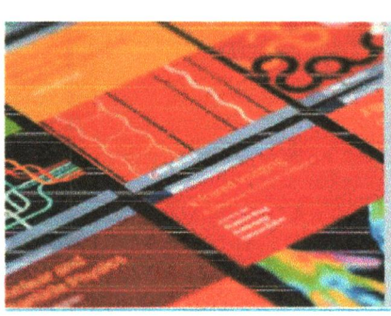
Estimation of Relative Risk of Dengue Fever in Makassar Using Localized Bayesian Autoregressive Conditional Model

by

Journal of Physics: Conference Series

Estimation of Relative Risk of Dengue Fever in Makassar Using Localized Bayesian Autoregressive Conditional Model



Estimation of Relative Risk of Dengue Fever in Makassar Using Localized Bayesian Autoregressive Conditional Model

S A Thamrin[1*], R Khaerati[1], A K Jaya[1] and Ansariadi[2]

[1]Department of Statistics, Hasanuddin University, Makassar, Indonesia
[2]Department of Epidemiology, Hasanuddin University, Makassar, Indonesia

*e-mail: tuti@unhas.ac.id

Abstract. Analysis of the relative risk of the spread of dengue fever (DF) in Makassar city, Indonesia, needs to be done to see the which areas are at high risk of DF. Bayesian Autoregressive Conditional (CAR) is used in the mapping model of this disease. This model is Able to model of relative risk by taking into account the smoothing of that relative risk and entering spatial information to reduce the error of the estimated parameters in order the reliable relative risk models is Obtained. In this study, the relative risk value of the spread of DF was Analyzed using the localized Bayesian models CAR. Under this model the geographical mapping of DF in Makassar can be identified for each sub-district and shows that Makassar is still very vulnerable to DF.

Keywords: Bayesian, conditional autoregressive, dengue fever, a relative risk reduction, localized models

1. Introduction

Dengue fever (DF) is an infectious disease caused by dengue virus and transmitted by Aedes Aegypty mosquitoes [1]. DF is still a major public health problem because of its spread, severity and material loss. This will increase if there is no early prevention and eradication of this disease [2].

Aedes Aegypty mosquitoes can only breed in tropical areas with temperatures above 16 ℃ and at altitudes less than 1,000 meters above sea level [3]. One of the regions in Indonesia included in the criteria is the Makassar city. Topographically, the Makassar city has temperatures ranging between 26.2 ℃ and 29.8 ℃ and is located between 1-25 meters above sea level. Therefore, the Makassar city is an area that is often visited by Aedes mosquitoes to breed. As a result, many cases of DF in the city occurred. The level of rainfall in the Makassar city ranges from an average of 2,729 mm with a number of rainy days starting from an average of 144 days [4].

Analysis of the relative risk of the spread of DF in Makassar city needs to be done to see which areas in Makassar city are at high risk of getting DF. Jaya, et al. [5] have conducted research on the Spatial Bayesian in estimating the relative risk with the R-Integrated Nested Laplace Approximation (INLA) approach. The INLA method is used as an alternative faster solution for estimating the posterior distribution that has so far used the MCMC approach. The relative risk assessment using the Bayesian CAR Besag-York-Mollie (BYM) model has been carried out by Samat and Mey [6] and Thamrin and Alimun [7]. In addition, Lee and Sarran [8] used a localized model with a clustering system to estimate the relative risk in the case of air pollution resulting in a finer relative risk value.

In this study, the relative risk value of the spread of DF will be analyzed using the Bayesian Conditional Autoregressive (CAR) method with a localized model. Under this model, disease mapping assumes that the number of people with DF follows the Poisson distribution and the existence of cluster structures in the data. Data changes can occur suddenly so that a flexible model is needed to overcome these changes without affecting the relative risks generated previously. This localized model is able to overcome this. With this model, the risk level of the spread of dengue disease every year for each district in Makassar can be known. Therefore, this study aims to identify the relative risk area of DF spread in Makassar city through a geographical map.

2. Material and Methods

2.1 Data source

This study uses data obtained from the Makassar City Health Office, South Sulawesi, Indonesia, covering all districts in period 2011-2016. The data taken is the number of dengue fever cases for each district in Makassar city. The variables used in this study were the number of dengue cases and locations (latitude and longitude), stating the districts in Makassar city.

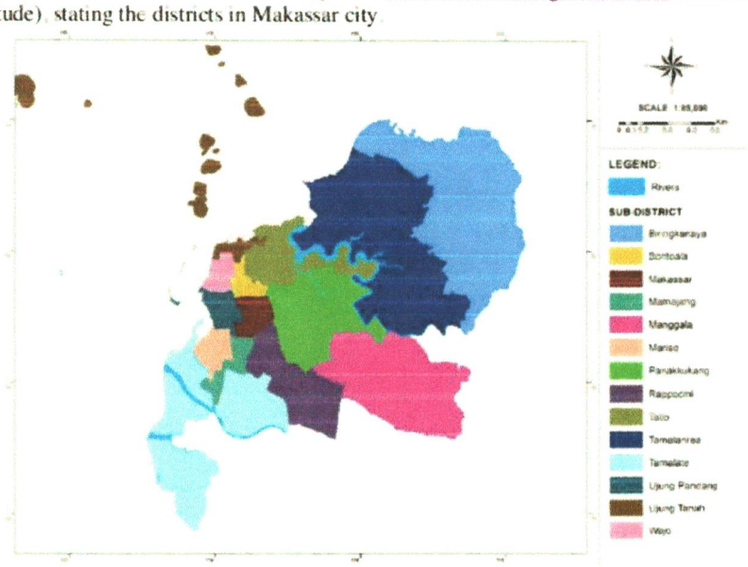


Figure 1. Local Map of Makassar, South Sulawesi, Indonesia

DF sufferers in Makassar city continue to occur every year. DF sufferers observed in the Makassar city are divided into 14 districts as shown in Figure 1 of the map of Makassar city. Rappocini sub-district has the highest total number of dengue cases, 172 cases, followed by Biringkanaya and Manggala sub-districts with 115 cases. Meanwhile, Ujung Tanah, Wajo, Mamajang and Ujung Pandang sub-districts are the sub-districts with the least number of DF sufferers in 2011 to 2016.

Autoregressive Conditional Localized Bayesian Model

Bayesian Conditional Autoregressive (CAR) is a technique in mapping diseases that models relative risk by taking into account the smoothing of the estimated relative risk. This method also includes spatial information to reduce errors from the estimated relative risk parameters so that a better relative risk estimate is obtained [9]. The term Bayesian refers to the concept of a smoothing model while CAR refers

to a model that allows entering spatial information in modeling. CAR was first introduced by Besag [10].

In the Bayesian approach, the information contained in the data is used, in this case the number of events in each region. The first step that must be taken is to determine the distribution of prior p (ψ) to obtain the relative risk value so as to provide information about the variability of events in the area.

Suppose the random variable $y = \{y_1, \ldots, y_m\}$is univariate vector which expresses the number of cases on the location of the k^{th} (k = 1, 2, ..., m). The general model CAR Bayes written as:

$$Y_k|\mu_k \sim f(y_k|\mu_k, v^2) \text{ for } k = 1, \ldots, n \quad (1)$$

$$g(\mu_k) = x_k^T\beta + \phi_k + O_k$$

Y_k an exponential family of distributions. In CAR Bayes family can be binomial, Gaussian or Poisson. Y_k is an exponential family of the distribution $f(y_k|\mu_k, v^2)$. In Bayes CAR it can be a binomial, Gaussian or Poisson family. The expected value of Y_k is denoted by $E(Y_k) = \mu_k$, while v^2 is the additional scale parameter needed if the Gaussian family is used. The link function for the Generalized Linear Model case is denoted by g (μ_k). In computing this link function, it can be a logit function (binomial family), identity (Gaussian family) or natural log (Poisson family). The regression parameter vector is denoted by $\beta = \{\beta_1, \ldots, \beta_p\}$, and the non-linear covariate effect can be entered into the equation model (1) through the natural cubic spline function or polynomial basis function [11].

In this study it is assumed that the conditional probability distribution is Poisson and CAR models. It will be used to estimate relative risk. To solve equation (1), the hierarchical Bayes approach is used through computational Integrated Nested Laplace Approximations (INLA) for estimating model parameters.

2.3 Spatial CAR Localized

Determination of the prior value (ϕ) in Besag et al. [12] produces spatial models based on their respective regions given by the following models.

$$COR(\phi_k, \phi_j|\phi_{-kj}) = \frac{\rho w_{kj}}{\sqrt{(\rho\sum_{i=1}^{n} w_{ki}+1-\rho)(\rho\sum_{i=1}^{n} w_{ji}+1-\rho)}} \quad (2)$$

For non-neighboring regions $w_{kj} = 0$, random effects are independent. Then, for neighboring regions, the spatial correlation is determined by ρ. However, the approach with equation (2) allows spatial local autocorrelation to occur which only focuses on the value of spatial autocorrelation in a particular area. As a result, it can reduce the occurrence of spatial autocorrelation between certain pairs of neighbors. To overcome this, it can be done by adding a series of spatial random effects that divide the area into several clusters, so that it allows quite significant differences in the average between adjacent area units in different groups. One model proposed by Lee and Sarran [8] that partition the area units into the maximum number of clusters determined by the researcher and symbolized by $(\lambda_1, \ldots, \lambda_G)$. This model is given as:

$$\psi_k = \phi_k + \lambda z_k \quad (3)$$

$$\phi_k|\phi_{-k}, \boldsymbol{W}, \tau^2 \sim N\left(\frac{\sum_{i=1}^{K} w_{ki}\phi_i}{\sum_{i=1}^{K} w_{ki}}, \frac{\tau^2}{\sum_{i=1}^{n} w_{ki}}\right)$$

$$\tau^2 \sim \text{Inverse} - \text{Gamma}\,(0.001, 0.001)$$

$$\lambda_i \sim \text{Uniform}\,(\lambda_{i-1}, \lambda_{i+1}) \quad \text{untuk } i = 1, \ldots, G$$

$$f(Z_k) = \frac{exp(-\delta(Z_k - G^*)^2)}{\sum_{r=1}^{G}\exp(-\delta(r - G^*)^2)}$$

$$\delta \sim \text{Uniform}(1, M = 10)$$

The cluster is represented by $(\lambda_1, \ldots, \lambda_G)$ where $\lambda_1 < \lambda_2 < \cdots < \lambda_G$. This requirement applies to prevent the problem of changes in value (such as the number of cases) in areas that are indeed at risk of having an increasing number of sufferers of diseases such as DF. The division of area units is denoted by K, and G is the maximum number of clusters desired by the researcher. In this model, the clusters are ordered by the resulting average values (λ_i) and $\delta\,(Z_k - G^*)^2$ are prior to Z_k, where $G^* = (G+1)/2$ if G is selected is odd and $G^* = G/2$ if the selected G is even. By using equation (3), clustered area units will produce different relative risk values. Therefore, areas that are close together and have different values, can still be calculated even though the area units are allocated to various groups. Lee and Sarran [8] recommend researchers to choose G in odd numbers. The selected G value must be of small value, because $(\lambda_1, \ldots, \lambda_G)$ is designed to estimate neighboring regions that have different random effects, but areas on different sides of the study area can have the same value $(\lambda_1, \ldots, \lambda_G)$. In addition, the spatial variations produced as a prior are modeled by $\phi = (\phi_1, \ldots, \phi_n)$, and with such spatial variations the models do not match the desired smoothing structure. Therefore, choosing G to be an odd number and as small as 3 or 5 is recommended.

3. Result

The results of the relative risk estimation based on the application of the localized model of equations (2) and (3) applied to the number of cases observed from the spread of DF in Makassar, Indonesia are displayed in the form of graphs, tables and maps. Figure 2 illustrates the time series of DF plots based on the number of cases for each sub-district in Makassar, Indonesia from 2011 to 2016. From Figure 2 it is shown that Rappocini sub-district has the highest number of DF sufferers compared to other subdistricts. While for Wajo sub-district, every year there are less than 10 cases. This is also similar to the subdistricts of Bontoala and Ujung Tanah, although in 2016 there were more than 10 cases.

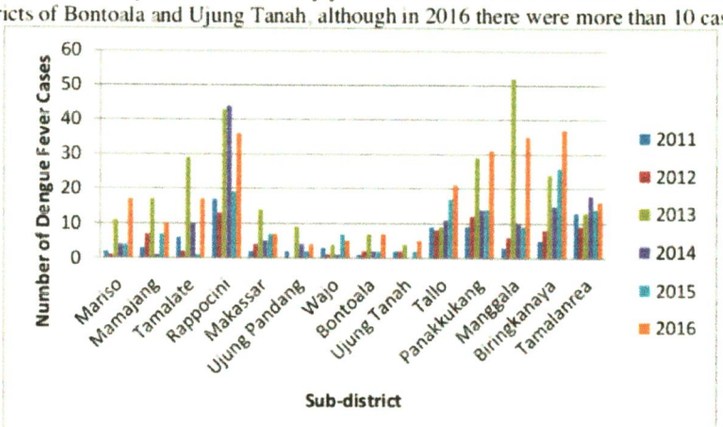


Figure 2. Number of DF cases for each district in the Makassar city in 2011-2016

Estimated relative risk based on the Local model for DF in all subdistricts in Makassar in 2011-2016 is presented in Table 1. Based on Table 1, Rappocini sub-district has the highest relative risk compared to other districts in 2014 but the number of cases decreased in 2015. Then Manggala and Tamalanrea subdistricts had the highest relative risk of spreading DF in 2013 and 2014. This shows that the Rappocini, Manggala and Tamalanrea subdistricts have a higher risk of occurrence compared to other subdistricts in Makassar. Then the sub-districts that had the smallest relative risk were in Ujung Tanah sub-district in 2014 and Tallo sub-district in 2013. From the low relative risk, it showed that the area was less likely to be affected by DF.

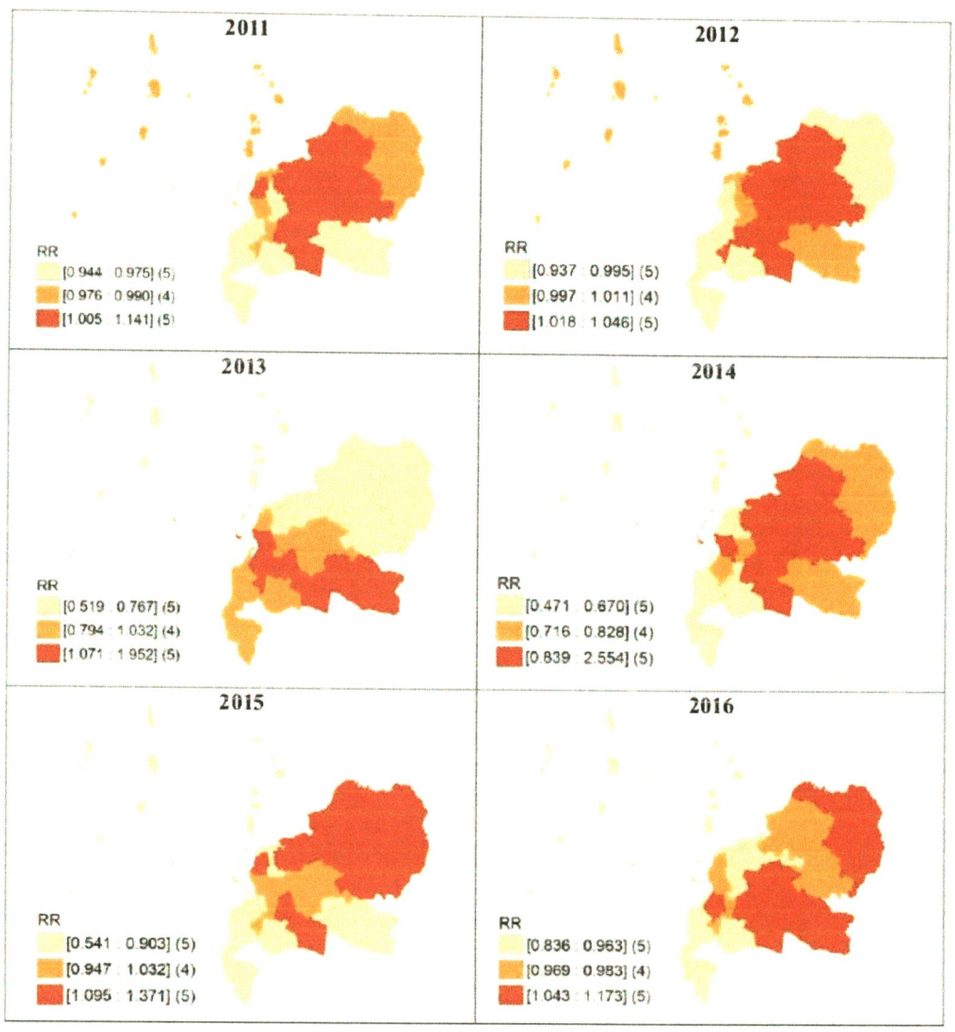


Figure 3. Relative risk of DF in Makassar city from 2011-2016

Figure 3 shows the relative risk map of dengue disease that can identify districts that have a high risk of disease. This map is obtained based on the results of the relative risk estimation with the localized model according to Table 1. To see the level of risk of disease in each subdistrict, the relative risk value is divided into 3 clusters for each year. The division of regions in 3 clusters was chosen to produce a better and smoother relative risk value. For example, in 2014, the relative risk area of DF was divided into 3 clusters with a range of values for the first cluster of 0.47 - 0.67, the second cluster was in the range 0.72 - 0.83 and the third cluster was in the range 0.84 - 2.55. Because DF cases in each subdistrict are different each year, the range of values of relative risk will also vary, but will still be divided into 3 clusters for each year. Maps for the 2015 value of the relative risks generated are all above 0.5 with the lowest value being in the subdistrict of Tamalate. In addition, Rappocini subdistrict has a high relative risk every year. This shows that this area is vulnerable to DF. The estimated relative risk obtained also

considers the population of each sub-district. Areas that have a high population with a small disease case, it will produce a relatively small risk value and vice versa.

Table 1. Relative risk of dengue every subdistrict in the Makassar city, Indonesia, in 2011-2016.

Sub-district	Relative Risk					
	2011	**2012**	**2013**	**2014**	**2015**	**2016**
Wajo	1.0053	0.9948	0.7938	0.6699	1.3713	0.9732
Ujung Pandang	0.9899	0.9840	1.2228	1.1534	0.9469	0.9690
headland	0.9776	0.9999	0.5956	0.4713	0.8877	0.9199
Tamalate	0.9575	0.9367	0.9186	0.6256	0.5409	0.8357
Tamalanrea	1.1413	1.0442	0.7556	1.5090	1.2116	0.9740
Tallo	1.0175	1.0176	0.5188	0.8388	1.1767	0.9630
Rappocini	1.1195	1.0425	1.3811	2.5539	1.0946	1.0909
Panakukkang	1.0098	1.0464	1.0318	0.9675	1.0322	1.0721
Mariso	0.9629	0.9698	1.0713	0.7881	0.8508	1.1217
Mangala	0.9747	1.0107	1.9521	0.8159	0.9032	1.1732
Mamajang	0.9764	1.0339	1.3176	0.5120	1.0113	0.9831
Makassar	0.9443	1.0040	0.9894	0.7159	0.9623	0.8927
Bontoala	0.9484	0.9968	0.7671	0.6143	0.8647	0.9423
Biring Kanaya	0.9865	0.9952	0.7093	0.8284	1.3342	1.0428

4. Conclusion

The localized model can be used to estimate the relative risk of dengue in the Makassar city, Indonesia. From the relative risk value, the map of the distribution of high and low risk areas can be known based on the incidence of DF in the Makassar city from 2011-2016. Areas that are vulnerable to dengue disease need to get more attention from the Makassar city government to overcome the spread of the disease to other regions.

Acknowledgments

This research was funded by the Ministry of Research, Technology and Higher Education through grants PDUPT No. 1740 / UN4.21 / PL.01.10 / 2019.

References

[1] Candra A. 2019. Dengue Epidemiology, Pathogenesis and Transmission Risk Factors. *Aspirator* **2** (2) 110-119

[2] Respati T, Raksanegara A, Djuhaeni H, Sofyan A, Agustian D, Faridah L, and Sukandar H. 2017. Various Factors Affecting Genesis DFdi Bandung. *Aspirator* **9** (2) 91-96

[3] Hidayati Y. 2017. *Relationship Between Aedes Aegypti Mosquito Breeding Sites with Dengue Cases in Sub Rajabasa Bandar Lampung*. Thesis State Islamic Institute Raden Intan Lampung

[4] Climate Data. 2018. Climate of Makassar. https://id.climate-data.org/asia/indonesia/south-sulawesi/ makassar-3646 / [Access online 09 Sept. 2019]

[5] Jaya IGNM, Tantular B, Zulhanif. 2017. Spatial Bayesian varying Coefficient Estimating Relative Risk Model in Diarrhea in Bandung, the National Conference of the Mathematical Research and Learning was II (KNPMP II) ISSN: 2502-6526

[6] Samat NA and Mey L W. 2017. Malaria Disease Mapping in Malaysia based on Besag-York-Mollie (BYM) Model. *Journal of Physics: Conference Series IOP Conf. Series: Journal of Physics: Conf. Series* **890** (2017) 012167

[7] Thamrin, SA, and Alimun, 2019. Geographical Mapping of Dengue Fever Incidence from 2012 to 2016 in Makassar, Indonesia. *IOP Conf. Series: Earth and Environmental Science* **279** (012013): 1-7

[8] Lee, D and Sarran, C. 2015. Controlling for unmeasured confounding and spatial misalignment in long-term studies of water pollution and health. *Environmetrics* **26**: 477-487

[9] Venkatesan P, Srinivasan R, and Dharuman. 2012. Auto Regressive Conditional Bayesian Model for Mapping Tuberculosis Prevalence in India International. *Journal of Pharmaceutical Studies and Research* **3** (1) 1-3

[10] Besag, J. 1974. Spatial Interaction and the Statistical Analysis System Lattice. *Journal of the Royal Statistics* **36**:192-236

[11] Lee, D. 2013. CAR Bayes: An R Package for Bayesian Spatial Modeling with Conditional Autoregressive Priors. *Journal of Statistical Software* **55**: 1-13

[12] Besag J, J York and Mollie, A. 1991. Bayesian image restoration with two applications of spatial statistics. *Annals of the Institute of Statistics and Mathematics* **43**: 1-59

Estimation of Relative Risk of Dengue Fever in Makassar Using Localized Bayesian Autoregressive Conditional Model

ORIGINALITY REPORT

7%	6%	3%	1%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	journal-old.unhas.ac.id Internet Source	2%
2	Aswi Aswi, Susanna Cramb, Earl Duncan, Wenbiao Hu, Gentry White, Kerrie Mengersen. "Climate variability and dengue fever in Makassar, Indonesia: Bayesian spatio-temporal modelling", Spatial and Spatio-temporal Epidemiology, 2020 Publication	1%
3	moam.info Internet Source	1%
4	rdrr.io Internet Source	1%
5	www.tandfonline.com Internet Source	1%
6	mran.microsoft.com Internet Source	<1%
7	www.neliti.com Internet Source	<1%

8	cran.r-project.org Internet Source	<1 %
9	downloads.hindawi.com Internet Source	<1 %
10	biomedres.us Internet Source	<1 %
11	Yulia Yusrini Djabir, Aryadi Arsyad, Mufidah Murdifin, Rosany Tayeb et al. "Kleinhovia hospita extract alleviates experimental hepatic and renal toxicities induced by a combination of antituberculosis drugs", Journal of Herbmed Pharmacology, 2020 Publication	<1 %
12	www.nature.com Internet Source	<1 %

Exclude quotes	On	Exclude matches	< 5 words
Exclude bibliography	On		