Bayesian spatio-temporal estimation of dengue fever: when areas are few

Authors: Aswi Aswi[1], Susanna Cramb[1], Kerrie Mengersen[1]

[1] ARC Centre of Excellence for Mathematical and Statistical Frontiers, Queensland University of Technology, Australia

Introduction

- Spatio-temporal modelling when there are few (<20) small areas can be challenging. Bayesian methods can be beneficial in this situation due to the ease of specifying structure and additional information through priors. However, care is needed as there are often fewer neighbours and more edges, which may influence results.
- Dengue fever is still a serious health problem in several countries including Indonesia. Makassar, Indonesia had 6882 new cases of dengue registered from 2002 to 2015, and there have been fluctuations over time. Previous Bayesian spatio-temporal modelling approaches for dengue fever when there were few areas included incorporating a linear temporal trend component.[1] Here we investigate Bayesian spatial and spatio-temporal model specification when there are few areas.

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.
- A range of Bayesian model specifications were considered, including the following:

Model-1: Independent model with only an unstructured (iid) random effect (v_i)

$$y_i \sim \text{Poisson}(e_i\theta_i)$$
$$log(\theta_i) = \alpha + v_i$$

Model-2: BYM model [2] with both an intrinsic conditional autoregressive (CAR) prior on spatially structured (u_i) and iid random effects (v_i)

$$y_i \sim \text{Poisson}(e_i\theta_i)$$
$$log(\theta_i) = \alpha + u_i + v_i$$

Model-3: Linear temporal trend model [3]

$$y_{it} \sim \text{Poisson}(e_{it}\theta_{it})$$
$$log(\theta_{it}) = \alpha + u_i + v_i + (\beta + \delta_i)t$$

β is the main linear trend

δ_i is the difference between the global trend β and the area-specific trend

$$\delta_i \sim \text{Normal}(0, 1/\tau_\delta)$$

Model-4: Nonparametric dynamic trend model [4]

$$y_{it} \sim \text{Poisson}(e_{it}\theta_{it})$$
$$log(\theta_{it}) = \alpha + u_i + v_i + \gamma_t + \phi_t$$

γ_t is a temporally structured effect using a random walk of order 2

and ϕ_t an unstructured temporal effect

- Models were run using R-INLA and compared using goodness-of-fit measures, such as Deviance Information Criterion (DIC) and Conditional Predictive Ordinate (CPO), as well as comparing the obtained estimates and their precision for each area.

Results

Table 1. DIC and CPO for four models

Model	DIC	-mean(log(CPO))
Model – 1	7101.03	18.25931
Model – 2	7101.01	18.25911
Model – 3	2075.85	5.532935
Model – 4	1526.63	4.149736

- The nonparametric dynamic trend spatio-temporal model had substantially better model fit.
- Using adjacent boundaries to define neighbours, the maximum number of neighbours was 6 and the minimum number was 1.

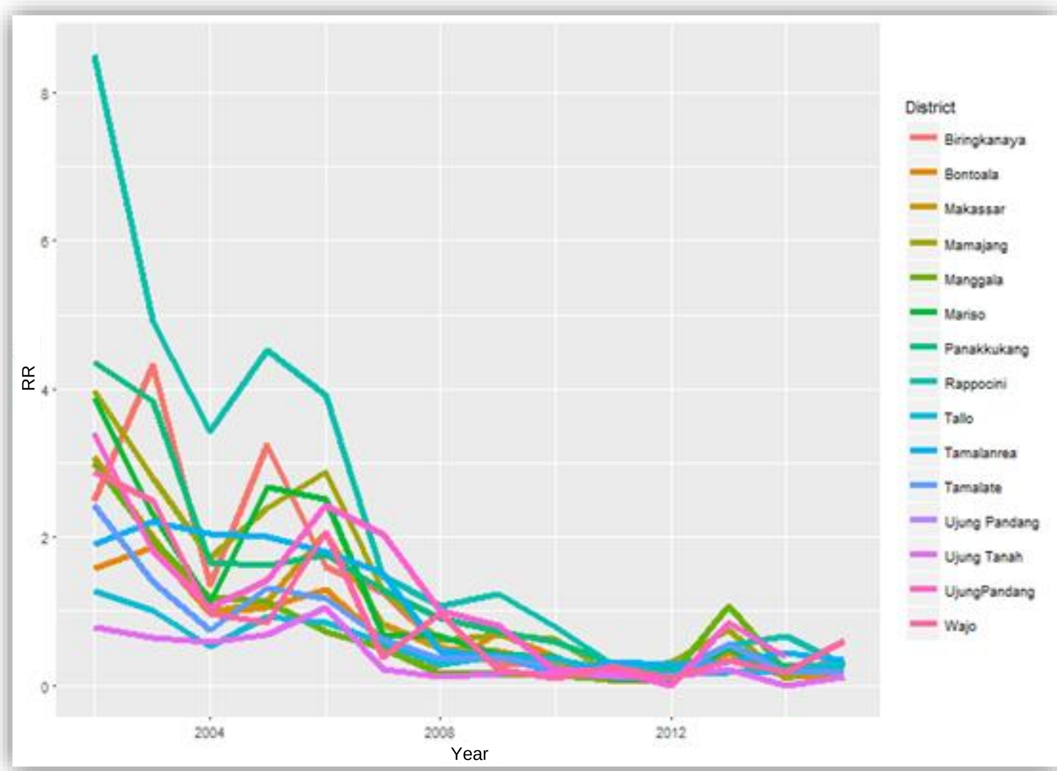


Figure 1. Crude risk estimates

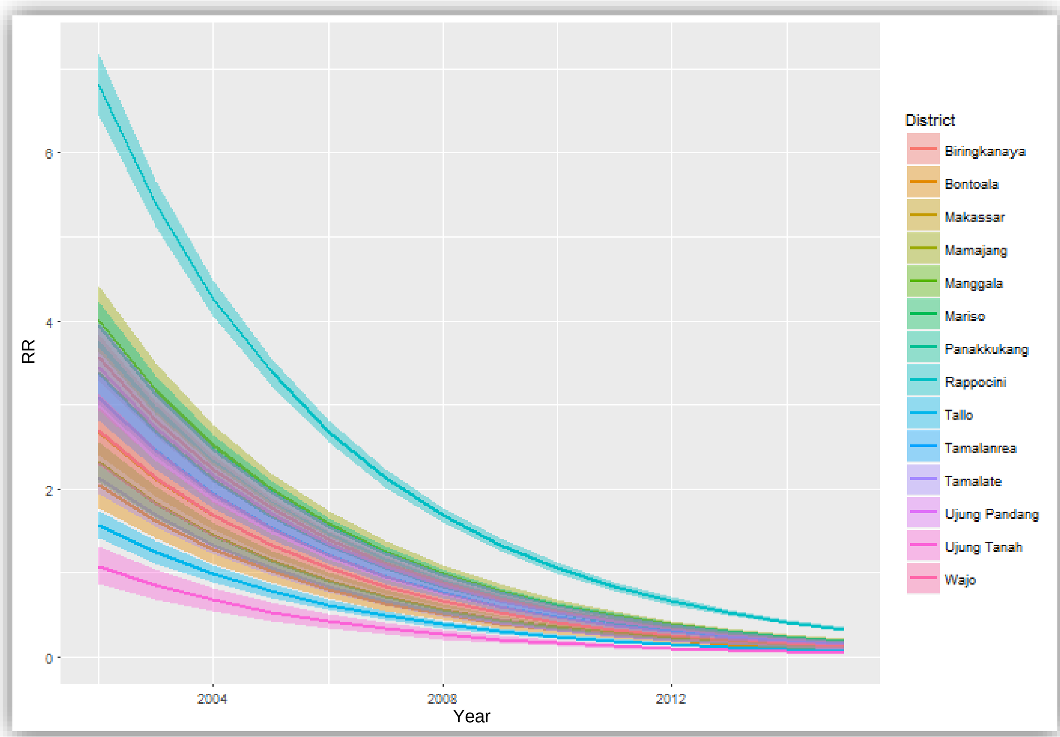


Figure 2. Linear temporal trend model

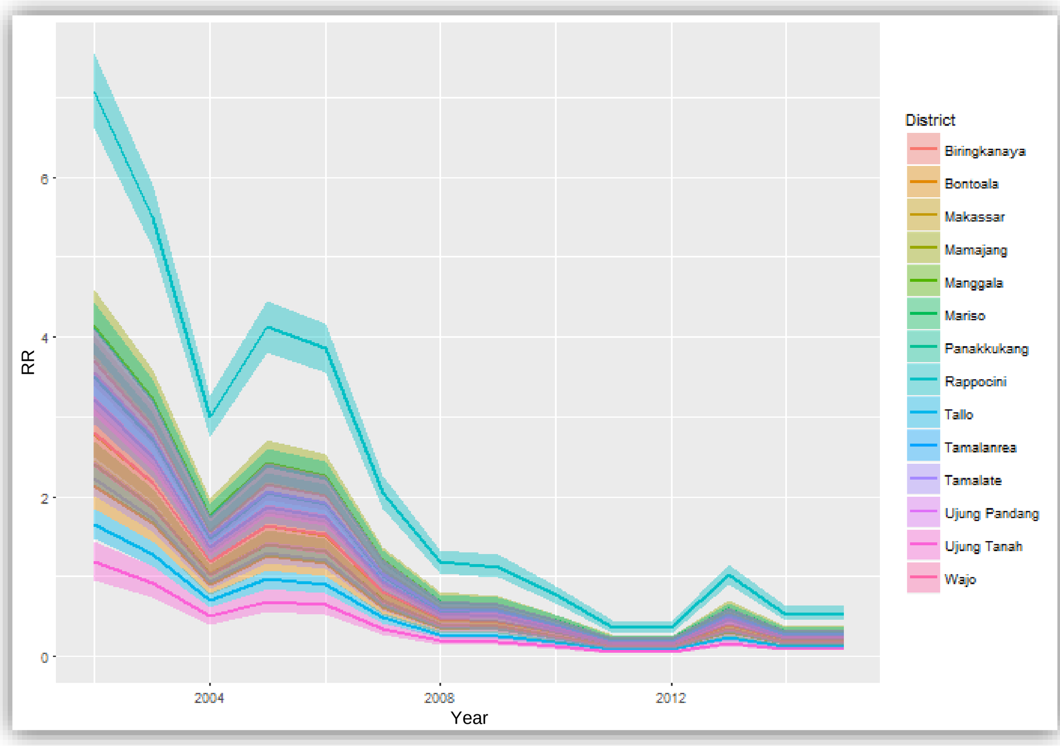


Figure 3. Nonparametric dynamic trend model

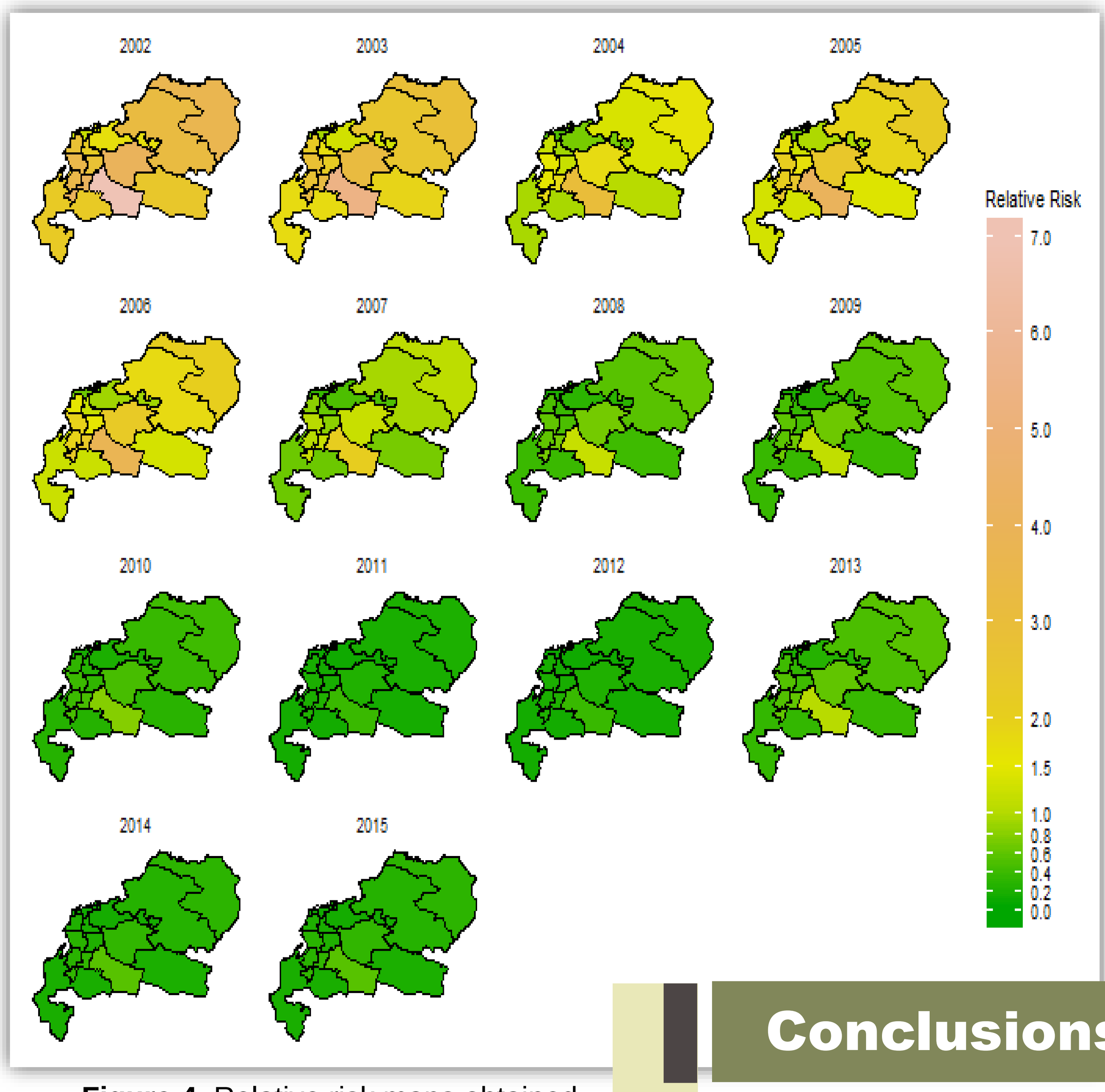
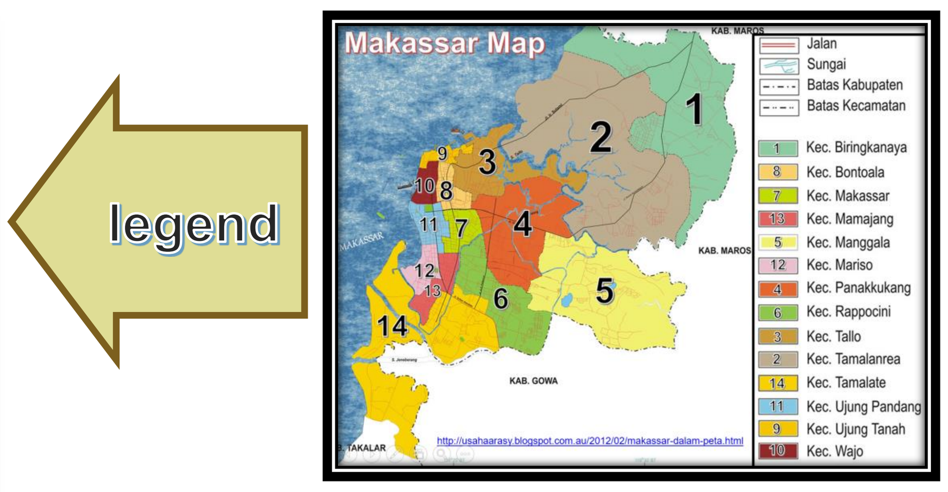


Figure 4. Relative risk maps obtained under the nonparametric dynamic trend model



- The median annual number of cases observed in a district was 16 (range: 0 - 419). Across all models, the highest relative risk (RR) was observed in Rappocini in 2002 and the estimate ranged from 2.12 to 7.07 between models.
- The RR in each area and year from the dynamic model had wider 95% CIs than the linear time trend spatio-temporal model.

Conclusions

Model choice had a large impact on results, and the nonparametric dynamic trend spatio-temporal model had a much better fit than other options. To ensure results are valid and reliable, careful exploration of a range of models is important, especially when there are few areas.

Acknowledgements

This work was supported by the ARC Centre of Excellence for Mathematical and Statistical Frontiers

References

[1] Mukhsar, et al., *Extended convolution model to bayesian spatio-temporal for diagnosing the DHF endemic locations.* Journal of Interdisciplinary Mathematics, 2016. **19**(2): p. 233-244.

[2] Besag, J., J. York, and A. Mollié, *Bayesian image restoration, with two applications in spatial statistics.* Annals of the institute of statistical mathematics, 1991. **43**(1): p. 1-20.

[3] Bernardinelli, L., Clayton, D., PascuttoBernardinelli, L., et al., *Bayesian analysis of space—time variation in disease risk.* Statistics in medicine, 1995. **14**(21-22): p. 2433-2443.

[4] Knorr-Held, L., *Bayesian modelling of inseparable space-time variation in disease risk.* 1999.