(1) dependence. This assumption added strength to the identification of these components, which could be observed by their narrow credible intervals. Relinquishing this assumption might pose troubles in the interpretation and identification of the temporal components.

When specifying the prior distribution of the spatial weights, one might specify a Multivariate

model in order to capture the possible correlation between the spatial weights of different components. One possible approach is the Multivariate Intrinsic Conditional Autoregressive Model, which assumes the following notation for the mean weight values, as proposed by Gelfand et al. (2003):

$$\alpha_{1il}|\alpha_{1i'l,i\neq i'} \sim N\left(\frac{1}{n_i}\sum_{i\neq i'} B_{ii'}\alpha_{1i'l}, \frac{1}{n_i}\boldsymbol{\Sigma_{\alpha_1}}\right)$$

where $\boldsymbol{\Sigma_{\alpha_1}}$ is the $L \times L$ positive definite variance-covariance matrix between the weights from different components. For this variance structure a Wishart prior distribution was assigned: $Wishart(R^{-1}, r)$, where R is a 315×315 - dimensional positive definite matrix and r represents the number of latent variables. The specification of this model would be referred to as MCAR($\boldsymbol{\Sigma_{\alpha_1}}$). When applying this prior distribution to the unnormalized weights, it was found that the convergence of the temporal components could not be guaranteed. A separate CAR model for each of the components provides us with an alternative, devoid of overspecification. When building the model, one could also take into account random effects for Gender and Feeding Type. However this was not investigated here due to an increase in computational time and a lack of need. Even though the addition of the random effects might give a better fit, the use of only the global time-dependent covariates provide us with a straightforward, biologically plausible interpretation. The use of a random intercept term for each municipality on the other hand might tamper with the (number of) temporal components as these random effects might borrow their information, possibly disrupting the interpretation of the spatial weights.

Also one might consider adding additional covariates to the model, such as gestational age, the time between birth and the heel prick, birth weight, the time from the moment of birth till the newborn was fed for the first time, the iodine status of the mother, etc. In the present setting these characteristics were not considered due to the fact that their inclusion prolonged the running time of the used programs up to the point that the deadline for this paper would have been jeopardized. However in future studies, it is recommended that these variables should be investigated as well.

Lastly, the use of the Normal distribution for the response variable ought to be addressed. The reader should convince himself that this is of course not the sanctified road which needs to be taken. For example a Poisson approach would also be usable, by counting the number of

newborns with high TSH levels according to a prespecified benchmark. However one should bear in mind that the amount of information is highly dependent on this cut-of value, posing a great responsibility on the determination of this value. Also the question “What is a high TSH value?” is not that straightforward as these values tend to be higher for newborns than for grown-ups. Remember also that “newborn” in this setting is a subjective concept as some babies were measured weeks, even months after their birth, causing a shift in their metabolism as compared to babies who were merely a couple of days old. Furthermore there is a general loss of information, which is present in every binarization process, making this approach less appealing.

5 References

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

Celeux G., Forbes F., Robert C., Titterington M. (2006). Deviance information criteria for missing data models. *Bayesian Analysis*. 1, 515-533.

Choi J., Lawson A.B., Cai B., Hossain M.M. (2011). Evaluation of Bayesian spatiotemporal latent models in small area health data. *Environmetrics*. 22 (8), 1008-1022.

Gelfand A., Vounatsou P. (2003). Proper multivariate conditional autoregressive models for spatial data. *Biostatistics*. 4, 11-25.

Irizarry L. (2014). Thyroid Hormone Toxicity. *Medscape*.
Available: http://emedicine.medscape.com/article/819692-overview. Last accessed 26/07/2014.

Lawson A.B., Song H.R., Cai B., Hossain M.M., Huang K. (2010). Space-time latent component modeling of geo-referenced health data. *Statistics in Medicine*. 29 (19), 2012-2027.

Lope V., Pollán M., Pérez-Gómez B., Aragonés N., Ramis R., Gómez-Barroso D., López-Abente G. (2006). Municipal mortality due to thyroid cancer in Spain. *BMC Public Health*. 6, 302.

Mutambanengwe C.K., Faes C., Aerts M., Gude F., Cadarso-Suárez C. (2013). On the Spatial Heterogeneity of Quantiles of TSH in Newborns for the Detection of Iodine Deficient Areas. *Spatial and Spatio-temporal Epidemiology*.

Roth E. (2013). TSH Hormone Levels in Infants. *Livestrong*.
Available: http://www.livestrong.com/article/230115-tsh-hormone-levels-in-infants. Last accessed 28/07/2014.

Sinha S. (2012). Pediatric Hypothyroidism Workup. *Medscape*.
Available: http://reference.medscape.com/article/922777-workup. Last accessed 28/07/2014.

Spiegelhalter D., Best N., Carlin B., Van Der Linde A. (2002). Bayesian deviance, the effective number of parameters and the comparison of arbitrarily complex models. *Journal of the Royal Statistical Society: Series B*. 64, 583-640.

Thureen S., (2011). The Normal Range for TSH Levels. *Answers*.
Available: http://thyroid.answers.com/thyroid-stimulating-hormone-tsh/what-is-the-normal-range-for-tsh-levels. Last accessed 28/07/2014.

Vitti P., Rago T., Aghini-Lombardi F., Pinchera A. Iodine deficiency disorders in Europe. *Public Health Nutrition* 4 (2B): 529-35.

World Health Organization (WHO). (2004). Iodine status worldwide: WHO Global Database on Iodine Deficiency. Geneva, Switzerland.

Auteursrechtelijke overeenkomst

Ik/wij verlenen het wereldwijde auteursrecht voor de ingediende eindverhandeling:
Space-time latent component modeling of TSH in newborns

Richting: **Master of Statistics-Biostatistics**
Jaar: **2014**

in alle mogelijke mediaformaten, - bestaande en in de toekomst te ontwikkelen - , aan de Universiteit Hasselt.

Niet tegenstaand deze toekenning van het auteursrecht aan de Universiteit Hasselt behoud ik als auteur het recht om de eindverhandeling, - in zijn geheel of gedeeltelijk -, vrij te reproduceren, (her)publiceren of distribueren zonder de toelating te moeten verkrijgen van de Universiteit Hasselt.

Ik bevestig dat de eindverhandeling mijn origineel werk is, en dat ik het recht heb om de rechten te verlenen die in deze overeenkomst worden beschreven. Ik verklaar tevens dat de eindverhandeling, naar mijn weten, het auteursrecht van anderen niet overtreedt.

Ik verklaar tevens dat ik voor het materiaal in de eindverhandeling dat beschermd wordt door het auteursrecht, de nodige toelatingen heb verkregen zodat ik deze ook aan de Universiteit Hasselt kan overdragen en dat dit duidelijk in de tekst en inhoud van de eindverhandeling werd genotificeerd.

Universiteit Hasselt zal mij als auteur(s) van de eindverhandeling identificeren en zal geen wijzigingen aanbrengen aan de eindverhandeling, uitgezonderd deze toegelaten door deze overeenkomst.

Voor akkoord,

Watjou, Kevin

Datum: **10/09/2014**