Bayesian spatial conditional autoregressive (CAR) Leroux model of Covid-19 cases in Makassar, Indonesia

Cite as: AIP Conference Proceedings **2662**, 020028 (2022); https://doi.org/10.1063/5.0108032
Published Online: 22 December 2022

Aswi, Muhammad Arif Tiro and Zulkifli Rais

Bayesian Spatial Conditional Autoregressive (CAR) Leroux Model of Covid-19 Cases in Makassar, Indonesia

Aswi[1 a)], Muhammad Arif Tiro[1], and Zulkifli Rais[1]

[1]*Statistics Department, Universitas Negeri Makassar, Indonesia*

[a)] Corresponding author: aswi@unm.ac.id

Abstract. The number of positive patients for the Coronavirus disease-2019 (Covid-19) is growing exponentially. South Sulawesi Province, one of the provinces in Indonesia, has the highest number of Covid-19 cases outside Java Island. Makassar City as the provincial capital of South Sulawesi has the highest number of positive confirmed Covid-19 cases in South Sulawesi. This study aims to estimate the relative risk of Covid-19 cases in Makassar by comparing several Bayesian Spatial Conditional Autoregressive (CAR) Leroux models. Data on the number of confirmed positive cases of Covid-19 from March 20, 2020, to August 30, 2021, in every sub-district (15 sub-districts) in Makassar City are used. In addition, data on the number of populations from each sub-district are also used to calculate the expected value of the occurrence of Covid-19 cases. The selection of the best model is based on several criteria, namely Deviance Information Criteria (DIC), Watanabe Akaike Information Criteria (WAIC), and residuals from Moran's *I* Modification (MMI). The Bayesian spatial CAR Leroux model with hyperprior Inverse-Gamma (0.5;0.05) is preferred to model the confirmed Covid-19 cases in Makassar city. Ujung Pandang has the highest relative risk of Covid-19 while Sangkarrang Island has the lowest relative risk of Covid-19. These results may help policy makers in decision-making.

INTRODUCTION

Since late 2019, the outbreak of coronavirus disease 2019 (Covid-19) has spread exponentially around the world [1]. Covid-19 is affecting 221 countries and territories with 219,621,970 confirmed cases; 196,405,045 recovered and 4,550,307 deaths by 2 September 2021 [2]. On 2 March 2020, the first case of Covid-19 in Indonesia was reported [3]. A total of 4,079,267 people were confirmed with Covid-19 in Indonesia by 30 August 2021 with 203,060 people in treatment; 3,743,716 recovered and 132,491 people died (https://infocorona.makassar.go.id/) [4]. In Indonesia, the highest number of COVID-19 cases outside Java Island was reported from South Sulawesi Province [5]. A total of 104,146 people were confirmed with Covid-19 in South Sulawesi Province with a known 94,831 recovered and 1,979 deaths. Makassar, the capital city of South Sulawesi Province, has the highest number of Covid-19 cases among 24 districts/cities. A total of 47,293 confirmed cases were reported with at least 44,243 recovered and 954 deaths from 19 March 2020 to 30 August 2021.

Research on Covid-19 modeling has been conducted. However, only a limited number of studies considered Bayesian hierarchical Spatial models. Bayesian hierarchical spatial BYM (Besag, York & Mollié) model has been used to investigate the association between socioeconomic factors and Covid-19 cases across neighborhoods in New York City [6]. They compared between Bayesian Poisson model and Bayesian Negative Binomial spatial models. Another research used Bayesian hierarchical Poisson spatial BYM models to investigate the association between the number of positive Covid-19 cases and the proportion of black/African American populations [7]. The association between Covid-19 deaths and NO2 and PM2·5 in England using a Bayesian hierarchical Poisson log-linear model has also been investigated [8]. A re-parametrization of the BYM CAR prior was used to model the spatial random effect [8].

Several Bayesian hierarchical spatial models are commonly used for disease mapping of areal unit data such as CAR BYM, CAR Leroux, CAR localised models and have been compared. When the Leroux, BYM, and Cressie models were compared, the Leroux model was better than other models. This is because the Leroux model allows

International Conference on Statistics and Data Science 2021
AIP Conf. Proc. 2662, 020028-1–020028-5; https://doi.org/10.1063/5.0108032
Published by AIP Publishing. 978-0-7354-4249-8/$30.00

having a range of spatial autocorrelation [9]. Different Bayesian spatial models included CAR BYM, CAR Leroux, CAR localised with G=2, and CAR localised with G=3 have been compared in the context of a small number of areas using both simulation study and dengue case study [10]. They concluded that overall, the Leroux model was preferred under every scenario. To our knowledge, the Bayesian spatial CAR Leroux model in analysing the relative risk of Covid-19 has not been explored yet. This paper aims to investigate the most suitable Bayesian Spatial CAR models in modelling the relative risk of confirmed Covid-19 cases in Makassar, Indonesia.

METHODS

Study Area

Makassar, the capital city of South Sulawesi Province, is situated between 5^0 8'6'19" South Latitude and 119^0 24'17'38" East Longitude. It covers an area of 175.77 km square and has 15 districts namely Mariso, Mamajang, Tamalate, Rappocini, Makassar, Ujung Pandang, Wajo, Bontoala, Ujung Tanah, Tallo, Panakkukang, Manggala, Biringkanaya, Tamalanrea,, and Sangkarrang Island [11]. Makassar has about 1,423,877 population ranging from 14,125 to 209,048 in 2020.

Covid-19 Data

Data on the number of confirmed cases of Covid-19 from 19 March 2020 to 30 August 2021 for each district were used in this study. These data were obtained from the official website of Makassar Health Office (https://infocorona.makassar.go.id/) [4]. Population data in each district were gathered from the Badan Pusat Statistik (BPS) [11].

Models

Standardised Incidence Ratio (SIR) is one measure of disease risk which is defined as the ratio of the number of observed counts (y_i) to the expected counts (E_i). The expected value is here defined as a total number of cases times the population of the area *i* divided by total population in all areas. SIR may be insufficient when the area has a small population and the diseases are rare. Bayesian disease models are recommended to estimate the risk of diseases as they allow to include information from surrounding areas as well as some covariates in the model.

The Bayesian spatial CAR Leroux model proposed by Leroux et al. [12] was used in estimating the number of confirmed cases of Covid-19 in Makassar, Indonesia. This Leroux model consists of a single spatial random effect, u_i, that allows varying strengths of spatial autocorrelation by a constant ρ. The Covid-19 counts were analysed using the CARBayes package version 5.2.3 [13] in the software R version 3.6.1 [14] modelled based on the Poisson distribution. The model is given as follows:

$$\begin{aligned} y_i &\sim \text{Poisson}(E_i\theta_i) \\ \log(\theta_i) &= \alpha + u_i \end{aligned} \tag{1}$$

where y_i and E_i are the number of confirmed cases of Covid-19 and the expected number of confirmed cases of Covid-19 in area *i*, respectively. θ_i is the relative risk of Covid-19 in the *i*th area and α is an overall constant rate. The relative risk of more than one means that the number of cases is higher than the expected number of cases (high risk). Conversely, the relative risk of less than one means that the number of cases is lower than the expected number of cases (low risk).

u_i is the spatial random effect which is modelled by a conditional autoregressive (CAR) prior.

$$\left(u_i \middle| \boldsymbol{u}_j, i \neq j, \tau_u^2\right) \sim N\left(\frac{\rho \sum_j u_j w_{ij}}{\rho \sum_j w_{ij} + 1 - \rho}, \frac{\tau_u^2}{\rho \sum_j w_{ij} + 1 - \rho}\right) \tag{2}$$

The spatial neighboring matrix can be defined in several ways. However, the binary weight matrix is the simplest way and is defined as follows [15]:

$$w_{ij} = \begin{cases} 1 & \text{if areas } i \text{ and } j \text{ are categorised as neighbours} \\ 0 & \text{otherwise.} \end{cases}$$

$w_{ij} = w_{ji}$ and $w_{ii} = 0$.

Different hyperpriors on the precision term τ_u^2 were used namely Inverse-Gamma(1, 0.01), which is the default hyperprior specification in CARBayes, Inverse-Gamma(1, 0.1), Inverse-Gamma(0.1, 0.01) and Inverse-Gamma(0.5, 0.005).

All models were analysed using CARBayes package version 5.2.3 [13] in the software R version 3.6.1 [14]. The posterior distribution for the parameter is obtained by Markov Chain Monte Carlo (MCMC) algorithms and was based on 25,000 iterations with 15,000 MCMC samples after a burn-in of 10,000 samples. The visualisation of the trace and density plots of a subsample was used to check the convergence of each model parameter. The selection of the best model is based on several criteria, namely Deviance Information Criteria (DIC) [16], Watanabe Akaike Information Criteria (WAIC) [17], and residuals from Moran's *I* Modification (MMI).

The Moran's *I* statistics [18] and Modified Moran's *I* [19] for the observed data were also calculated to detect the spatial autocorrelation. A detailed explanation regarding the formula and the R program for both Moran's *I* and MMI has been provided [20].

To investigate the influence of the priors on the posterior estimation, a sensitivity analysis was performed. Four distinct options for the prior on the variance terms were used namely Inverse-Gamma(1, 0.01) which is the default hyperprior specification in CARBayes, Inverse-Gamma(1, 0.1), Inverse-Gamma(0.1, 0.01), and Inverse-Gamma(0.5, 0.005).

RESULTS

The number of confirmed Covid-19 cases in Makassar (19 March 2020-30 August 2021) shows that a total of 44,803 confirmed Covid-19 cases were reported with a mean (2,987), median (2,010), and variance (4,918,042) respectively. Rappocini has the highest number of confirmed Covid-19 cases (6,602) followed by Biringkanaya (6,166) and Tamalate (5,497). On the other hand, Sangkarrang Island has the lowest number of confirmed Covid-19 cases (38) followed by Ujung Tanah (639). Makassar has approximately 1,423,877 population ranging from 14,125 (Sangkarrang) to 209,048 (Biringkanaya) in 2020. Moran's *I* value for observed data is 0.46 with p-value = 0.0009 and MMI value is 0.42. This indicates that there is spatial autocorrelation in data.

The results of DIC and WAIC using Bayesian Spatial CAR Leroux models with four distinct hyperpriors on the precision terms (τ_u^2) are given in Table 1 and depict the insensitivity to the hyperprior.

TABLE 1. *DIC, and WAIC for CAR Leroux Model with Different Hyperpriors.*

Hyperprior (τ_u^2)	DIC	WAIC	Residual using MMI
Inverse-Gamma (0.5,0.05)	176.95	176.50	0.17
Inverse Gamma (0.1,0.01)	177.18	177.11	-0.13
Inverse Gamma (1,0.1)	176.95	177.22	-0.37
Inverse Gamma (1,0.01)	177.70	179.68	-0.11

From Table 1, it concludes that CAR Leroux with hyperprior Inverse-Gamma (0.5;0.05) has the lowest DIC (176.95) and WAIC (176.50). Overall, based on the criteria used, the best model to explain the relative risk of Covid-19 cases is a CAR Leroux model with hyperprior Inverse-Gamma (0.5,0.05). By using the best model, the relative risk values for each district were calculated. The number of confirmed Covid-19 cases, the fitted values, and the relative risk (RR) values for each district with a hyperprior Inverse-Gamma (0.5,0.05) are given in Table 2.

TABLE 2. *The Number of Confirmed Covid-19 Cases and the Relative Risk Value for each istrict*

ID	Districts	Number of Covid-19 Cases	Fitted Values	RR
1	Mamajang	1884	1884.78	1.07
2	Manggala	5042	5040.07	1.09
3	Mariso	1648	1648.73	0.90
4	Sangkarrang	38	43.26	0.09
5	Rappocini	6602	6602.48	1.45
6	Tamalate	5497	5497.47	0.96
7	Makassar	2010	2011.91	0.78
8	Ujung Pandang	1323	1321.19	1.71
9	Panakukkang	5064	5057.94	1.15
10	Bontoala	1184	1183.47	0.68
11	Wajo	1106	1103.32	1.16
12	Ujung Tanah	639	639.27	0.56
13	Tallo	2093	2094.76	0.45
14	Tamalanrea	4507	4512.05	1.39
15	Biringkanya	6166	6172.58	0.94

Based on Table 2, the relative risk in Ujung Pandang was the highest (1.71), followed by Tamalanrea (1.39) and Rappocini (1.45). On the other hand, Sangkarrang has the lowest relative risk followed by Tallo (0.56) and Ujung Tanah (0.45). The relative risk map of confirmed Covid-19 using a Bayesian spatial CAR Leroux model with hyperprior Inverse-Gamma (0.5,0.05) is given in Figure 1.

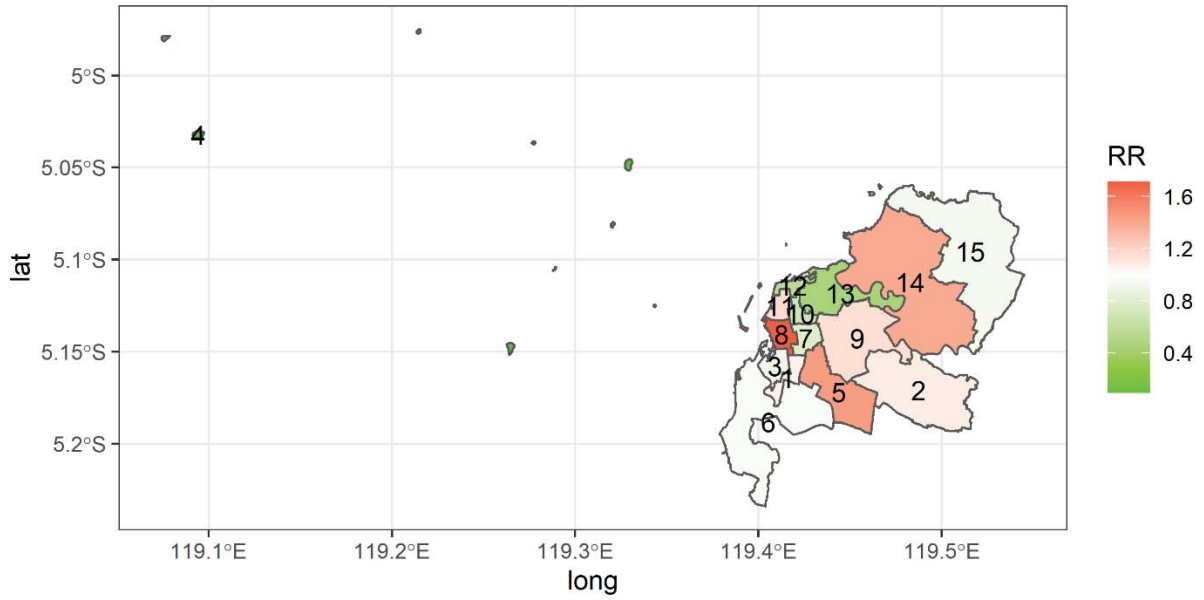

FIGURE 1. *The map of Relative Risk of Confirmed Covid-19 in Makassar for Each District.*

From Figure 1, districts with a higher than average risk (RR > 1) were Ujung Pandang, Rappocini, Tamalanrea, Wajo, Panakukkang, Manggala and Mamajang, while districts with a lower than average risk (RR < 1) were Sangkarrang, Tallo, Ujung Tanah, Bontoala, Makassar, Mariso, Biringkanya, and Tamalate. Ujung Pandang has the highest relative risk of Covid-19. This may be because Ujung Pandang is the capital city of Makassar as well as the administrative center of Makassar city [11] so that the population mobility is high. However, further investigation is needed. The association between population mobility within a sub-city area and Covid-19 cases in Latin American cities has been conducted [21]. They found that there is a strong positive correlation between changes in population movement within a sub-city area and subsequent Covid-19 cases. [21]. On the other hand, Sangkaraang island has the

lowest relative risk of Covid-19. This may be because Sangkarrang Island is situated separately and is not directly adjacent to any other districts. Sangkarang Island district is a division of Ujung Tanah district and the distance from Ujung Pandang is about 20 km.

CONCLUSION

In conclusion, our results suggest the Bayesian spatial CAR Leroux model with hyperprior Inverse-Gamma (0.5;0.05) is more appropriate to model the confirmed Covid-19 cases in Makassar city. Ujung Pandang has the highest relative risk of Covid-19, followed by Rappocini and Tamalanrea, while Sangkarrang Island has the lowest relative risk of Covid-19, followed by Tallo and Ujung Tanah. No covariates were included in this research. It is acknowledged that including covariates such as population mobility, density population, and climatic factors may affect the results and could be possible future work.

ACKNOWLEDGMENTS

The authors wish to acknowledge the Universitas Negeri Makassar. This work was funded by DIPA Universitas Negeri Makassar Number: SP DIPA 023.17.2.677523/2021 November 2020, according to the Decree of the Rector of Universitas Negeri Makassar Number: 550/UN36/HK/2021 on May 05, 2021.
Dr. Susanna Cramb is also acknowledged for helping with editing the manuscripts.

REFERENCES

1. Shereen M A, Khan S, Kazmi A, Bashir N and Siddique R 2020 *Journal of advanced research* **24** 91-98
2. Worldometer 2021 Covid-19 corona virus pandemic.
3. WHO 2021 Weekly Epidemiological and Weekly Operational Updates February 2021.
4. Dinkes K M 2021 Info Penanggulangan Covid-19 Kota Makassar.
5. CNN I 2020 Sulsel Jadi Wilayah Kasus Corona Tertinggi di Luar Pulau Jawa.
6. Whittle R S and Diaz-Artiles A 2020 *BMC medicine* **18** 271-271
7. DiMaggio C, Klein M, Berry C and Frangos S 2020 *Annals of epidemiology* **51** 7-13
8. Konstantinoudis G, Padellini T, Bennett J, Davies B, Ezzati M and Blangiardo M 2021 *Environment international* **146** 106316
9. Lee D 2011 *Spat. Spatio-temporal Epidemiol.* **2** 79-89
10. Aswi A, Cramb S, Duncan E and Mengersen K 2020 *Int. J. Health Geogr.* **19** 39-39
11. Statistik B P 2021 Makassar Municipality in Figures 2021. (Makassar: BPS)
12. Leroux B G, Lei X and Breslow N 2000 *Statistical Models in Epidemiology, the Environment, and Clinical Trials* **116** 179-191
13. Lee D 2013 *J. Stat. Softw.* **55** 1-24
14. R Core Team 2019 R: A language and environment for statistical computing. (Vienna, Austria: R Foundation for Statistical Computing)
15. Waller L A and Gotway C A 2004 *Applied spatial statistics for public health data* (Hoboken, N.J: John Wiley & Sons)
16. Spiegelhalter D J, Best N G, Carlin B P and Van Der Linde A 2002 *Journal of the Royal Statistical Society. Series B, Statistical methodology* **64** 583-639
17. Watanabe S 2010 *J. Mach. Learn. Res.* **11** 3571-3594
18. Moran P A P 1950 *Biometrika* **37** 17
19. Carrijo T B and Da Silva A R 2017 *Geogr. Anal.* **49** 451-467
20. Aswi A, Cramb S, Duncan E and Mengersen K 2021 *Journal of physics. Conference series* **1899** 12098
21. Josiah L Kephart, Xavier Delclòs-Alió, Daniel A Rodríguez, Olga L Sarmiento, Tonatiuh Barrientos-Gutiérrez, Manuel Ramirez-Zea, D Alex Quistberg, Usama Bilal and Roux A V D 2021 *The Lancet. Digital health*