Comparison of different Bayesian spatio-temporal models using R packages

Authors: Aswi Aswi[1], Kerrie Mengersen[1], Susanna Cramb[1], Wenbiao Hu[2], Gentry White[3]
[1] ARC Centre of Excellence for Mathematical and Statistical Frontiers, Queensland University of Technology, Australia; [2] School of Public Health and Social Work, Queensland University of Technology, Australia; [3]School of Mathematical Science, Queensland University of Technology, Australia

Introduction

A number of R packages are available for implementing spatio-temporal Bayesian modelling such as spTimer, R-INLA, CARBayesST and spTDyn. The main advantage of the Bayesian approach for modelling spatio-temporal structures resides in its taking into account uncertainty in the estimates or predictions, and its ease in specifying spatial and temporal structure in prior distributions. Here, we will focus on two R packages, namely R-INLA (Integrated Nested Laplace Approximation)[1] and CARBayesST[2] and apply these packages through a case study of dengue fever in Makassar, Indonesia. The number of dengue cases in Makassar fluctuates between locations and years. Since dengue cases vary in location and time, the spatial and temporal component must be taken into consideration. Here, we investigate Bayesian spatial and spatio-temporal model specification.

Methods

Annual dengue fever incidence data for Makassar, Indonesia (14 geographic areas) during 2002-2015 were obtained from the City Health Department of Makassar, South Sulawesi Province.
The following range of Bayesian model specifications were considered. All models used a Poisson distribution for the count data.

Model-1: Spatio-temporal CAR linear[3]

$$y_{ij} \sim \text{Poisson}(e_{ij}\theta_{ij})$$
$$log(\theta_{ij}) = \alpha + \alpha_1 + u_i + (\beta + \delta_i)\frac{j-\bar{j}}{N}$$

$(\alpha_1 + u_i)$ and $(\beta + \delta_i)$ are spatially varying intercept and slope respectively

Model-2: Spatio-temporal CAR ANOVA[4]

$log(\theta_{ij}) = \alpha + u_i + \delta_j + \gamma_{ij}$; u_i is the overall spatial effect to all time, (δ_j) is overall temporal trend random effect, γ_{ij} is Independent space-time interactions

Model-3: Spatio-temporal CAR separate spatial [5]

$log(\theta_{ij}) = \alpha + u_{ij} + \delta_j$; u_{ij} is separate spatial for each time period

Model-4: Spatio-temporal CAR AR[6]

$log(\theta_{ij}) = \alpha + u_{ij}$

Model-5: Spatio-temporal CAR adaptive[7]

Overall model structure is the same as for ST CAR AR model, but is more suitable when the residual spatial autocorrelation in the response is consistent over time but has a localised structure.

Model-6: Spatio-temporal CAR localised[8]

$log(\theta_{ij}) = \alpha + u_{ij} + \lambda_{Z_{ij}}$; $\lambda_{Z_{ij}}$ is piecewise constant clustering component

Model-7: Linear temporal trend model [3]

$log(\theta_{ij}) = \alpha + u_i + v_i + (\beta + \delta_i)j$; β is the main linear trend, δ_i is the difference between the global trend β and the area-specific trend

Model-8: Nonparametric dynamic trend(RW1) model [4]

$log(\theta_{ij}) = \alpha + u_i + v_i + \gamma_j + \phi_j$; γ_j and ϕ_j are temporally structured random effects using a random walk of order 1 (RW1) and unstructured random effect respectively

Model-9: Nonparametric dynamic trend(RW2) model

Model is the same as Model 8, except temporally structured random effect γ_j uses a random walk of order 2 (RW2).

Model-10 : Nonparametric dynamic trend(AR1) model

Model is the same as Model 8, except temporally structured random effect γ_j uses a first order of autoregressive (AR1).

Model-11 : Nonparametric dynamic trend (AR2) model

Model is the same as Model 8, except temporally structured random effect γ_j using a second order of autoregressive (AR2).

Models were run using R-INLA and CARBayes packages and compared using goodness-of-fit measures, such as Deviance Information Criterion (DIC) and Watanabe-Akaike Information Criterion (WAIC), as well as comparing the obtained estimates and their precision for each area. For models run under CARBayes package, we used burnin=20000 and n.sample=120000.

Results

Table 1. DIC, WAIC and Time used for every model

Model	DIC	WAIC	Time (seconds)
Model – 1	2012.91	2230.25	44.40
Model – 2	9374.02	45374.63	54.70
Model – 3	9499.17	Inf	71.80
Model – 4	1632.36	1884.21	33.00
Model – 5	1923.21	2111.74	162.90
Model – 6: G=2	**1367.07**	1927.39	117.10
Model – 6: G=3	1438.99	1892.07	128.80
Model – 7	2075.85	2164.67	4.10
Model – 8	1526.40	**1614.61**	4.27
Model – 9	1526.63	1615.17	8.23
Model – 10	1526.31	1614.75	4.51
Model – 11	1526.35	1614.93	5.12

The spatio-temporal CAR localised model with G=2 had substantially better model fit (has the lowest DIC). However, Spatio-temporal CAR separate spatial model had the highest DIC.

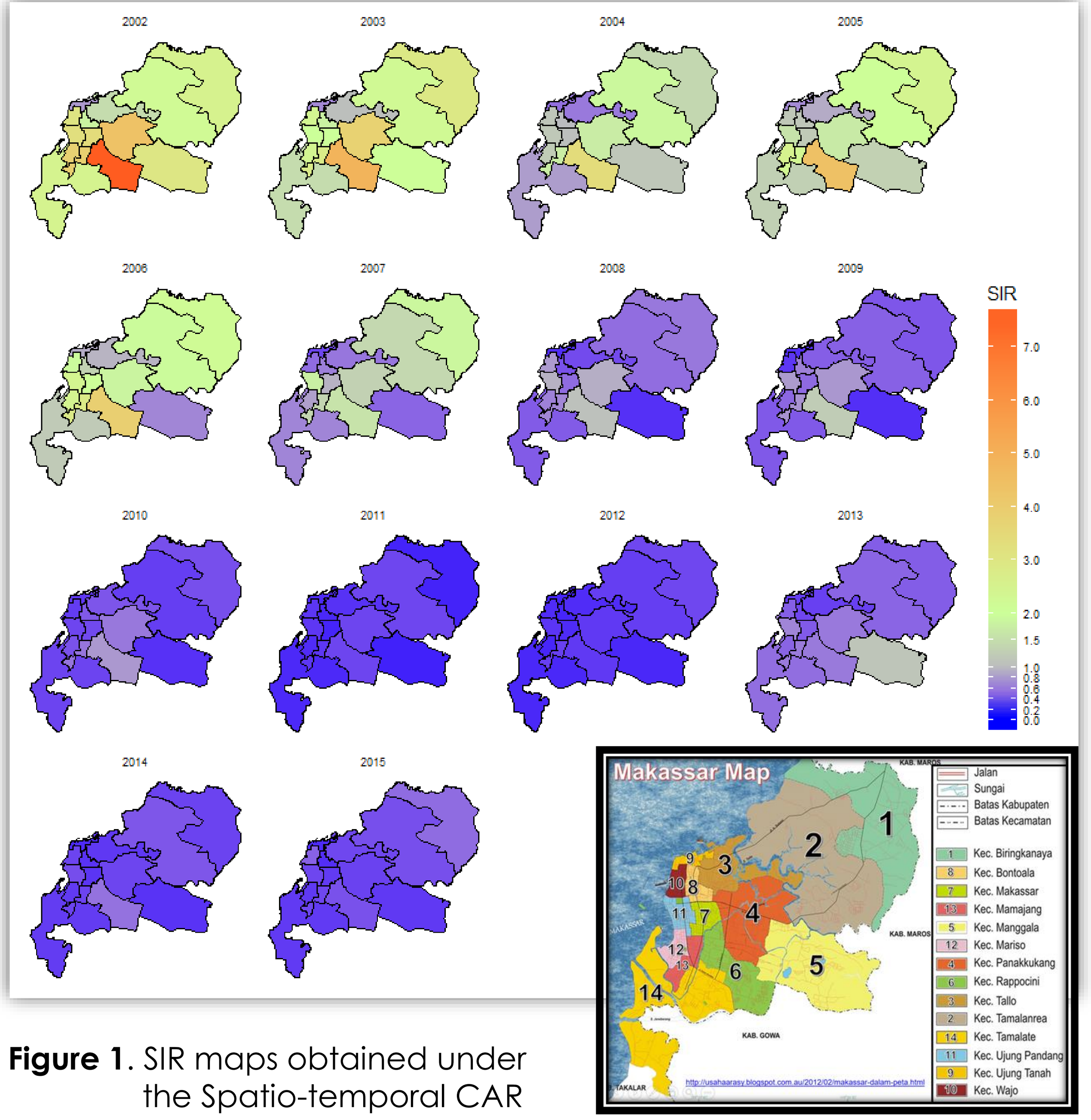

Figure 1. SIR maps obtained under the Spatio-temporal CAR localised with G=2

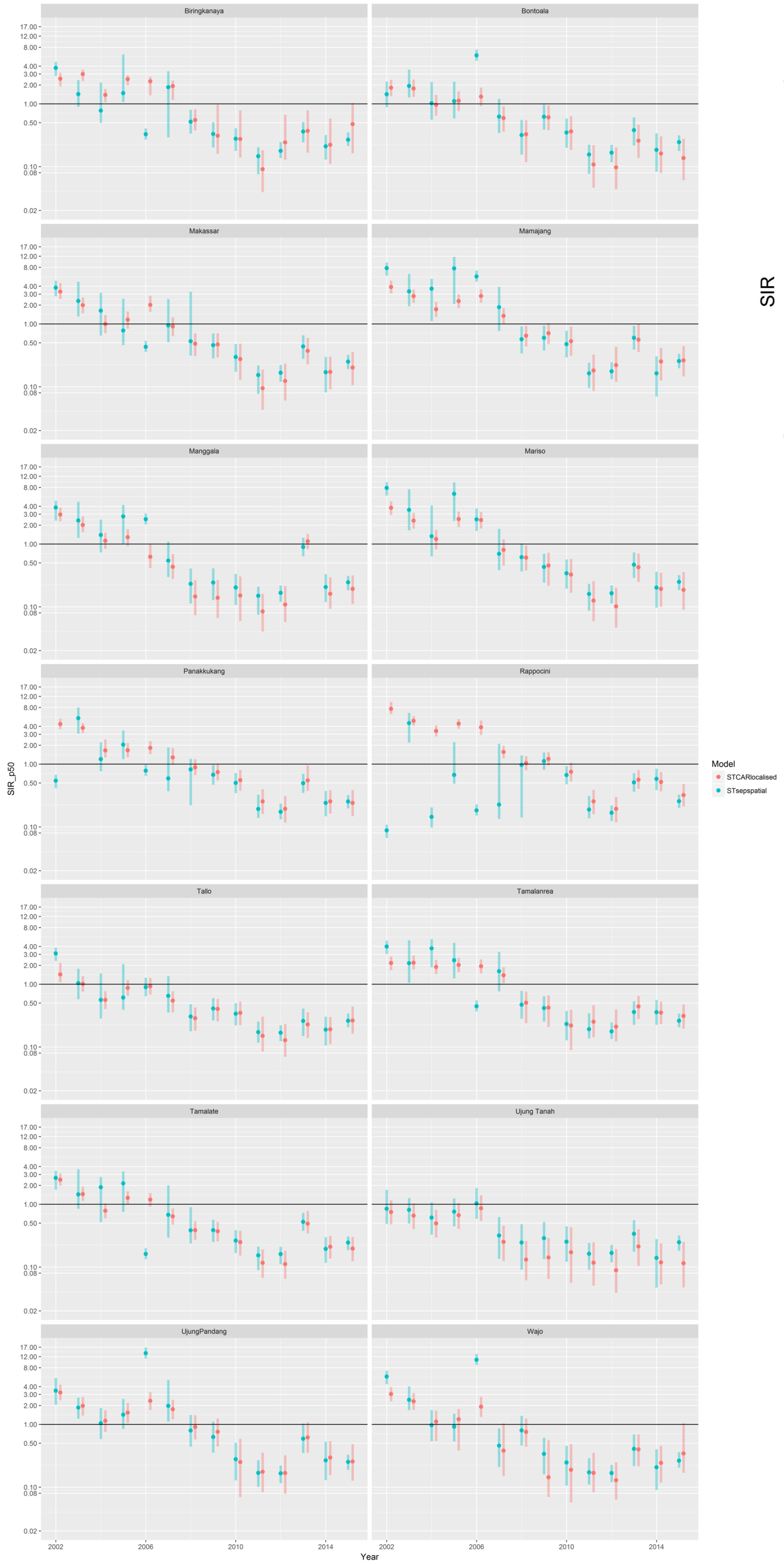

Figure 2. SIR plot under ST CAR localised (the lowest DIC) and ST separate spatial (the highest DIC) for every area as well as 95%CIs

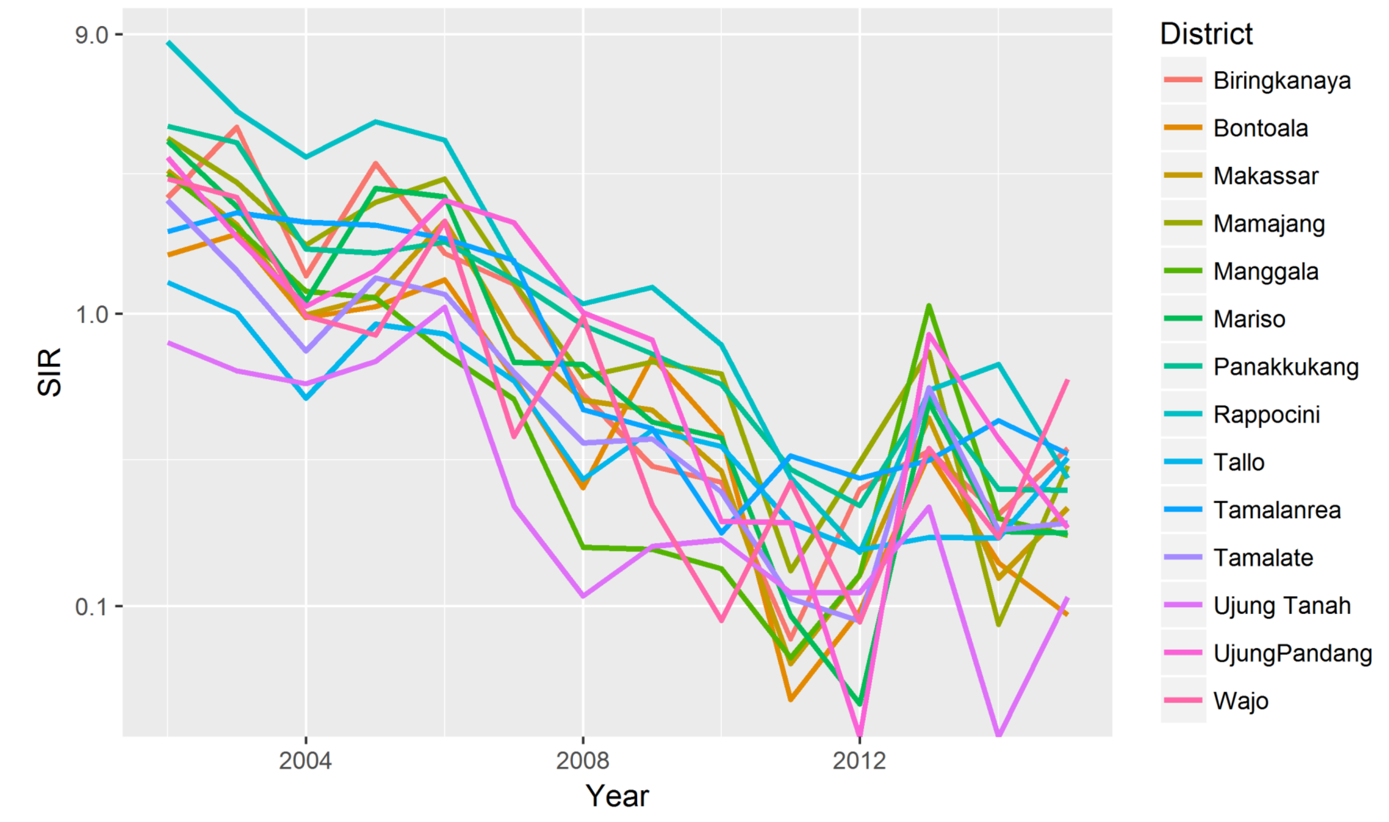

Figure 3. Raw SIR plot

Conclusions

Spatio-temporal CAR localised with G=2 model had a much better fit than other models. The computational speed and ease of using these packages makes them a very attractive option for Bayesian spatio-temporal modelling.

Acknowledgements

This work was supported by the ARC Centre of Excellence for Mathematical and Statistical Frontiers and the Statistical Society of Australia (SSA).

References

[1] Blangiardo M, Cameletti M. Spatial and spatio-temporal Bayesian models with R-INLA: John Wiley & Sons; 2015.

[2] Lee D, Rushworth A,Naoier G (2016). CARBayesST version 2.2: An R Package for Spatio-temporal Areal Unit Modelling with Conditional Autoregressive Priors, URL https://www.researchgate.net/profile/Duncan_Lee/publication/295906974

[3] Bernardinelli L, Clayton D, Pascutto C, Montomoli C, Ghislandi M, Songini M (1995). Bayesian Analysis of Space-Time Variation in Disease Risk. Statistics in Medicine, 14,2433–2443.

[4] Knorr-Held L (2000). Bayesian Modelling of Inseparable Space-Time Variation in Disease Risk. Statistics in Medicine, 19, 2555–2567.

[5] Napier G, Lee D, Robertson C, Lawson A, Pollock K (2016). A Model to Estimate the Impact of Changes in MMR Vaccination Uptake on Inequalities in Measles Susceptibility in Scotland. Statistical Methods in Medical Research, 25, 1185–1200.

[6] Rushworth A, Lee D, Mitchell R (2014). A Spatio-Temporal Model for Estimating the Long-Term Effects of Air Pollution on Respiratory Hospital Admissions in Greater London. Spatial and Spatio-temporal Epidemiology, 10, 29–38.

[7] Rushworth A, Lee D, Sarran C (2017). An Adaptive Spatio-Temporal Smoothing Model for Estimating Trends and Step Changes in Disease Risk. Journal of the Royal Statistical Society C, 66, 141–157.

[8] Lee D, Lawson A (2016). Quantifying the Spatial Inequality and Temporal Trends in Maternal Smoking Rates in Glasgow. Annals of Applied Statistics, 10, 1427–1446.