Available online at http://scik.org
Commun. Math. Biol. Neurosci. 2023, 2023:28
https://doi.org/10.28919/cmbn/7898
ISSN: 2052-2541

SPATIAL CLUSTERING OF STUNTING CASES IN INDONESIA: A BAYESIAN APPROACH

ANDI ASMAWATI AZIS[1], ASWI ASWI[2,*]

[1]Department of Biology, Universitas Negeri Makassar, Makassar, Indonesia

[2]Department of Statistics, Universitas Negeri Makassar, Makassar, Indonesia

Abstract: Stunting is one of the major public health problems, especially in developing countries such as Indonesia. In the Southeast Asian region, Indonesia has the third highest prevalence with an average prevalence of 36.4% (2005-2017). Indonesian Nutritional Status Study (SSGI) in 2021 stated that the percentage of stunting in Indonesia is 24.4%. Research on spatial modelling of stunting has been done, but the use of the Bayesian spatial Conditional Autoregressive (CAR) model is still rare. This article aims to provide the most appropriate Bayesian spatial CAR localised (clustering) model and identify the relative risk (RR) of stunting in each province in Indonesia. Data on the number of toddlers 0-59 months whose height is measured and the number of stunted toddlers in each province in Indonesia in 2021 were used. The best model is based on the Deviance Information Criterion, Watanabe Akaike Information Criterion and Modified Moran's I value for residuals. The results indicated that the Bayesian spatial CAR Localised with hyperprior Inverse-Gamma (0.5, 0.05) and Inverse-Gamma (1, 0.1) are preferred for two and three clusters, respectively. Our results identified the high-risk areas for stunting. Approximately 56% of provinces in Indonesia are at a high risk of stunting. Sulawesi Barat has the highest RR for stunting followed by Nusa Tenggara Timur dan Papua Barat. In contrast, Jakarta has the lowest RR of stunting followed by Sulawesi Utara and Sumatera Selatan. Government should pay more attention to areas that are most at high risk of stunting.

Keywords: stunting; Bayesian method; relative risk; CAR localised.

2020 AMS Subject Classification: 62P10.

---

*Corresponding author

E-mail address: aswi@unm.ac.id

Received January 30, 2023

1. INTRODUCTION

Stunting is a problem of chronic malnutrition in children which is influenced by various factors. The height of a stunted child is shorter than the normal height of children of their age. Likewise, the brain of a stunted child will likely never develop to normal cognitive potential. Globally, approximately 144 million children under five are stunted. It is stated by the World Health Organization (WHO) that the average prevalence of stunting in Indonesia from 2005-2017 is 36.4% which exceeds the WHO threshold (30%) [1]. In 2017, the prevalence of stunted toddlers aged 0-59 months in Indonesia was 29.6%. Indonesian Nutritional Status Study (SSGI) in 2021 stated that the percentage of stunting in Indonesia is 24.4%. One indicator to determine whether a child is stunted or normal is to look at height for age. The number of toddlers 0-59 months whose height was measured in Indonesia in 2021 was 24958590 with the number of stunting cases being 2373712 (9.51%).

Some research on spatial modelling of stunting has been done [2-7], but the use of spatial analysis using the Bayesian method is still rare. The risk of stunting in South Sulawesi Province has been investigated by using the Bayesian spatial Conditional Autoregressive (BSCAR) Leroux model and found that the best model for modeling the relative risk (RR) cases of stunting under five in 2019 is the BSCAR Leroux model with hyperprior Inverse Gamma IG (0.5; 0.0005). They found that Toraja district and Pare-Pare have the highest RR (2.02) of stunting followed by Enrekang district (1.94). In contrast, Gowa district has the lowest RR (0.46) of stunting followed by Makassar City (0.52), and Pinrang district (0.67) [8]. Another research on modelling stunting cases used the BSCAR Leroux model in South Sulawesi Province in 2020 involving some covariates in the model [9]. They found that factors that significantly influence the incidence of stunting are the percentage of poverty and undernourished children aged 0-59 months. These both factors have a positive effect on stunting incidence. However, these studies have not implemented the BSCAR localised model in modelling the RR and focused only on one province in Indonesia. This study aims to provide the most suitable BSCAR localised (clustering) model and identify the RR of stunting in each province in Indonesia.

2. MATERIALS AND METHOD

2.1. Data

Data on the number of toddlers 0-59 months whose height is measured and the number of stunted toddlers in each of the 34 provinces in Indonesia in 2021 were used. These data were gathered from Badan Pusat Statistik [10].

2.2. Models

A BSCAR localised model [11] was used to estimate the risk of stunting and examine the clusters of stunting cases. This model is similar to BSCAR Leroux [12] model but there is a clustering component (λ_{Z_i}) in the BSCAR localised model which allow adjacency different over areas.

The number of stunting cases (y_i) are modelled by using a Poisson log-linear model, commonly applied for disease mapping. This model is explained as follows:

$$y_i \sim \text{Poisson}(E_i\theta_i) \text{ for } i = 1, 2, 3, \ldots, 34 \text{ provinces}$$

$$log(\theta_i) = \beta_0 + u_i + \lambda z_i \quad (1)$$

E_i and θ_i are the number of expected cases and the RR in the i^{th} area, respectively. β_0 is the overall level of RR. The spatial random effect (u_i) is modelled using an intrinsic CAR prior as follows:

$$\left(u_i \middle| u_j, i \neq j, \tau_u^2\right) \sim N\left(\frac{\sum_j u_j \omega_{ij}}{\sum_j \omega_{ij}}, \frac{\tau_u^2}{\sum_j \omega_{ij}}\right) \quad (2)$$

ω_{ij} is the spatial adjacency matrix based on queen contiguity. ω_{ij} is defined by using the simplest form of weights (binary spatial weight) where $\omega_{ij} = 1$ if locations *i* and *j* are neighbours. Otherwise, $\omega_{ij} = 0$.

Five hyperpriors on the variance component τ_u^2 were used, that is, IG(1, 0.01) [13], IG(0.1, 0.1), IG (1, 0.1), IG (0.5, 0.05), and IG(0.5, 0.0005). The areas are partitioned into maximum G clusters which have their form of intercept and it is ordered as $\lambda_1 < \lambda_2 \ldots < \lambda_G$.

$$\lambda_k \sim \text{Uniform}(\lambda_{k-1}, \lambda_{k+1}) \text{ for } k = 1, 2, \ldots, G, \quad (3)$$

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

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

$$f(Z_i) = \frac{\exp(-\delta(Z_i - G^*)^2)}{\sum_{r=1}^{G} \exp(-\delta(r - G^*)^2)} \quad (4)$$

where $\delta \sim$ Uniform (1, 10).

$G^* = \frac{G+1}{2}$ if G is odd and $G^* = \frac{G}{2}$ if G is even.

G is suggested to be a small and odd number [11]. We choose different BSCAR localised models with a different maximum number of clusters: two clusters (G=2), three clusters (G=3), and five clusters (G=5) for each hyperprior.

All models were fit using the CARBayes package version 5.3 [13] in R software version 4.2.0 [14]. To check the convergence of the MCMC sample, trace plots, as well as density plots, were used. MCMC samples were generated based on 106,000 iterations. The burn-in of 6,000 samples was discarded. As a result, 100,000 MCMC samples were collected. The goodness of fit of the model formulation was compared using some criteria: Watanabe Akaike Information Criterion (WAIC) [15], the Deviance Information Criterion (DIC) [16], and Modified Moran's *I* (MMI) [17, 18] for the residuals. A model with a smaller value of DIC and WAIC fits the model well. Furthermore, a closer to zero of MMI residual signifies a better model fit.

Moran's *I* statistics on raw data were used to check the spatial autocorrelation and it is calculated as follows:

$$I = \frac{n \sum_{i=1}^{n} \sum_{j=1}^{n} \omega_{ij} (Y_i - \bar{Y})(Y_j - \bar{Y})}{\sum_{i=1}^{n} \sum_{j=1}^{n} \omega_{ij} (Y_i - \bar{Y})^2}$$

n is the number of areas (provinces), Y_i and Y_j are the observed data in the certain province *i* and province j, $\bar{X}$ is the average of all the X values over the n provinces, ω_{ij} is the spatial adjacency matrix. Moran's *I* statistics on residuals can also be used to detect model goodness of fit. However, it is recommended to use MMI for a small number of areas. An MMI was expanded [17] for detecting spatial dependence which works even for a small number of areas. Research on MMI is available in some pieces of literature [17, 18]. MMI statistic is calculated as follows:

$$I_{\text{Mod}} = \frac{\sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} (Y_i - \bar{Y})(\sum_{\boldsymbol{j=1}}^{\boldsymbol{n}} \omega_{ij} Y_j - \bar{Y})}{[\sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} (Y_i - \bar{Y})^2]^{1/2} \left[\sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \left(\sum_{\boldsymbol{j=1}}^{\boldsymbol{n}} \omega_{ij} \boldsymbol{Y_j} - \boldsymbol{\bar{Y}}\right)^{\boldsymbol{2}}\right]^{\boldsymbol{1/2}}}$$

3. Main Results

3.1. Descriptive Analysis

The total of children under five measured their height in Indonesia in 2021 is 24958590 with a mean (734076.2), median (404569.5), and standard deviation (1006036). Meanwhile, the total number of cases of stunting under five in Indonesia is 2373712 cases, with a mean (69815.06), a median (39629.5), and a standard deviation (92772.44). The provinces with the highest and the lowest stunting under-five cases are West Java (411000 cases) and Bangka Belitung (6171 cases), respectively.

Moran's I statistics for observed data is 0.38. The expected value and standard deviation for observed data are -0.03 and 0.15, respectively with p-value = 0.007. Given a p-value of Moran's I is 0.007 with a Moran's I value is 0.38, it indicates that there is a positive spatial dependence between areas. Furthermore, the value of MMI for observed data is also calculated (0.598) which is greater than Moran's I value.

3.2. Bayesian Spatial CAR localised Model

A BSCAR localised model was used with two (G =2), three (G=3), and five (G =5) clusters with different five hyperpriors. The DIC, WAIC, and MMI values for the residuals and the number of areas (provinces) included in each cluster are given in Table 1.

Table 1 indicates that the BSCAR localised model with G=2 in each hyperprior has a lower DIC value than the model with G=3 and G=5 except for hyperprior IG(0.5, 0.0005) (M15). Localised models with G=2 also have lower WAIC values compared to models with G=3 and G=5 for the five hyperpriors. The models with G=2 (M1=M4=M7=M10=M13) and G=3 (M2=M5=M8=M11=M14) have the same clustering structure for five different hyperpriors. Meanwhile, the model with G=5 has a different grouping structure for each hyperprior (M3=M6≠M9≠M12≠M15). The BSCAR localised model with G=3 with hyperprior IG(1, 0.1) (M5) has an MMI value for the residual that is the closest to zero (-0.32) and it is relatively similar to Model M10 (-0.41).

Table 1. The DIC, WAIC, and MMI for residuals, and the Number of Areas in the cluster

Hyperprior	Model		DIC	WAIC	MMI	Number of Areas in the cluster				
					Residual	G1	G2	G3	G4	G5
IG (1, 0.01)	M1	G = 2	**501.71**	**502.43**	-0.67	11	23	-	-	-
	M2	G = 3	510.25	525.57	-0.40	11	13	10	-	-
	M3	G = 5	506.80	590.67	-0.72	11	4	3	11	5
IG (1, 0.1)	M4	G = 2	**502.14**	**503.15**	-0.51	11	23	-	-	-
	M5	G = 3	510.66	528.56	**-0.32**	11	13	10	-	-
	M6	G = 5	511.60	604.04	-0.72	11	4	3	11	5
IG (0.1, 0.1)	M7	G = 2	**501.29**	**501.65**	-0.58	11	23	-	-	-
	M8	G = 3	510.34	524.95	-0.61	11	13	10	-	-
	M9	G = 5	531.18	641.32	-0.64	11	7	4	7	5
IG (0.5, 0.05)	M10	G = 2	**501.11**	**500.97**	**-0.41**	11	23	-	-	-
	M11	G = 3	510.30	524.97	-0.56	11	13	10	-	-
	M12	G = 5	523.39	586.60	-0.69	11	7	2	9	5
IG (0.5, 0.0005)	M13	G = 2	502.13	**502.71**	-0.48	11	23	-	-	-
	M14	G = 3	509.46	524.25	-0.64	11	13	10	-	-
	M15	G = 5	**430.73**	640.02	-0.78	11	7	10	1	5

Generally, based on the criteria used in this research, the preferred model for estimating the RR of stunting cases is BSCAR localised with hyperprior IG (0.5, 0.05) (M10) with two clusters (G=2). The localised structure (LS), as well as RR values for each province for G=2 with hyperprior IG (0.5, 0.05) (M10), were given in Table 2. The LS and RR values for BSCAR localised model with hyperprior IG (1, 0.1) for G=3 (M5) were also provided.

Table 2. The LS and RR values for each province for G=2 (M10) and G=3 (M5)

ID	Provinces	G=2 (M10)		G=3 (M5)	
		LS	RR	LS	RR
26	Sulawesi Barat	2	**2.63**	3	**2.63**
22	Nusa Tenggara Timur	2	2.23	3	2.23
24	Papua Barat	2	2.17	3	2.17
21	Nusa Tenggara Barat	2	2.06	3	2.06
12	Kalimantan Barat	2	1.81	3	1.81
23	Papua	2	1.51	3	1.51
15	Kalimantan Timur	2	1.50	2	1.50
10	Sulawesi Tengah	2	1.48	2	1.48
14	Kalimantan Tengah	2	1.44	3	1.44
1	Aceh	2	1.42	2	1.42
29	Sulawesi Tenggara	2	1.40	3	1.40
34	DI Yogyakarta	2	1.39	3	1.39
20	Maluku Utara	2	1.39	3	1.39
16	Kalimantan Utara	2	1.32	2	1.32
31	Sumatera Barat	2	1.28	2	1.28
10	Jawa Tengah	2	1.19	2	1.19
13	Kalimantan Selatan	2	1.15	2	1.15
19	Maluku	2	1.03	2	1.03
11	Jawa Timur	2	1.01	2	1.01
6	Gorontalo	2	0.99	2	0.99
27	Sulawesi Selatan	2	0.95	2	0.95
4	Banten	2	0.91	2	0.91
9	Jawa Barat	2	0.89	2	0.89
17	Kepulauan Riau	1	0.67	1	0.67
5	Bengkulu	1	0.63	1	0.63
25	Riau	1	0.63	1	0.63
33	Sumatera Utara	1	0.62	1	0.62
18	Lampung	1	0.57	1	0.57
8	Jambi	1	0.48	1	0.48
2	Bali	1	0.46	1	0.46
3	Bangka Belitung	1	0.46	1	0.46
32	Sumatera Selatan	1	0.41	1	0.41
30	Sulawesi Utara	1	0.33	1	0.33
7	DKI Jakarta	1	0.32	1	0.32

From Table 2, it shows that Sulawesi Barat has the highest RR of stunting (RR=2.63), followed by Nusa Tenggara Timur province (RR=2.23) and Papua Barat province (RR=2.17). In contrast, Jakarta has the lowest RR of stunting (RR=0.32), followed by Sulawesi Utara province (RR=0.33) and Sumatra Selatan provinces (RR=0.41). The RR values for both G=2 and G=3 were similar (Table 2). More than half of the provinces in Indonesia (19 out of 34) have a relative risk above the average (RR>1). In another word, about 56 % of provinces in Indonesia are at a high risk of stunting. While the other 15 provinces have a relative risk below the average (RR<1). The visualization of the clustering structure of Model M10 and Model M5 is provided in Figure 1 and Figure 2, respectively.

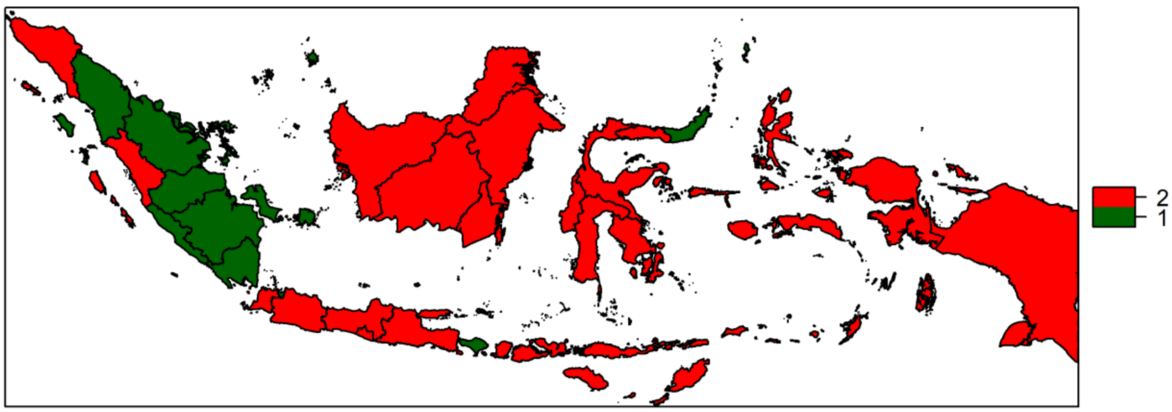

Fig. 1. Localised map based on the BSCAR localised model with G=2 (M10).

Based on Figure 1, eight out of ten provinces on Sumatera island are included in cluster 1(green). Only 2 provinces (Aceh, and Sumatera Barat) are included in cluster 2 (red). Overall, cluster 1 is dominated by provinces on Sumatera Island. On the other hand, all provinces in Papua Island (the first biggest island in Indonesia) are included in cluster 2, likewise Kalimantan Island (the second biggest island in Indonesia). In Sulawesi Island, all provinces are included in cluster 2 except Sulawesi Utara.

When using the BSCAR localised model with the number of clusters is 3 (G=3), all provinces in Papua Island are included in cluster 3 (Figure 2). The provinces which are included in cluster 1 are the same for both model BSCAR localised model with G=2 and BSCAR localised model with G=3.

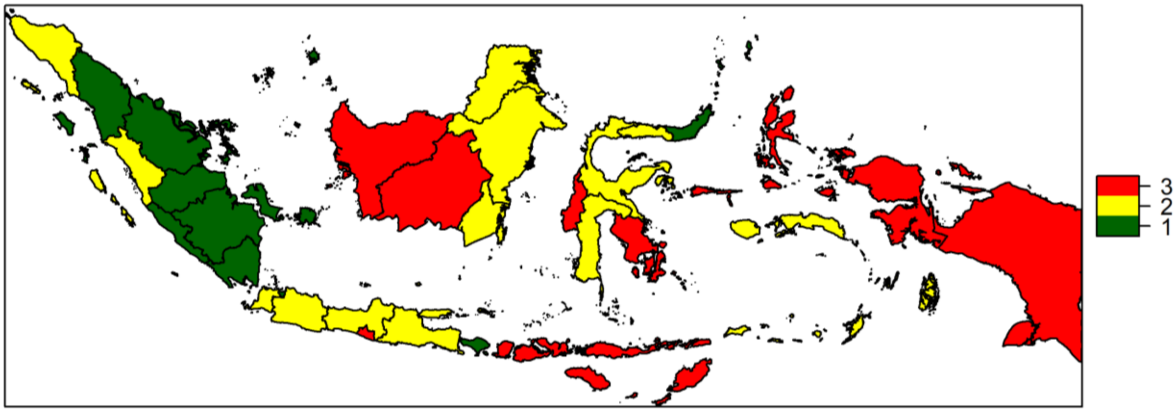

Fig. 2. Localised maps based on the BSCAR localised model with G=3 (M5)

The RR maps of stunting cases based on the BSCAR localised model with G=2 (M10) are given in Figure 3. Out of ten provinces on the Sumatera island, only two provinces (Aceh, and Sumatera Barat) have a relative risk above the average (RR>1). The RR values of Aceh province (ID=1) and Sumatera Barat province (ID=31) are 1.42 and 1.28, respectively.

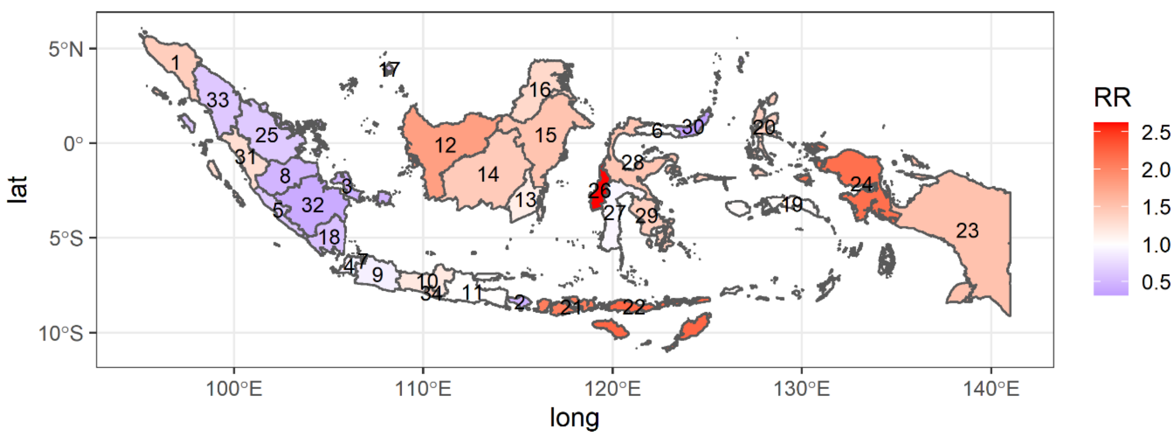

Fig. 3. The RR map based on BSCAR localised model with G=3 (M5)

4. Conclusions

The results indicated that the BSCAR localised with hyperprior IG (0.5, 0.05) and IG (1, 0.1) are preferred for two and three clusters, respectively. Our results identified the high-risk areas for stunting. Approximately 56% of provinces in Indonesia are at a high risk of stunting. Sulawesi

Barat has the highest RR for stunting followed by Nusa Tenggara Timur dan Papua Barat. In contrast, Jakarta has the lowest RR of stunting followed by Sulawesi Utara and Sumatera Selatan. Government should pay more attention to areas that are most at high risk of stunting. Including some covariates in the model could be possible for future work.

ACKNOWLEDGMENTS

The authors would like to thank Nurwan, S.Stat for his valuable contributions.

CONFLICT OF INTERESTS

The authors declare that there is no conflict of interests.

REFERENCES

[1] R.I. Kemenkes, Situasi balita pendek (Stunting) di Indonesia, Bul Jendela Data dan Inf Kesehatan. 2018.

[2] T. Eryando, T. Sipahutar, M.P. Budhiharsana, et al. Spatial analysis of stunting determinants in 514 Indonesian districts/cities: Implications for intervention and setting of priority, Geospat Health. 17 (2022), 1055. https://doi.org/10.4081/gh.2022.1055.

[3] S. Shaw, J. Khan, B. Paswan, Spatial modeling of child malnutrition attributable to drought in India, Int. J. Public Health. 65 (2020), 281–290. https://doi.org/10.1007/s00038-020-01353-y.

[4] B.T. Seboka, S. Hailegebreal, T.T. Mamo, et al. Spatial trends and projections of chronic malnutrition among children under 5 years of age in Ethiopia from 2011 to 2019: a geographically weighted regression analysis, J. Health Popul. Nutr. 41 (2022), 28. https://doi.org/10.1186/s41043-022-00309-7.

[5] J.M.K. Aheto, G.A. Dagne, Geostatistical analysis, web-based mapping, and environmental determinants of under-5 stunting: evidence from the 2014 Ghana Demographic and Health Survey, The Lancet Planetary Health. 5 (2021), e347–e355. https://doi.org/10.1016/s2542-5196(21)00080-2.

[6] O.A. Egbon, A.M. Belachew, M.A. Bogoni, Risk factors of concurrent malnutrition among children in Ethiopia: a bivariate spatial modeling approach, All Life. 15 (2022), 512–536. https://doi.org/10.1080/26895293.2022.2067251.

[7] L.N. Kazembe, N.M. Nickanor, Spatial modelling of the relationship between socio-economic disadvantage and

child health in Namibia, Spat. Demogr. 5 (2016), 1–24. https://doi.org/10.1007/s40980-016-0022-8.

[8] A. Aswi, S. Sukarna, and Nurhilaliyah, Pemetaan risiko relatif kasus stunting di provinsi sulawesi selatan, Sainsmat: Jurnal Ilmiah Ilmu Pengetahuan Alam, 11 (2022), 11-20.

[9] A. Aswi, Sukarna, Pemodelan spasial bayesian dalam menentukan faktor yang mempengaruhi kejadian stunting di provinsi sulawesi selatan, J. Math. Comput. Stat. 5 (2022), 1-11.

[10] Badan Pusat Statistik, Sulawesi Selatan dalam Angka 2021, (2021).

[11] D. Lee, C. Sarran, Controlling for unmeasured confounding and spatial misalignment in long-term air pollution and health studies, Environmetrics. 26 (2015), 477–487. https://doi.org/10.1002/env.2348.

[12] B.G. Leroux, X. Lei, N. Breslow, Estimation of disease rates in small areas: a new mixed model for spatial dependence, in: M.E. Halloran, D. Berry (Eds.), Statistical Models in Epidemiology, the Environment, and Clinical Trials, Springer New York, New York, NY, 2000: pp. 179–191. https://doi.org/10.1007/978-1-4612-1284-3_4.

[13] D. Lee, CARBayes: An R package for Bayesian spatial modeling with conditional autoregressive priors, J. Stat. Soft. 55 (2013), 1-24. https://doi.org/10.18637/jss.v055.i13.

[14] R Core Team, R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, 2019.

[15] S. Watanabe, Asymptotic equivalence of bayes cross validation and widely applicable information criterion in singular learning theory, J. Mach. Learn. Res. 11 (2010), 3571-3594.

[16] D.J. Spiegelhalter, N.G. Best, B.P. Carlin, et al. Bayesian measures of model complexity and fit, J. R. Stat. Soc. B. 64 (2002), 583–639. https://doi.org/10.1111/1467-9868.00353.

[17] T.B. Carrijo, A.R. da Silva, Modified Moran's I for small samples, Geogr. Anal. 49 (2017), 451–467. https://doi.org/10.1111/gean.12130.

[18] A. Aswi, S. Cramb, E. Duncan, et al. Detecting spatial autocorrelation for a small number of areas: a practical example, J. Phys.: Conf. Ser. 1899 (2021), 012098. https://doi.org/10.1088/1742-6596/1899/1/012098.