

PAPER • OPEN ACCESS

## Effects of Climatic Factors on Dengue Incidence: A Comparison of Bayesian Spatio-Temporal Models

To cite this article: Aswi Aswi *et al* 2021 *J. Phys.: Conf. Ser.* **1863** 012050

View the [article online](#) for updates and enhancements.



**240th ECS Meeting** ORLANDO, FL

Orange County Convention Center **Oct 10-14, 2021**

Abstract submission deadline extended: April 23rd

**SUBMIT NOW**

# Effects of Climatic Factors on Dengue Incidence: A Comparison of Bayesian Spatio-Temporal Models

Aswi Aswi<sup>1\*</sup>, Sukarna<sup>2</sup>, Susanna Cramb<sup>3,4</sup>, and Kerrie Mengersen<sup>3</sup>

<sup>1</sup>Statistics Department, Universitas Negeri Makassar, Indonesia

<sup>2</sup>Mathematics Department, Universitas Negeri Makassar, Indonesia

<sup>3</sup>Centre for Data Science, Queensland University of Technology, Brisbane, Australia

<sup>4</sup>Centre for Healthcare Transformation, School of Public Health and Social Work, Queensland University of Technology, Brisbane, Australia

\* E-mail: aswi@unm.ac.

**Abstract.** Considering only the spatial component of diseases can identify areas with reduced or elevated risk, but not capture anything about temporal variation of risk which could be more or equally crucial. Hence, both spatial and temporal components of diseases need to be considered. Bayesian methods are useful due to the ease of specifying additional information, including temporal or spatial structure, through prior distributions. Here, we examine a range of different Bayesian spatio-temporal models available using CARBayes. Combinations of model formulations and climatic covariates were compared using goodness-of-fit measures, such as Watanabe Akaike Information Criterion (WAIC). Comparisons were made in the context of a substantive case study, namely monthly dengue fever incidence from January 2013 to December 2017 and climatic covariates in 14 geographic areas of Makassar, Indonesia. A spatio-temporal conditional autoregressive adaptive model combining rainfall and average humidity provided the most suitable model.

## 1. Introduction

Spatial or spatio-temporal models using Bayesian methods are useful in modelling dengue fever. Despite this, Aswi et al. [1] found through a systematic review that only a limited number of studies included Bayesian spatio-temporal random effects when modelling dengue fever. When they were used, the spatial random effects were commonly assigned a conditional autoregressive (CAR) prior, while the temporal effects were commonly used the first order autoregressive AR(1).

Subsequently, Aswi et al. [2] compared six different Bayesian spatio-temporal CAR models using annual dengue data across Makassar, Indonesia from 2002 to 2015. The six models implemented in CARBayesST were compared: Spatio-temporal conditional Autoregressive (ST CAR) linear, ST CAR Autoregressive (AR), ST CAR adaptive, ST CAR separate spatial and ST CAR localised models, but no climatic variables were included. Given the mosquito vector for dengue fever has a lifecycle that requires certain temperatures and the availability of water, examining sub-annual timeframes, such as weekly or monthly, and the influence of climate is helpful. Another study used only Bayesian ST CAR localised models and some climatic covariates in examining monthly and annual dengue data [3]. They



found ST CAR localised models [4] performed well for annual data, identifying multiple distinct clusters or groups, but not for monthly data, as only a single group was identified.

Since the localised model was not recommended for monthly dengue fever data, this paper aims to examine the most suitable Bayesian ST CAR models in modelling monthly dengue fever with and without climatic factors.

## 2. Methods

### 2.1 Study Area

Makassar is the capital city of South Sulawesi Province that has 14 districts and covers an area of 175.77 km square. It has approximately 1.5 million population in 2017 [5]. The city is made up of 14 districts namely Biringkanaya, Bontoala, Mamajang, Manggala, Mariso, Makassar, Panakkukang, Rappocini, Tamalanrea, Tamalate, Tallo, Ujung Pandang, Ujung Tanah, and Wajo districts.

### 2.2 Dengue data and Climatic data

The monthly DHF incidence was acquired from the Health office of Makassar city from January 2013 to December 2017 for every district. Climatic data were obtained from the Meteorology, Climatology, and Geophysical Agency from January 2013 to December 2017 which consist of daily maximum, minimum, and mean temperature, rainfall amounts, and mean humidity (<http://dataonline.bmkg.go.id/home>). There are only four rainfall stations in Makassar, so the rainfall for each area is based on the nearest rainfall station. Calculations were performed in R using the `fields` package [6]. As the scale of climatic factors differs, all climatic data were standardised to have a mean zero and standard deviation one.

### 2.3. Models

Four Bayesian ST CAR models were used, namely ST CAR linear [7], ST CAR ANOVA [8], ST CAR AR [9], and ST CAR adaptive [10] in estimating the dengue fever risk and quantifying the risk associated between dengue fever and climatic covariate in Makassar, Indonesia. The dengue fever counts were modelled using the Poisson distribution. All these model combinations were analysed using the `CARBayesST` package [11] in the software package R [6]. These models are explained as follows.

$$y_{ij} \sim \text{Poisson}(E_{ij}\theta_{ij})$$

where  $y_{ij}$  and  $E_{ij}$  are the number of dengue cases and the expected number of dengue cases in area  $i$  and time  $j$ , respectively.  $\theta_{ij}$  is the relative risk of dengue.

Model formulations and combinations of climatic covariates were compared using the 95% posterior credible interval (considered substantive when the interval does not contain zero), and the goodness of fit measure, Watanabe Akaike Information Criterion (WAIC) [12]. Each model is given in detail below.

**2.3.1. Spatio temporal CAR Linear Model.** This model consists of four components, namely the intercept ( $\alpha_1$ ), spatial effect for all time ( $u_i$ ), temporal effect for all areas ( $\beta_j$ ), and space-time interaction ( $\delta_{ij}$ ). This model is given as follows [7] :

$$\log(\theta_{ij}) = \alpha_1 + u_i + (\beta + \delta_i) \frac{j - \bar{j}}{J}$$

where  $\mathbf{u}$  represent normally distributed random effects that describe spatial variation and  $\delta$  represent the interaction between spatial effects and temporal effects. The random-effects used Leroux priors as follows:

$$(u_i | u_{-i}, \mathbf{W}) \sim N\left(\frac{\rho_{\text{int}} \sum_{k=1}^I \omega_{ik} u_k}{\rho_{\text{int}} \sum_{k=1}^I \omega_{ik} + 1 - \rho_{\text{int}}}, \frac{\tau_{\text{int}}^2}{\rho_{\text{int}} \sum_{k=1}^I \omega_{ik} + 1 - \rho_{\text{int}}}\right),$$

$$(\delta_i | \delta_{-i}, \mathbf{W}) \sim N\left(\frac{\rho_{\text{slo}} \sum_{k=1}^I \omega_{ik} \delta_k}{\rho_{\text{slo}} \sum_{k=1}^I \omega_{ik+1} - \rho_{\text{slo}}}, \frac{\tau_{\text{slo}}^2}{\rho_{\text{slo}} \sum_{k=1}^I \omega_{ik+1} - \rho_{\text{slo}}}\right).$$

The adjacency matrix  $\mathbf{W} = (\omega_{ik})$  is defined using the simplest spatial weight matrix namely the binary neighbourhood matrix as follows:

$$w_{ik} = \begin{cases} 1 & \text{if areas } i \text{ and } k \text{ are adjacent} \\ 0 & \text{otherwise.} \end{cases}$$

$\tau_{\text{int}}^2$  and  $\tau_{\text{slo}}^2$  are the precision terms for the intercept and slope of the regression, respectively.  $\rho_{\text{int}}$  and  $\rho_{\text{slo}}$  are spatial dependence parameters with values ranging from zero to one. The default priors in the CARBayesST package were used as follows.

$\rho_{\text{int}}, \rho_{\text{slo}} \sim \text{Uniform}(0, 1)$ ;  $\beta \sim N(0, 1000)$  and  $\tau_{\text{int}}^2, \tau_{\text{slo}}^2 \sim \text{Inverse-Gamma}(1, 0.01)$ . In addition for the prior on the precision terms, we tried Inverse-Gamma(0.5, 0.005), and Inverse-Gamma(0.1, 0.01).

**2.3.2. Spatio temporal CAR ANOVA Model.** This model consists of three components, namely the spatial random effect over all time ( $u_i$ ), temporal random effect over all areas ( $\delta_j$ ), and space-time interaction ( $\gamma_{ij}$ ). This model is given as follows [8]:

$$\log(\theta_{ij}) = u_i + \delta_j + \gamma_{ij}$$

The priors for the  $u$  and  $\delta$  terms are as follows:

$$(u_i | u_{-i}, \mathbf{W}) \sim N\left(\frac{\rho_S \sum_{k=1}^I \omega_{ik} u_k}{\rho_S \sum_{k=1}^I \omega_{ik+1} - \rho_S}, \frac{\tau_S^2}{\rho_S \sum_{k=1}^I \omega_{ik+1} - \rho_S}\right),$$

$$(\delta_j | \delta_{-j}, \mathbf{D}) \sim N\left(\frac{\rho_T \sum_{k=1}^J d_{jk} \delta_k}{\rho_T \sum_{k=1}^J d_{jk+1} - \rho_T}, \frac{\tau_T^2}{\rho_T \sum_{k=1}^J d_{jk+1} - \rho_T}\right),$$

The adjacency matrix  $\mathbf{D} = (d_{jk})$  indicates the adjacency between times  $j$  and  $k$  and is defined as follows:

$$d_{jk} = \begin{cases} 1 & \text{if } |k - j| = 1 \\ 0 & \text{otherwise.} \end{cases}$$

An independent normal prior is used for  $\gamma$  namely  $\gamma_{ij} \sim N(0, \tau_\gamma^2)$ .  $\rho_S, \rho_T \sim \text{Uniform}(0, 1)$ .

The prior on the precision terms were used Inverse-Gamma(1, 0.01), Inverse-Gamma(0.5, 0.005), and Inverse-Gamma(0.1, 0.01).

**2.3.3 Spatio temporal CAR AR Model.** This model consists of one component only, namely the spatial random effect for each time ( $u_{ij}$ ) as follows [9]:

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

$$(u_j | u_{j-1},) \sim N(\rho_T u_{j-1}, \tau^2 \mathbf{Q}(\mathbf{W}, \rho_S)^{-1}) \quad j = 2, \dots, J,$$

$$u_1 \sim N(\mathbf{0}, \tau^2 \mathbf{Q}(\mathbf{W}, \rho_S)^{-1})$$

$$\rho_S, \rho_T \sim \text{Uniform}(0, 1).$$

The prior on the precision terms  $\tau^2$  were used three different priors as mentioned above.

*2.3.4 Spatio temporal CAR Adaptive Model.* ST CAR adaptive model is an extension of ST CAR AR. When the residual spatial dependence in the response is consistent over time but has a localised structure, ST CAR adaptive is suitable. The model structure of ST CAR adaptive is the same as CAR AR but nonzero (spatial) parts of the adjacency matrix (W) can vary locally. ST CAR adaptive elude the restrictive assumption that the estimation of the two adjacent areas must be similar [10].

*2.3.5 Sensitivity Analysis.* To examine the influence of the priors on the estimation of the posterior distribution, a sensitivity analysis was conducted. We used three distinct options for the prior on the variance terms namely Inverse-Gamma(1, 0.01), the default hyperprior specification in CARBayesST, Inverse-Gamma(0.5, 0.005), and Inverse-Gamma(0.1, 0.01).

### 3. Results

The results of all four Bayesian ST CAR models with and without climatic data for monthly dengue cases from January 2013 to December 2017 with three distinct options for the prior on the precision terms (see Section 2.3.5 Sensitivity Analysis) are given in Tables 1 to 3, and demonstrate insensitivity to the choice of hyperprior.

**Table 1.** Bayesian ST CAR models without and with climatic data for all 4 models for monthly dengue cases from January 2013 to December 2017 using Inverse-Gamma(1, 0.01)

ST CAR Linear		ST CAR ANOVA		ST CAR AR		ST CAR Adaptive		
Models	WAIC	Models	WAIC	Models	WAIC	Models	WAIC	
1	Without Covariates	2763.66	Without Covariates	1986.09	Without Covariates	2016.24	Without Covariates	2015.42
2	R*+AT*+AH*	<b>2592.57</b>	R+AT+AH*	1989.55	R*+AT+AH*	2008.69	R* +AT+AH*	2007.97
3	R*+ AH*	2599.71	R + AH	1987.84	R* + AH*	2009.31	R* + AH*	<b>2007.70</b>
4	R	2760.91	R	1985.01	R	2014.16	R	2011.80
5	AH*	2698.37	AH	1985.44	AH*	<b>2007.94</b>	AH	2009.04
6	AT	2772.14	AT	1991.74	AT	2013.48	AT	2010.44
7	R + AT	2776.18	R + AT	1991.14	R + AT	2013.57	R + AT	2015.63
8	R + MinT	2709.29	R + MinT	1983.50	R + MinT	2014.79	R + MinT	2018.19
9	R+ MaxT*	2762.61	R + MaxT	1985.61	R + MaxT	2013.81	R + MaxT	2015.75

\*95% posterior credible interval for the coefficient does not contain zero.

R, AH, AT, MaxT, MinT, are rainfall, average humidity, average temperature, maximum temperature, and minimum temperature, respectively.

**Table 2.** Bayesian ST CAR models without and with climatic data for all 4 models for monthly dengue cases from January 2013 to December 2017 using Inverse-Gamma(0.5, 0.005)

ST CAR Linear		ST CAR ANOVA		ST CAR AR		ST CAR Adaptive		
Models	WAIC	Models	WAIC	Models	WAIC	Models	WAIC	
1	Without Covariates	2773.18	Without Covariates	1986.32	Without Covariates	2012.95	Without Covariates	2014.18
2	R*+AT*+AH*	<b>2592.54</b>	R+AT+AH	1985.15	R+AT+AH*	2004.09	R* +AT+AH*	2004.42
3	R*+ AH*	2597.14	R + AH*	1987.57	R + AH*	2003.84	R* + AH*	<b>2008.00</b>
4	R	2777.93	R	1985.02	R	2016.04	R	2015.73
5	AH*	2701.96	AH	1986.04	AH*	<b>2009.57</b>	AH*	2012.95
6	AT	2772.37	AT	1988.93	AT	2010.67	AT	2015.59
7	R + AT	2784.28	R + AT	1985.48	R + AT	2008.54	R + AT	2019.12
8	R + MinT*	2718.72	R + MinT	1984.43	R + MinT	2011.27	R + MinT	2015.54
9	R+ MaxT*	2767.02	R + MaxT	1991.03	R + MaxT	2016.20	R + MaxT	2019.11

\*95% posterior credible interval for the coefficient does not contain zero.

R, AH, AT, MaxT, MinT, are rainfall, average humidity, average temperature, maximum temperature, and minimum temperature, respectively.

**Table 3.** Bayesian ST CAR models without and with climatic data for all 4 models for monthly dengue cases from January 2013 to December 2017 using Inverse-Gamma (0.1, 0.01)

ST CAR Linear		ST CAR ANOVA		ST CAR AR		ST CAR Adaptive		
Models	WAIC	Models	WAIC	Models	WAIC	Models	WAIC	
1	Without Covariates	2779.06	Without Covariates	1983.12	Without Covariates	2008.79	Without Covariates	2009.59
2	R*+AT*+AH*	<b>2602.96</b>	R+AT+AH	1991.26	R*+AT+AH*	2004.29	R* +AT+AH*	1999.49
3	R*+ AH*	2611.92	R + AH	1987.95	R* + AH*	2006.28	R* + AH*	<b>2003.39</b>
4	R	2783.11	R	1987.41	R	2012.16	R	2017.86
5	AH*	2718.13	AH	1988.41	AH*	<b>2006.90</b>	AH*	2006.41
6	AT	2782.77	AT	1981.02	AT	2012.96	AT	2012.74
7	R + AT	2786.54	R + AT	1988.72	R + AT	2014.61	R + AT	2008.96
8	R + MinT*	2720.42	R + MinT	1985.24	R + MinT	2013.82	R + MinT	2016.41
9	R+ MaxT*	2784.65	R + MaxT	1988.83	R + MaxT	2011.23	R + MaxT	2017.35

\*95% posterior credible interval for the coefficient does not contain zero.

R, AH, AT, MaxT, MinT, are rainfall, average humidity, average temperature, maximum temperature and minimum temperature, respectively.

Overall, results differ by model, but not by hyperprior choice. The lowest WAIC overall (1981.02) is the ST CAR ANOVA with average temperature and the highest WAIC was the ST CAR linear model (2786.54) with rainfall and average temperature (Table 3). However, both of these models found these variables were not considered significant (defined as the 95% credible interval including zero).

Considering additionally the significance of the included covariates by model type, the ST CAR linear model including rainfall, average temperature, and the average humidity is the best of the linear models as it has a small WAIC and 95% credible intervals for these three climatic covariates does not contain zero. The same conclusion has been found for all three different priors on the precision. ST CAR ANOVA model was the worst model in this case as the 95% CI for all climatic covariates contains zero. ST CAR AR model with the inclusion of average humidity and rainfall was the best model (2006.28), and the WAIC is indistinguishable from that for the ST CAR AR model including average humidity (2006.90) (Table 3). Under ST CAR adaptive model with the inclusion of rainfall and average humidity was the best model. Overall, based on 95% CI and the smallest WAIC, the best model in modelling dengue fever is ST CAR adaptive model with rainfall and average humidity incorporated (2003.39).

Our results indicated that rainfall amounts and average humidity significantly influenced the relative risk of dengue. There was a negative correlation between rainfall and the dengue relative risk. However, the correlation between average humidity and the relative risk of dengue was positive. This importance of climatic covariates (rainfall, and average humidity) is similar to some previous research [3, 13, 14].

#### 4. Conclusion and Future Work

In conclusion, our results suggest that we need to be careful in choosing Bayesian ST CAR models as they can have different results. When no covariates were included, the CAR ANOVA performed well. However, trying more than one model is recommended, especially when including explanatory variables. Based on the results, it can be concluded that the Bayesian ST CAR adaptive model incorporating rainfall and average humidity performed well, as did the ST CAR AR including average humidity. Considering other approaches in determining which variables to include in the model such as the least absolute shrinkage and selection operator (Lasso) and Bayesian Lasso methods could be possible future work.

#### Acknowledgements

The authors wish to acknowledge the city health department of Makassar for allowing the dengue fever dataset to be used in this study. The authors also acknowledge the Meteorology, Climatology, and Geophysical Agency for providing climatic data.

#### References

- [1] Aswi A, Cramb S M, Moraga P and Mengersen K 2019 *Epidemiol Infect* **147**
- [2] Aswi A, Cramb S, Hu W, White G and Mengersen K 2020 *In Case Studies in Applied Bayesian Data Science*, ed K Mengersen, *et al.* (Switzerland: Springer) pp 229-244
- [3] Aswi A, Cramb S, Duncan E, Hu W, White G and Mengersen K 2020 *Spat. Spatio-temporal Epidemiol.* **33**
- [4] Lee D and Lawson A 2016 *Ann. Appl. Stat* **10** 1427-1446
- [5] Badan Pusat Statistik 2018 Makassar Municipality in figures 2018 (Makassar: BPS)
- [6] R Core Team 2019 R: A language and environment for statistical computing (Vienna, Austria: R Foundation for Statistical Computing)
- [7] Bernardinelli L, Clayton D, Pascutto C, Montomoli C, Ghislandi M and Songini M 1995 *Stat. Med* **14** 2433-2443
- [8] Knorr-Held L 1999 *Stat. Med* **19** 2555-2567
- [9] Rushworth A, Lee D and Mitchell R 2014 *Spat. Spatio-temporal Epidemiol* **10** 29-38
- [10] Rushworth A, Lee D and Sarran C 2017 *J. R. Stat. Soc. Ser. C* **66** 141-157

- [11] Lee D, Rushworth A and Napier G 2018 *J. Stat. Softw.* **84** 1-39
- [12] Watanabe S 2010 *J. Mach. Learn. Res.* **11** 3571-3594
- [13] Lowe R, Cazelles B, Paul R and Rodó X 2016 *Stoch. Environ. Res. Risk Assess* **30** 2067-2078
- [14] Malik A, Yasar A, Tabinda A B, Zaheer I E, Malik K, Batool A and Mahfooz Y 2017 *Environ Monit Assess* **189** 1-20