Media Statistika

p-ISSN 1979 – 3693 e-ISSN 2477 – 0647
MEDIA STATISTIKA 15(1) 2022: 48-59
http://ejournal.undip.ac.id/index.php/media_statistika

THE INTERPLAY BETWEEN CLUSTERS, COVARIATES, AND SPATIAL PRIORS IN SPATIAL MODELLING OF COVID-19 IN SOUTH SULAWESI PROVINCE, INDONESIA

Aswi Aswi[1], Muhammad Arif Tiro[1], Sudarmin[1], Sukarna[2], Susanna Cramb[3]
[1] Statistics Department, Universitas Negeri Makassar, Indonesia
[2] Mathematics Department, Universitas Negeri Makassar, Indonesia
[3] Australian Centre for Health Services Innovation & Centre for Healthcare Transformation, Queensland University of Technology, Australia

e-mail: aswi@unm.ac.id

DOI: 10.14710/medstat.15.1.48-59

Article Info:
Received: 9 July 2022
Accepted: 10 July 2022
Available Online: 27 July 2022

Keywords:
Bayesian CAR Localised; Clustering; Covid-19; Relative Risk
.

Abstract: A number of previous studies on Covid-19 have used Bayesian spatial Conditional Autoregressive (CAR) models. However, basic CAR models are at risk of over-smoothing if adjacent areas genuinely differ in risk. More complex forms, such as localised CAR models, allow for sudden disparities, but have rarely been applied to modelling Covid-19, and never with covariates. This study aims to evaluate the most suitable Bayesian spatial CAR localised models in modelling the number of Covid-19 cases with and without covariates, examine the impact of covariates and spatial priors on the identified clusters and which factors affect the Covid-19 risk in South Sulawesi Province. Data on the number of confirmed cases of Covid-19 (19 March 2020 -25 February 2022) were analyzed using the Bayesian spatial CAR localised model with a different number of clusters and priors. The results show that the Bayesian spatial CAR localised model with population density included fits the data better than a corresponding model without covariates. There was a positive correlation between the Covid-19 risk and population density. The interplay between covariates, spatial priors, and clustering structure influenced the performance of models. Makassar city and Bone have the highest and the lowest relative risk (RR) of Covid-19 respectively.

1. INTRODUCTION

The spread of the Coronavirus disease 2019 (Covid-19) epidemic continues to increase worldwide. On March 2nd, 2020, the first Covid-19 case was reported in Indonesia. Within two years, by February 25th, Indonesia had recorded 5,457,775 confirmed Covid-19 cases, 4,736,234 people recovered, and 147,586 deaths (http://covid19.go.id/). In South Sulawesi province, a total of 131,826 confirmed Covid-19 cases were reported in this period.

A number of previous studies have modelled Covid-19 using a Bayesian spatial Conditional Autoregressive (CAR) considering socioeconomic influences. The association of socioeconomic, and environmental variables with the incidence of Covid-19 in the 30 provinces in mainland China has been identified using a Bayesian spatial CAR Besag, York

& Mollié (BYM) model (Peng et al., 2022). Their results suggested that the risk of Covid-19 was positively correlated with the economic development level and population movements. Whittle and Diaz-Artiles (2020) also used a Bayesian spatial CAR BYM to assess the relationship between the number of Covid-19 cases in New York City and socioeconomic factors and found significant associations with population density, race, and household income. A study in Columbia used a Bayesian spatial CAR Leroux model to model Covid-19 considering the index of multidimensional poverty, which significantly influenced the inequalities risk of dying from Covid-19 (Polo, Soler-Tovar, Villamil Jimenez, Benavides-Ortiz, & Mera Acosta, 2022).

Bayesian spatial CAR models have also been used to examine population density and distance to cities in several places. A Bayesian spatial CAR Leroux examined these factors in South Sulawesi Province and found that population density was positively correlated with the risk of Covid-19 (Aswi & Sukarna, 2022). A Bayesian spatial CAR Leroux has also been used to examine the association of population density and the distance to the city with the risk of Covid-19 in Makassar city (Tiro, Aswi, & Rais, 2021). However, since Makassar city covers only a small area, the results differed. They concluded that population density was not statistically correlated with the relative risk of Covid-19, but the distance to the capital city was negatively correlated with the risk of Covid-19. Another study has investigated the association between population density, distance from the virus epicenter and the Covid-19 incidence in Iranian Provinces by using linear regression analysis (Dadar, Fakhri, Bjørklund, & Shahali, 2020). Their results showed that the association between population density and the Covid-19 incidence was not statistically significant, but the incidence of Covid-19 was strongly negatively associated with the distance from the virus epicenter. The results regarding the association of population density and Covid-19 are location-specific.

While a number of studies have evaluated the effect of population density and distance to the capital city, to our knowledge, the Bayesian spatial CAR localised model has not been used in modeling Covid-19 considering covariates. The Bayesian spatial CAR localised model has been explored in modelling Covid-19 risk in South Sulawesi Province (Aswi, Mauliyana, Tiro, & Bustan, 2022), but it did not include covariates in the model. This study aims to evaluate the most suitable Bayesian spatial CAR localised models in modelling the number of Covid-19 cases with and without covariates (the distance to the capital city and population density) in South Sulawesi Province and examine the impact of covariates and spatial priors on the identified clusters. Factors that affect the risk of Covid-19 in South Sulawesi Province also were identified.

2. LITERATURE REVIEW

2.1. Spatial Dependence

Moran's *I* is the common indicator used to measure the degree of spatial dependence (Moran, 1950) in ordinal and interval data. Moran's *I* can be used on raw data or fitted counts to check for spatial autocorrelation, and on residuals from a spatial model to check model goodness of fit (if the model accounted for the spatial structure appropriately then Moran's *I* should be close to 0).

Moran's *I* is computed as the ratio of spatial covariation to the total variation where values range from -1 to +1. The positive value indicates positive spatial dependence, while the negative value indicates negative spatial dependence, and the 0 value indicates no spatial dependence.

Moran's I statistics are calculated as follows:

$$I = \frac{n\sum_{i=1}^{n}\sum_{k=1}^{n}\omega_{ik}(Y_i - \bar{Y})(Y_k - \bar{Y})}{\sum_{i=1}^{n}\sum_{k=1}^{n}\omega_{ik}(Y_i - \bar{Y})^2}$$

where n is the number of locations, Y_i and Y_k are the observed value in the particular location i and another location k, $\bar{X}$ is the average of all the X values over the n locations, ω_{ik} is the spatial connectivity/weight matrix.

As Moran's I tend to underestimate spatial autocorrelation when there are less than 100 areas, a modified Moran's I (MMI) was developed (Carrijo & Da Silva, 2017) to detect spatial dependence which works even for a few areas. MMI statistic is calculated as follows:

$$I_{\text{Mod}} = \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})(\sum_{k=1}^{n}w_{ik}Y_k - \bar{Y})}{[\sum_{i=1}^{n}(Y_i - \bar{Y})^2]^{1/2}\left[\sum_{i=1}^{n}\left(\sum_{k=1}^{n}w_{ik}Y_k - \bar{Y}\right)^2\right]^{1/2}}$$

A detailed explanation of MMI can be seen in some works of literature (Aswi, Cramb, Duncan, & Mengersen, 2021; Carrijo & Da Silva, 2017).

2.2. Relative Risk

The Standardised Incidence Ratio (SIR) is calculated by dividing the number of Covid-19 (y_i) cases with the number of expected cases in each area (E_i). The expected number of Covid-19 cases is here calculated as the overall incidence rate for the entire South Sulawesi Province multiplied by the population at risk in each location (pop_i) and it is given as follows:

$$E_i = \frac{\sum_i y_i}{\sum_i pop_i} pop_i$$

Usually, this would be calculated by age groups and summed together, but data by age were not available.

In estimating the relative risk (RR) across small areas, Bayesian methods such as Bayesian hierarchical models are preferred rather than raw SIRs as they are able to incorporate information from neighboring locations through prior distributions as well as adjust for covariates in the model.

2.3. Bayesian Hierarchical Models

Bayesian Hierarchical models include random-effects models, multilevel models, and generalised linear (mixed) models (Ntzoufras, 2011). In a hierarchical Bayesian model, the parameter distributions are conditional on parameters existing at the next level of the hierarchy. Several advantages of the hierarchical model are (1) they enable increased robustness of Bayes estimators (Robert, 2007); (2) They allow for sharing of information across experimental units (Gelman, 2013), and (3) they enable easy construction of complicated models (Mugglin, Cressie, & Gemmell, 2002).

3. MATERIAL AND METHOD

3.1. Study Area

South Sulawesi Province is located between 0°12' and 8° South latitude, and between 116°48' and 112°36' East longitude. It consists of 24 districts, three of which are cities: Makassar (capital city), Palopo, and Pare-Pare. The area of South Sulawesi Province is 46717.48 km^2 with a population of 9.074 million in 2020, equating to an average population density of 640.88 people per km^2 (Badan Pusat Statistik, 2021). Makassar city has the highest population density (8100.80 people/km^2), while Luwu Timur has the lowest population density of 42.73 people/km^2 (Badan Pusat Statistik, 2021). Luwu Timur has the longest distance to Makassar city (565 km) while Makassar has the shortest distance (0 km), followed by Gowa (11 km) (Badan Pusat Statistik, 2015).

3.2. Data

Data on the number of confirmed cases of Covid-19 (19 March 2020 -25 February 2022) for each of the 24 districts used in this study was obtained from the official website "Ministry of Health of the Republic of Indonesia" https://infeksiemerging.kemkes.go.id/. Data on population numbers were from the Badan Pusat Statistik (Badan Pusat Statistik, 2021) and used to calculate the expected counts. Population density data in each district (Badan Pusat Statistik, 2021) and the distance to the capital city of South Sulawesi Province (Badan Pusat Statistik, 2015) were used as covariates.

3.3. Model Formulation

In this paper, the Bayesian spatial CAR localised model (Lee & Sarran, 2015) was used to estimate the risk of Covid-19 and examine the clusters of Covid-19 cases with and without covariates. The population density of each district is calculated as the ratio of the number of populations in each area to the corresponding area.

The Bayesian spatial CAR localised model has two key components, namely a spatial random effect (u_i) and the clustering components (λ_{Z_i}) which enables neighbourhood random effects different over locations. The locations are partitioned into a pre-specified maximum of G clusters and include a cluster-specific mean in the model. A Poisson log-linear model, as is commonly used for mapping the relative risk of diseases (Aswi, Cramb, Moraga, & Mengersen, 2019) was used to model the number of confirmed Covid-19 cases (y_i) as follows:

$$y_i \sim \text{Poisson}(E_i\theta_i) \text{ for i = 1, 2, 3, …, 24 locations}$$

$$log(\theta_i) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + u_i + \lambda z_i$$

where E_i is the number of expected cases, and θ_i is the relative risk in the i^{th} location. β_0 is the overall level of RR, while β_1 and β_2 are the covariate coefficients. The spatial structured random effect is modelled using an intrinsic conditional autoregressive prior as follows:

$$(u_i|u_k, i \neq k, \tau_u^2) \sim N\left(\frac{\sum_k u_k\omega_{ik}}{\sum_k \omega_{ik}}, \frac{\tau_u^2}{\sum_k \omega_{ik}}\right)$$

where ω_{ik} is the spatial weight matrix defined using binary spatial matrix and first-order adjacency weight matrix. A sensitivity analysis was performed by using four hyperpriors on the variance component τ_u^2 namely: Inverse-Gamma IG (1, 0.01) as the default hyperprior of CARBayes, IG (1, 0.1), IG (0.5, 0.0005), and IG (0.001, 0.001).

The locations are partitioned into maximum G clusters and include a cluster-specific mean in the model. Cluster G has its own form of intercept which is ordered as $\lambda_1 < \lambda_2 < \ldots < \lambda_G$.

$$\lambda_k \sim \text{Uniform}(\lambda_{k-1}, \lambda_{k+1}) \text{ for } k = 1, 2, \ldots, G$$

where $\lambda_0 = -\infty$ and $\lambda_{G+1} = +\infty$

A variable Z_i assigns the allocation of the i^{th} location to a cluster,

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

where $\delta \sim$ Uniform (1, 10);
If G is odd then $G^* = \frac{G+1}{2}$, while if G is even then $G^* = \frac{G}{2}$. It is recommended that G is small and odd numbers (Lee & Sarran, 2015).

A set of combination model formulations were used to examine the interplay between the number of clusters, covariates, and spatial priors. We choose different spatial CAR localised models allowing a maximum of two, three, and five clusters with and without covariates included, using different priors.

All analyses were conducted in R software version 4.1.2 (R Core Team, 2019) using the CARBayes package version 5.2.5 (Lee, 2013) to estimate model parameters. We generated Markov chain Monte Carlo (MCMC) samples based on 20,000 iterations with 12,000 MCMC samples collected after a burn-in of 8,000 samples. A visualization of MCMC trace and density plots was performed to check the MCMC convergence.

The goodness-of-fit of model formulation and combination of covariates were compared using the Deviance Information Criterion (DIC) (Spiegelhalter, Best, Carlin, & Van Der Linde, 2002), Watanabe Akaike Information Criterion (WAIC) (Watanabe, 2010), MMI (Aswi et al., 2021; Carrijo & Da Silva, 2017) for the residuals, and considering whether the 95% posterior credible interval contains zero. A smaller value of DIC, WAIC, and MMI for residuals indicates a better model fit. R code used in this study is available upon request.

4. RESULTS AND DISCUSSION

4.1. Descriptive Analysis

A total of 131,826 positive confirmed Covid-19 cases in the Province of South Sulawesi (March 19, 2020- February 25, 2022) were identified with a mean (5,057), median (2,522), and variance (121,218,441). The lowest numbers of confirmed Covid-19 cases were in Enrekang (805 cases), Selayar (1,303 cases), and Bantaeng (1,331 cases). In contrast, the highest numbers of confirmed Covid-19 cases were in Makassar (56,041 cases), Gowa (9,560 cases), and Luwu Timur (5,124 cases).

The values of Moran's *I* statistics, expectation and variance for observed data are 0.056, -0.043, and 0.002, respectively with Z-score = 2.089 and p-value = 0.018. Given a Moran's *I* value of observed data of 0.056 with p-value=0.018, the null hypothesis stating no spatial autocorrelation is rejected. This indicates that the areal pattern for Covid-19 cases is statistically significant with a positive spatial autocorrelation. MMI value is 0.042.

4.2. Bayesian Spatial CAR Localised models

The Bayesian spatial CAR localised models with G=2, G=3, and G=5 with four different hyperpriors were used in this study. The DIC, WAIC, MMI for residuals, posterior quantities for covariates as well as the number of areas included in each cluster for Bayesian spatial CAR localised models with G=2, G=3, and G=5 are provided in Tables 1, 2, and 3, respectively.

Table 1 shows that the Bayesian spatial CAR localised models with G=2 with the inclusion of population density have the smallest values of DIC and WAIC (Model M3, M7, M11, M15) for all four hyperpriors. The number of areas included in each cluster (G1 and G2) for different model combinations differed (M1, M2, M3, and M4) but the clustering structure for different hyperpriors was the same (M1=M5=M9=M13; M2=M6=M10=M14; M3=M7=M11=M15; M4=M8=M12=M16). The model with the inclusion of distance to the capital province was considered significantly associated with the risk of Covid-19 as the 95% posterior Credible Interval (CI) does not contain zero (M2, M6, M10, M14) and it indicates that the distance to the capital province was negatively significant associated with Covid-19 risk. Furthermore, the population density was shown to be positively and significantly associated with the risk of Covid-19 (M3, M7, M11, M15). However, the inclusion of both covariates causes the relationship between the distance to the capital province and Covid-19 risk to change sign (positive).

Table 1. The DIC, WAIC, MMI for residuals, posterior quantities for covariates, and the number of areas included in each cluster for G=2.

G = 2									
Hyperpriors	Models		DIC	WAIC	MMI residual	Posterior Quantities for Covariates		Number of areas in the cluster	
						2.5%	97.5%	G1	G2
IG(1, 0.01)	M1	Without Covariate	290.59	300.83	-0.40	-	-	11	13
	M2	Distance*	294.88	324.10	-0.29	-1.01	-0.69	17	7
	M3	Density*	286.98	285.82	-0.66	0.16	0.24	10	14
	M4	Distance*+ Density*	288.03	287.23	-0.45	0.06 0.23	0.20 0.31	10	14
IG(1, 0.1)	M5	Without Covariate	289.95	304.85	-0.50	-	-	11	13
	M6	Distance*	293.89	313.20	-0.34	-1.08	-0.75	17	7
	M7	Density*	287.02	285.52	-0.35	0.17	0.29	10	14
	M8	Distance*+ Density*	287.66	287.46	-0.19	0.10 0.16	0.29 0.25	10	14
IG(0.5, 0.0005)	M9	Without Covariate	290.83	297.99	-0.21	-	-	11	13
	M10	Distance*	294.23	331.19	-0.35	-0.98	-0.68	17	7
	M11	Density*	287.59	286.22	-0.51	0.17	0.27	10	14
	M12	Distance*+ Density*	287.88	287.50	-0.39	0.08 0.26	0.27 0.35	10	14
IG(0.001, 0.001)	M13	Without Covariate	292.02	309.67	-0.62	-	-	11	13
	M14	Distance*	295.12	320.20	-0.33	-0.72	-0.46	17	7
	M15	Density*	286.60	284.52	-0.55	0.20	0.29	10	14
	M16	Distance*+ Density*	287.81	287.87	-0.49	0.09 0.21	0.20 0.38	10	14

Table 2 describes the results of the Bayesian spatial CAR localised models with G=3 with and without the inclusion of covariates. The results showed that the model with the inclusion of both covariates with the hyperprior IG (1, 0.01) has the smallest values of DIC (M20). However, the model with the inclusion of population density with the hyperprior IG (0.5, 0.0005) has the smallest values of WAIC as well as the smallest MMI for residual (M27). The number of areas included in each cluster (G1, G2, and G3) for different model combinations are different (M17, M18, M19, M20) but the clustering structure for different hyperpriors was the same except for the model with the inclusion of both covariates (M20 ≠M24 ≠M28 ≠M32). The number of areas for each cluster for Model M20: G1=5, G2=7, and G3=12, while for Model M24 are (G1=4, G2=8, G3=12).

The model which included distance to the capital province found it was negatively and significantly associated with the risk of Covid-19, whereas population density was positively and significantly associated with the risk of Covid-19. These relationships were consistent whether the covariates were alone or in combination.

Table 2. The DIC, WAIC, MMI for residuals, posterior quantities for covariates, and the number of areas included in each cluster for G=3.

Hyperpriors	Models		DIC	WAIC	MMI residual	Posterior Quantities for Covariates		Number of areas in the Cluster		
						2.5%	97.5%	G1	G2	G3
IG(1, 0.01)	M17	Without Covariate	301.62	372.65	-0.20	-	-	10	10	4
	M18	Distance*	301.65	357.32	-0.75	-0.62	-0.44	14	7	3
	M19	Density*	293.73	308.23	-0.37	0.17	0.34	10	8	6
	M20	Distance*+ Density*	226.49	346.50	-0.39	-0.40 0.25	-0.07 0.34	5	7	12
IG(1, 0.1)	M21	Without Covariate	299.17	384.52	-0.27	-	-	10	10	4
	M22	Distance*	299.64	340.95	-0.66	-0.49	-0.28	14	7	3
	M23	Density*	294.04	308.73	-0.31	0.14	0.32	10	8	6
	M24	Distance*+ Density*	281.58	342.08	-0.49	-0.24 0.22	-0.03 0.40	4	8	12
IG(0.5, 0.0005)	M25	Without Covariate	301.06	379.22	-0.62	-	-	10	10	4
	M26	Distance*	300.47	351.59	-0.34	-0.63	-0.34	14	7	3
	M27	Density*	292.21	304.13	-0.22	0.13	0.27	10	8	6
	M28	Distance*+ Density*	252.01	327.00	-0.63	-0.29 0.29	-0.02 0.48	6	6	12
IG(0.001, 0.001)	M29	Without Covariate	299.40	361.25	-0.45	-	-	10	10	4
	M30	Distance*	299.39	339.52	-0.73	-0.68	-0.35	14	7	3
	M31	Density*	292.92	307.93	-0.60	0.23	0.38	10	8	6
	M32	Distance*+ Density*	297.36	356.32	-0.70	-0.20 0.38	-0.02 0.49	5	7	12

Table 3 depicts the results of the Bayesian spatial CAR localised models with G=5 with and without the inclusion of covariates. While all 4 hyperpriors had the lowest DIC when both covariates were included, the lowest WAIC was consistently for the model without covariates, but this model often had relatively high MMI residual values, suggestive of poor fit. The lowest MMI residuals were consistently for the models including population density. While individually the covariates showed the same patterns (negative association with Covid-19 risk for distance, positive for population density), when combined often the distance had a 95% CI which included 0. Only M48 with hyperprior of IG(0.001, 0.001) retained the original negative association for distance.

Table 3. The DIC, WAIC, MMI for residuals, posterior quantities for covariates, and the number of areas included in each cluster for G=5.

G = 5												
Hyper-priors	Models		DIC	WAIC	MMI residual	Posterior Quantities for Covariates		Number of areas in the Cluster				
						2.5%	97.5%	G1	G2	G3	G4	G5
IG(1, 0.01)	M33	Without Covariate	285.23	284.49	-0.72	-	-	5	5	8	5	1
	M34	Distance*	311.31	401.20	-0.27	-0.82	-0.64	8	6	5	4	1
	M35	Density*	300.88	347.23	-0.23	0.26	0.39	3	7	5	8	1
	M36	Distance*+ Density*	162.17	746.84	-0.67	-0.12 0.34	0.02 0.47	1	7	4	8	4
IG(1, 0.1)	M37	Without Covariate	284.83	284.07	-0.75	-	-	5	5	8	5	1
	M38	Distance*	310.75	391.86	-0.36	-0.67	-0.51	8	6	5	4	1
	M39	Density*	277.69	322.57	-0.35	0.13	0.26	3	7	7	6	1
	M40	Distance+ Density*	276.52	403.51	-0.65	-0.05 0.37	0.10 0.45	4	5	3	8	4
IG(0.5, 0.0005)	M41	Without Covariate	285.85	286.35	-0.76	-	-	5	5	8	5	1
	M42	Distance*	310.20	389.06	-0.53	-0.85	-0.63	7	7	5	4	1
	M43	Density*	294.10	310.18	-0.49	0.14	0.28	3	7	7	6	1
	M44	Distance+ Density*	244.13	428.58	-0.62	-0.08 0.37	0.06 0.54	1	3	8	8	4
IG(0.001, 0.001)	M45	Without Covariate	285.53	285.57	-0.69	-	-	5	5	8	5	1
	M46	Distance*	310.29	395.30	-0.19	-0.85	-0.70	5	9	5	4	1
	M47	Density*	298.51	343.62	-0.17	0.23	0.34	3	7	5	8	1
	M48	Distance*+ Density*	-273.5	3018.44	-0.72	-0.13 0.36	-0.02 0.47	1	4	7	8	4

The number of areas included in each cluster (G1, G2, G3, G4, and G5) for different model combinations of covariates is different (M33, M34, M35, M36) and the clustering structure for different hyperpriors was different except for the model without covariates (M33 =M37 =M41 =M45). The number of areas for each cluster for Model M33 =M37 =M41 =M45 are G1=5, G2=5, G3=8, G4=5, and G5=1. The model found distance to the capital province was negatively and significantly associated with the risk of Covid-19. Furthermore, the model with the inclusion of population density was positively and significantly associated with the risk of Covid-19. However, the inclusion of both covariates caused the relationship between the distance to the capital province and Covid-19 risk to not be significant, except in model M48.

Overall, based on all the model selection criteria used in this study, the preferred model to estimate the relative risk of Covid-19 across South Sulawesi Province is a Bayesian spatial CAR localised with hyperprior IG(0.5, 0.0005) model with the inclusion of population density (M27) that allows up to three clusters (G=3). There was a positive association between the Covid-19 risk and the population density. The importance of population density agrees with other studies (Arbel, Fialkoff, Kerner, & Kerner, 2021; Moosa & Khatatbeh, 2021; Sy, White, & Nichols, 2021; Wong & Li, 2020).

Table 4. The localised structure (LS) and RR values for each district are based on the preferred model with population density for G=3 (M27) as well as for G=2 (M7), plus the same without covariates (M25 and M5).

ID	Districts	LS for G=2		RR for M7	LS for G=3		RR for M27
		M5	M7		M25	M27	
1	Barru	2	2	0.72	2	2	0.72
2	Bone	1	1	0.26	1	1	0.26
3	Bulukumba	1	1	0.36	1	1	0.36
4	Enrekang	1	1	0.27	1	1	0.27
5	Gowa	2	2	0.93	2	2	0.93
6	Jeneponto	1	1	0.45	1	1	0.45
7	Luwu Timur	2	2	1.29	2	3	1.29
8	Luwu Utara	2	2	0.79	2	2	0.80
9	Luwu	1	1	0.31	1	1	0.30
10	Makassar	2	2	2.94	3	2	2.94
11	Maros	2	2	0.89	2	3	0.89
12	Palopo	2	2	1.19	3	3	1.19
13	Pangkep	2	2	0.76	2	2	0.76
14	Parepare	2	2	1.40	3	3	1.40
15	Pinrang	1	1	0.32	1	1	0.32
16	Selayar	2	2	0.71	2	2	0.71
17	Sidrap	1	1	0.38	1	1	0.38
18	Sinjai	2	2	1.09	2	3	1.10
19	Soppeng	2	2	0.81	2	2	0.81
20	Takalar	1	2	0.62	2	2	0.62
21	Toraja Utara	1	1	0.39	1	1	0.39
22	Toraja	2	2	1.06	3	3	1.06
23	Wajo	1	1	0.33	1	1	0.32
24	Bantaeng	1	1	0.50	1	1	0.50

Despite differences in localised structure, the RR values were similar for both G=2 and G=3 when population density was included (Table 4). However, the switching of the effect of distance to the capital city when included with population density under only G=2 confirms that even numbers for G are best avoided since switching issues are recognised when numbers are even (Lee & Sarran, 2015). The visualisation of the localised clustering structure of the preferred model with (M27) and without (M25) population density included is presented in Figure 1 and shows slight differences. Districts which changed cluster when covariates were included are Luwu Timur, Makassar, Maros, and Sinjai.

The RR map of confirmed Covid-19 cases based on the Bayesian Spatial CAR localised with hyperprior IG(0.5, 0.0005) with G=3 and population density included (M27) is given in Figure 2, and along with Table 4 shows a far higher risk in Makassar city than elsewhere.

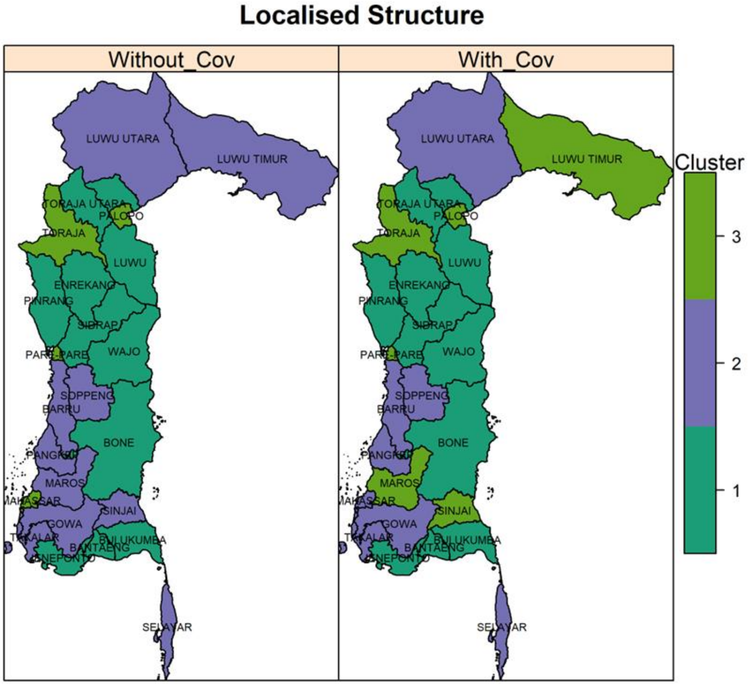

Figure 1. Localised maps were obtained under the Bayesian spatial CAR localised model with G=3 without covariates and with population density as a covariate.

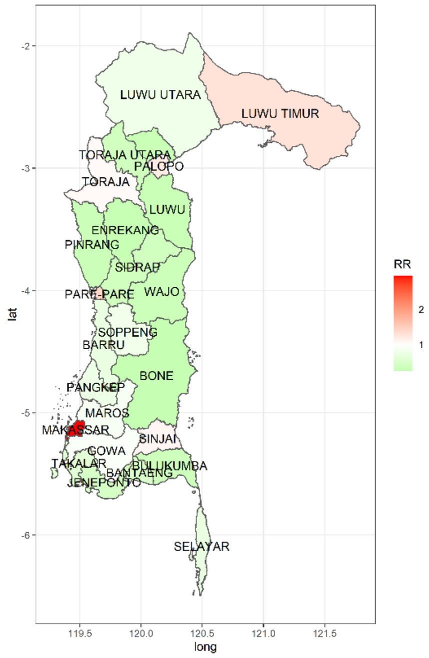

Figure 2. Localised maps were obtained under the Bayesian spatial CAR localised model with G=3 with population density as a covariate.

Overall, the number of areas included in each cluster for different model combinations of covariates is different and the clustering structure for different hyperpriors was also different. Makassar city and Bone have the highest and the lowest RR respectively.

5. CONCLUSION

The inclusion of covariates causes group structure to alter in the localised model. A spatial CAR localised model with three clusters and incorporation of population density provided the best fit. The interplay between covariates, spatial priors, and clustering structure influenced the performance of models for modeling Covid-19 cases. There was a positive correlation between the Covid-19 risk and population density. Using appropriate Bayesian spatial models enables the identification of different clusters of areas and the impact of covariates which may help inform policy decisions for future planning. Makassar city which has the highest population density has the highest RR while Bone has the lowest RR for Covid-19. Considering other covariates could be possible for future work.

ACKNOWLEDGMENT

The authors wish to acknowledge the Universitas Negeri Makassar (UNM), Indonesia. This work was funded by DIPA UNM Number: SP DIPA 023.17.2.677523/2021, 23 November 2021, according to the Decree of the Rector of UNM Number: 508/UN36/HK/2021 on April 28.

REFERENCES

Arbel, Y., Fialkoff, C., Kerner, A., & Kerner, M. (2021). Do Population Density, Socio-Economic Ranking and Gini Index of Cities Influence Infection Rates from Coronavirus? Israel as a Case Study. *The Annals of Regional Science, 68*(1), 181-206. doi:10.1007/s00168-021-01073-y

Aswi, A., Cramb, S., Duncan, E., & Mengersen, K. (2021). Detecting Spatial Autocorrelation for a Small Number of Areas: a Practical Example. *Journal of Physics. Conference Series, 1899*(1), 12098. doi:10.1088/1742-6596/1899/1/012098

Aswi, A., Cramb, S. M., Moraga, P., & Mengersen, K. (2019). Bayesian Spatial and Spatio-Temporal Approaches to Modelling Dengue Fever: a Systematic Review. *Epidemiology and Infection, 147*. doi:10.1017/S0950268818002807

Aswi, A., Mauliyana, A., Tiro, M. A., & Bustan, M. N. (2022). Relative Risk of Coronavirus Disease (Covid-19) In South Sulawesi Province, Indonesia: Bayesian Spatial Modeling. *Media Statistika, 14*(2), 158-169. doi:10.14710/medstat.14.2.158-169

Aswi, A., & Sukarna, S. (2022). Factors Affecting the Covid-19 Risk in South Sulawesi Province, Indonesia: A Bayesian Spatial Model. *Inferensi, 5*(1), 51-58.

Badan Pusat Statistik. (2015). *Sulawesi Selatan dalam Angka 2015*.

Badan Pusat Statistik. (2021). *Sulawesi Selatan dalam Angka 2021*.

Carrijo, T. B., & Da Silva, A. R. (2017). Modified Moran's I for Small Samples. *Geographical Analysis, 49*(4), 451-467. doi:10.1111/gean.12130

Dadar, M., Fakhri, Y., Bjørklund, G., & Shahali, Y. (2020). The Association Between the Incidence of COVID-19 and the Distance from the Virus Epicenter in Iran. *Archives of Virology, 165*(11), 2555. doi:10.1007/s00705-020-04774-5

Gelman, A. (2013). *Bayesian Data Analysis, Third Edition* (3rd ed.). Hoboken: CRC Press.

Lee, D. (2013). CARBayes: an R Package for Bayesian Spatial Modeling with Conditional Autoregressive Priors. *Journal of Statistical Software, 55*(13), 1-24.

Lee, D., & Sarran, C. (2015). Controlling for Unmeasured Confounding and Spatial Misalignment in Long-Term Air Pollution and Health Studies. *Environmetrics, 26*(7), 477-487.

Moosa, I. A., & Khatatbeh, I. N. (2021). The Density Paradox: are Densely-Populated Regions More Vulnerable to Covid-19? *The International Journal of Health Planning and Management, 36*(5), 1575-1588. doi:10.1002/hpm.3189

Moran, P. A. P. (1950). Notes on Continuous Stochastic Phenomena. *Biometrika, 37*(1-2), 17. doi:10.1093/biomet/37.1-2.17

Mugglin, A. S., Cressie, N., & Gemmell, I. (2002). Hierarchical Statistical Modelling of Influenza Epidemic Dynamics in Space and Time. *Statistics in Medicine, 21*(18), 2703-2721. doi:10.1002/sim.1217

Ntzoufras, I. (2011). *Bayesian Modeling Using WinBUGS* (Vol. 698): John Wiley & Sons.

Peng, D., Qian, J., Wei, L., Luo, C., Zhang, T., Zhou, L., Yin, F. (2022). COVID-19 Distributes Socially in China: A Bayesian Spatial Analysis. *PLoS ONE, 17*(4), e0267001-e0267001. doi:10.1371/journal.pone.0267001

Polo, G., Soler-Tovar, D., Villamil Jimenez, L. C., Benavides-Ortiz, E., & Mera Acosta, C. (2022). Bayesian Spatial Modeling of COVID-19 Case-Fatality Rate Inequalities. *Spatial and Spatio-temporal Epidemiology, 41*. doi:10.1016/j.sste.2022.100494

R Core Team. (2019). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org

Robert, C. (2007). *The Bayesian Choice From Decision-Theoretic Foundations to Computational Implementation* (2nd ed. ed.). New York, NY: Springer New York.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van Der Linde, A. (2002). Bayesian Measures of Model Complexity and Fit. *Journal of the Royal Statistical Society. Series B, Statistical methodology, 64*(4), 583-639. doi:10.1111/1467-9868.00353

Sy, K. T. L., White, L. F., & Nichols, B. E. (2021). Population Density and Basic Reproductive Number of COVID-19 Across United States Counties. *PLoS ONE, 16*(4), e0249271-e0249271. doi:10.1371/journal.pone.0249271

Tiro, M. A., Aswi, A., & Rais, Z. (2021). Association of Population Density and Distance to the City with the Risks of COVID-19: A Bayesian Spatial Analysis. *Journal of physics. Conference series, 2123*(1), 12001. doi:10.1088/1742-6596/2123/1/012001

Watanabe, S. (2010). Asymptotic Equivalence of Bayes Cross Validation and Widely Applicable Information Criterion in Singular Learning Theory. *Journal of Machine Learning Research, 11*, 3571-3594.

Whittle, R. S., & Diaz-Artiles, A. (2020). An Ecological Study of Socioeconomic Predictors in Detection of COVID-19 Cases Across Neighborhoods in New York City. *BMC medicine, 18*(1), 271-271. doi:10.1186/s12916-020-01731-6

Wong, D. W. S., & Li, Y. (2020). Spreading of COVID-19: Density Matters. *PLoS ONE, 15*(12), e0242398-e0242398. doi:10.1371/journal.pone.0242398