Available online at http://scik.org Commun. Math. Biol. Neurosci. 2023, 2023:95 https://doi.org/10.28919/cmbn/8087 ISSN: 2052-2541

ANALYZING THE IMPACT OF HEALTHCARE FACILITIES ON THE SPREAD OF COVID-19: A SPATIAL AUTOREGRESSIVE EXOGENOUS MODEL APPROACH

ALMEIRA TSANAWAFA¹, AJENG BERLIANA SALSABILA¹, SRI PURWANI^{2,*}

¹Mathematics Postgraduate Program, Universitas Padjadjaran, Sumedang 45363, Indonesia

²Departement of Mathematics, Universitas Padjadjaran, Sumedang 45363, Indonesia

Copyright © 2023 the author(s). This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract: This paper investigates the relationship between healthcare facilities and the spread of COVID-19 in Indonesia using a spatial autoregressive exogenous (SAR-X) model. The study aims to understand how the availability and accessibility of healthcare facilities influence the transmission dynamics of the virus at a regional level. The analysis utilizes a comprehensive dataset on COVID-19 cases, healthcare. The findings of this study contribute to a better understanding of the role of healthcare facilities in mitigating the spread of COVID-19 in Indonesia and can be a policy decision to strengthen healthcare infrastructure and resource allocation, facility locations, and other relevant variables. The SAR-X model allows for the incorporation of exogenous factors, such as healthcare capacity, population density, and mobility patterns, in examining their impact on the spatial patterns of COVID-19 transmission.

Keywords: COVID-19; healthcare facilities; SAR-X model; spatial autoregressive; transmission dynamics; spatial analysis.

2020 AMS Subject Classification: 62P10, 92D25.

*Corresponding author E-mail address: sri.purwani@unpad.ac.id Received July 03, 2023

1. INTRODUCTION

The COVID-19 pandemic has had a profound impact on global health systems and economies [1], including in Indonesia. As the number of COVID-19 cases continues to rise, understanding the factors that contribute to the spread of the virus is crucial for effective public health interventions. One such factor that plays a significant role in disease transmission is the availability and quality of healthcare facilities. Access to adequate healthcare services and infrastructure is essential for timely diagnosis, treatment, and containment of the virus [2]. Therefore, investigating the impact of healthcare facilities on the spread of COVID-19 in Indonesia is of utmost importance. In recent years, spatial analysis techniques have gained popularity in epidemiology research for understanding the spatial patterns of disease transmission [3]. Spatial Autoregressive (SAR) models have proven effective in capturing spatial dependencies and exploring the impact of exogenous factors on disease spread [4]. In this study, we employ a SAR-X model to analyze the relationship between healthcare facilities and the spread of COVID-19 in Indonesia.

Previous research stated that the environmental sector played an important role in increasing positive confirmed cases of Covid-19 [5]–[13]. In addition, the age factor also greatly influences the spread of Covid-19 [10], [12], [14]. The primary objective of this research is to investigate how the availability and accessibility of healthcare facilities influence the spatial patterns of COVID-19 transmission in Indonesia. We hypothesize that areas with better healthcare infrastructure and higher capacity for testing and treatment will exhibit lower rates of COVID-19 transmission. By incorporating exogenous variables such as the number of hospitals, clinics, and health centers, the SAR-X model allows us to examine the independent and combined effects of these factors on the spread of the virus. To accomplish our research objectives, we will utilize a comprehensive dataset on COVID-19 cases and healthcare facilities. By mapping the spatial distribution of COVID-19 cases and healthcare facilities, we can identify hotspots with high transmission rates and areas with limited access to healthcare services. The results of this analysis will provide valuable insights into the relationship between healthcare facilities and the spread of COVID-19 in Indonesia, informing policymakers and healthcare authorities on

strategies for improving healthcare infrastructure and resource allocation to mitigate the impact of the pandemic.

2. MATERIALS AND METHODS

The purpose of this paper aims to contribute to the existing literature on COVID-19 transmission by examining the impact of healthcare facilities in Indonesia using a SAR-X model.

2.1. Spatial Weighted Matrix

This matrix is used to define the weight between observed locations which is based on the neighborhood relationship between locations [15]. Neighborhood can be defined using the concept of queen contiguity which is illustrated in Figure 1.



Figure 1. Queen Contiguity

Figure 1 shows that the neighboring areas are determined based on the angles and sides that touch each other. There are two ways in obtaining the spatial weighting matrix W, that is the standardized and the unstandardized weighting matrix. The former is obtained by giving an equal weight to the nearest neighbors and zero to the others, whereas the latter is obtained by assigning a weight to the nearest neighbors and zero to the others.

2.2. Moran's Index

To determine the presence of spatial autocorrelation between locations, a spatial dependency test was used [16]. The spatial dependence test is performed using the Moran's index.

Hypothesis:

 H_0 : spatial autocorrelation between locations rejected.

 H_1 : spatial autocorrelation between locations.

Test statistics uses:

$$Z_{\text{value}} = \frac{I - E(I)}{\sqrt{\text{Var}(I)}}$$

with

$$I = \frac{n\sum_{i=1}^{n}\sum_{j=1}^{n}w_{ij}c_{ij}}{\sum_{i=1}^{n}\sum_{j=1}^{n}w_{ij}(x_{i}-\bar{x})^{2}}; E(I) = -\frac{1}{n-1}; Var(I) = \frac{n^{2}S_{1}-nS_{2}+3S_{0}^{2}}{(n^{2}-1)S_{0}^{2}} - [E(I)]^{2};$$
$$c_{ij} = (x_{i}-\bar{x})(x_{j}-\bar{x}); S_{0} = \sum_{i=1}^{n}\sum_{j=1}^{n}w_{ij}; S_{1} = \frac{1}{2}\sum_{j\neq 1}^{n}(w_{ij}+w_{ji})^{2}; S_{2} = \sum_{j\neq 1}^{n}(w_{ij}+w_{ji})^{2}$$

I: Moran's index value

Var(*I*): Moran index variance

E(I): Moran's index expected value

n: the number of incident locations

 x_i : value at the i

 x_i : value at the j

 \overline{x} : the average of variables

 w_{ij} : standardized spatial weight matrix i, j

Decision:

 H_0 rejected if $Z_{\text{value}} < -Z_{\frac{\alpha}{2}}$ or $Z_{\text{value}} > Z_{\frac{\alpha}{2}}$.

2.3. Spatial Autoregressive Exogenous Model

The SAR-X model is used to describe and predict the effect of location and exogenous variables on dependent variables with location-based data [17]. Referring to [18] in [16] it can be explained that the general model of the SAR-X regression can be written as follows:

$$\mathbf{y} = \rho \mathbf{W} \mathbf{y} + \mathbf{X} \boldsymbol{\beta} + \boldsymbol{\varepsilon}, \, \boldsymbol{\varepsilon} \sim \mathbf{N}(0, \sigma^2 \mathbf{I})$$

iid

with,

y: dependent variable vector of size $n \times 1$

 ρ : spatial lag parameter coefficient of dependent variable

W: spatial weighted matrix of size $n \times n$

X: independent variable matrix of size $n \times (p + 1)$

β: regression parameter coefficient vector of size $(p + 1) \times 1$

 $\boldsymbol{\varepsilon}$: error vector of size $n \times 1$

2.4. Parameter estimation using Maximum Likelihood Method

The random error variable in the SAR-X model has a normal distribution and is assumed to be normal, so that the Maximum Likelihood Estimation (MLE) method will be used in the parameter estimation of the SAR-X model. Referring to [19] unknown parameters in the MLE method will be obtained by maximizing a likelihood function.

SAR-X model can be written as follows:

$$\mathbf{\varepsilon} = \mathbf{y} - \rho \mathbf{W} \mathbf{y} - \mathbf{X} \boldsymbol{\beta}$$

The probability density function is defined as follows:

$$f(\boldsymbol{\varepsilon}|\rho,\boldsymbol{\beta}) = \frac{1}{(2\pi)^{\frac{n}{2}}(\sigma^2)^{\frac{n}{2}}} \exp\left(-\frac{(\mathbf{y}-\rho\mathbf{W}\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^{\mathrm{T}}(\mathbf{y}-\rho\mathbf{W}\mathbf{y}-\mathbf{X}\boldsymbol{\beta})}{2\sigma^2}\right)$$

where the likelihood function:

$$L(\rho, \boldsymbol{\beta}|\boldsymbol{\varepsilon}) = f(\boldsymbol{\varepsilon}|\rho, \boldsymbol{\beta})$$
$$= \frac{1}{(2\pi)^{\frac{n}{2}} (\sigma^2)^{\frac{n}{2}}} \exp\left(-\frac{(\mathbf{y} - \rho \mathbf{W}\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^{\mathrm{T}} (\mathbf{y} - \rho \mathbf{W}\mathbf{y} - \mathbf{X}\boldsymbol{\beta})}{2\sigma^2}\right).$$

Thus, the log likelihood function is given as follows:

$$\ln L(\rho, \boldsymbol{\beta} | \boldsymbol{\varepsilon}) = \ln \left(\frac{1}{(2\pi)^{\frac{n}{2}} (\sigma^2)^{\frac{n}{2}}} \exp \left(-\frac{(\mathbf{y} - \rho \mathbf{W} \mathbf{y} - \mathbf{X} \boldsymbol{\beta})^{\mathrm{T}} (\mathbf{y} - \rho \mathbf{W} \mathbf{y} - \mathbf{X} \boldsymbol{\beta})}{2\sigma^2} \right) \right)$$
$$= -\frac{n}{2} \ln (2\pi) - \frac{n}{2} \ln \sigma^2 - \frac{(\mathbf{y} - \rho \mathbf{W} \mathbf{y} - \mathbf{X} \boldsymbol{\beta})^{\mathrm{T}} (\mathbf{y} - \rho \mathbf{W} \mathbf{y} - \mathbf{X} \boldsymbol{\beta})}{2\sigma^2}$$

The estimation parameters ρ and β obtained by maximizing the log likelihood function are given as follows:

$$\hat{\rho} = ((\mathbf{W}\mathbf{y})^{\mathsf{T}}\mathbf{W}\mathbf{y})^{-1}(\mathbf{W}\mathbf{y})^{\mathsf{T}}(\mathbf{y})$$
$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^{\mathsf{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathsf{T}}(\mathbf{y} - \rho\mathbf{W}\mathbf{y})$$

2.5. Parameter Significance Test with Wald

Parameter significance test is carried out to determine the role of independent variables in the

model [20].

The hypothesis used:

 H_0 : significant parameter rejected

 H_1 : significant parameter

The test statistics is as follows:

Wald =
$$\frac{\widehat{\beta}^2}{\operatorname{Var}(\widehat{\beta})}$$

with,

$$\operatorname{Var}(\widehat{\boldsymbol{\beta}}) = \operatorname{Var}\left((\mathbf{X}^{\mathsf{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathsf{T}}(\mathbf{y} - \rho \mathbf{W}\mathbf{y})\right)$$

Decision:

 H_0 is rejected if Wald value > $\chi^2_{(1,1-\alpha)}$.

2.6. Mean Absolute Percentage Error (MAPE)

Referring to Lewis in [21], MAPE is an evaluation forecasting method that considers the effect of actual values. The lower the MAPE value, the more accurate the method accuracy is.

MAPE =
$$\frac{\left(\sum_{i=1}^{n} \frac{|y_i - \hat{y}_i|}{y_i}\right)}{n} \times 100\%$$

with,

 y_i : actual data

 \hat{y}_i : prediction data

n: the amount of data.

Referring to Lewis in [21], there is a scale for assessing MAPE accuracy shown in Table 1.

Scale MAPE	Accuracy Score
≤ 10%	Highly accurate prediction
$10\% < MAPE \le 20\%$	Good prediction
$20\% < MAPE \le 50\%$	Reasonable Prediction
> 50%	Inaccurate Prediction

 Table 1. MAPE score scale.

2.7. Coefficient of Determination

The coefficient of determination is used to see the suitability of the model used [22]. Defining the total squared sum (TSS)

$$TSS = \sum_{i}^{n} (y_i - \bar{y})^2$$

and the sum of squared error (SSE)

$$SSE = \sum_{i}^{n} (y_i - \widehat{y}_i)^2,$$

gives the general formula for the coefficient of determination as follows:

$$R^2 = 1 - \frac{SSE}{TSS}$$

Referring to [22], the closer R^2 to 1, the more suitable the model is, conversely, the closer R^2 to 0, the less suitable the model is.

3. RESULTS AND DISCUSSION

The data used is the daily positive confirmed Covid-19 patients in West Java Province in 2022 and the exogenous variable is the number of health facilities (hospitals, clinics, health centres) in 2020.



Figure 2. Map of West Java Province

Figure 2 shows a map of West Java Province consisting of 9 cities (Bandung, Banjar, Bekasi, Bogor, Cimahi, Cirebon, Depok, Sukabumi, Tasikmalaya) and 18 regencies (Bandung Barat, Bekasi, Bogor, Ciamis, Cianjur, Cirebon, Garut, Indramayu, Karawang, Kuningan, Majalengka, Pangandaran, Purwakarta, Subang, Sukabumi, Sumedang, Tasikmalaya). Applying the concept of queen contiguity to data in Figure 2, results in the spatial weighting matrix **W**.

0 1 0 0 0 0 0 0 0 0 0 1 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 0 0 0 1 1 0 1 1 **W** = 0 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 L۵ 0 0 0 0 0 0 0 0 0

Results of applying spatial dependency testing using the Moran's I index with $\alpha = 0.05$, are shown in Table 2.

 Table 2. Moran Index

Ι	E(I)	Var(I)	Z _{value}	H_0	Autocorrelation
0.5175357	-0.2	0.1224539	2.050500	Reject	\checkmark

Based on Table 2, the computed $Z_{value}(2.050500) \ge Z_{\frac{\alpha}{2}}(1.96)$, indicating the presence of spatial autocorrelation. This is shown in Table 3. The estimated parameter results of the SAR-X

model were obtained using MLE.

Table 3. Estimation result		
ρ	β	
0.1983125	7.7181327	

The positive value of $\hat{\rho}$ in Table 3 indicates the presence of spatial lag in the confirmed COVID-19 cases in West Java Province. The positive value of $\hat{\beta}$ indicates that, holding other variables constant, an increase of 1 unit in the number of healthcare facilities is associated with an increase in the dependent variable. The significance testing of the parameters using the Wald test is shown in Table 4.

Table 4. Wald test

Wald value	H_0	Significance of parameters
3.4990	Rejected	Significant

Table 4 shows that the obtained Wald value is greater than $\chi^2_{1,9.5}$, indicating that the healthcare facilities variable is significant in relation to the confirmed COVID-19 cases. The MAPE obtained from the SAR-X model and the SAR-X prediction model for daily confirmed positive COVID-19 cases in West Java Province is shown in Table 5.

Table 5. MAPE	

MAPE	Accuracy Score
14.5678%	Good prediction

Table 5 shows that the prediction result is accurate, indicating that the SAR-X model is suitable for predicting the impact of location and the number of healthcare facilities on daily confirmed positive COVID-19 cases. This is supported by the generated coefficient of determination of 78%.

4. CONCLUSION

The results of the spatial dependence test indicate the presence of spatial autocorrelation in daily confirmed cases of Covid-19. The Wald test results show that the number of healthcare facilities

has a significant effect on daily confirmed cases of Covid-19 in West Java Province. The application of the SAR-X model can be used to predict the impact of location and the number of healthcare facilities on daily confirmed cases of Covid-19. This is supported by a coefficient of determination of 78%. The findings of this study can serve as a recommendation for the Health Department in addressing the challenges of managing Covid-19 patients.

ACKNOWLEDGMENTS

We would like to thank the Rector and the DRPM of Universitas Padjadjaran for financial support of the article publication.

CONFLICT OF INTERESTS

The authors declare that there is no conflict of interests.

References

- [1] A.V. Devereaux, H. Backer, A. Salami, et al. Oxygen and ventilator logistics during California's COVID-19 surge: when oxygen becomes a scarce resource, Disaster Med. Public Health Prep. 17 (2021), E33. https://doi.org/10.1017/dmp.2021.267.
- I. Mustapha, N. Khan, M.I. Qureshi, et al. Assessing hospital management performance in intensive care units (ICUs) during the COVID-19: a study from the pandemic outbreak perspective, Int. J. Online Biomed. Eng. 18 (2022), 154-168. https://doi.org/10.3991/ijoe.v18i10.32733.
- [3] C. de A. Medeiros, M.B.A. Silva, A.L.S. de Oliveira, et al. Spatial analysis of the natural infection index for Triatomines and the risk of Chagas disease transmission in Northeastern Brazil, Rev. Inst. Med. Trop. S. Paulo. 65 (2023), e32. https://doi.org/10.1590/s1678-9946202365032.
- [4] R.P. Narayanan, J. Nordlund, R.K. Pace, et al. Demographic, jurisdictional, and spatial effects on social distancing in the United States during the COVID-19 pandemic, PLoS ONE. 15 (2020), e0239572. https://doi.org/10.1371/journal.pone.0239572.
- [5] S.R. Ocampo, E.B. Barrios, Sparse spatial autoregressive and spatio-temporal models for COVID-19 incidence

in the Philippines, Philippine J. Sci. 151 (2022), 1979-1994.

- [6] B. Sung, A spatial analysis of the effect of neighborhood contexts on cumulative number of confirmed cases of COVID-19 in U.S. Counties through October 20 2020, Prevent. Med. 147 (2021), 106457. https://doi.org/10.1016/j.ypmed.2021.106457.
- Y. Sun, X. Hu, J. Xie, Spatial inequalities of COVID-19 mortality rate in relation to socioeconomic and environmental factors across England, Sci. Total Environ. 758 (2021), 143595. https://doi.org/10.1016/j.scitotenv.2020.143595.
- [8] S. Bourdin, L. Jeanne, F. Nadou, et al. Does lockdown work? A spatial analysis of the spread and concentration of Covid-19 in Italy, Region. Stud. 55 (2021), 1182-1193. https://doi.org/10.1080/00343404.2021.1887471.
- [9] W. Xi, T. Pei, Q. Liu, et al. Quantifying the time-lag effects of human mobility on the COVID-19 transmission: a multi-city study in China, IEEE Access. 8 (2020), 216752-216761. https://doi.org/10.1109/access.2020.3038995.
- [10] R. Breen, J. Ermisch, The distributional impact of Covid-19: Geographic variation in mortality in England, Demographic Res. 44 (2021), 397-414. https://doi.org/10.4054/demres.2021.44.17.
- [11] M. Lomuscio, Exploring local dynamics of COVID-19 in Italy: local labour markets and the cases of Bergamo and Brescia, J. Entr. Organ. Diversity. 10 (2021), 22-47. https://doi.org/10.5947/jeod.2021.002.
- M. Myck, M. Oczkowska, C. Garten, et al. Deaths during the first year of the COVID-19 pandemic: insights from regional patterns in Germany and Poland, BMC Public Health. 23 (2023).
 https://doi.org/10.1186/s12889-022-14909-9.
- [13] Y. Chen, Y. Li, S. Feng, et al. Gravitational scaling analysis on spatial diffusion of COVID-19 in Hubei Province, China, PLoS ONE. 16 (2021), e0252889. https://doi.org/10.1371/journal.pone.0252889.
- [14] H. Yu, J. Li, S. Bardin, et al. Spatiotemporal dynamic of COVID-19 diffusion in China: A dynamic spatial autoregressive model analysis, ISPRS Int. J. Geo-Inf. 10 (2021), 510. https://doi.org/10.3390/ijgi10080510.
- [15] R. Kosfeld, Spatial econometrics, Institute of Economics, University of Kassel, Germany. (2006).
- [16] K.A. Kuse, D.D. Debeko, Spatial distribution and determinants of stunting, wasting and underweight in children under-five in Ethiopia, BMC Public Health. 23 (2023), 641. https://doi.org/10.1186/s12889-023-15488-z.

- [17] G.M. Robinson, Time series analysis, in: International Encyclopedia of Human Geography, Elsevier, 2020: pp. 291-300. https://doi.org/10.1016/B978-0-08-102295-5.10614-6.
- [18] D.A. Griffith, L. Anselin, Spatial econometrics: methods and models, Econ. Geography. 65 (1989), 160. https://doi.org/10.2307/143780.
- [19] J. Tang, X. Shi, X. Hu, Analysis of air pollution in Fenwei Plain in China based on functional spatial autoregressive combined model, PLoS ONE. 18 (2023), e0283336. https://doi.org/10.1371/journal.pone.0283336.
- [20] W. Yamaka, R. Shi, P. Maneejuk, et al. Spatial spillover effects of internet development on foreign trade in China, Sustainability. 15 (2023), 4213. https://doi.org/10.3390/su15054213.
- [21] K.D. Lawrence, R.K. Klimberg, S.M. Lawrence, Fundamentals of forecasting using Excel, Industrial Press, New York, 2009.
- [22] S. Hosseini, R. Pourmirzaee, D.J. Armaghani, et al. Prediction of ground vibration due to mine blasting in a surface lead–zinc mine using machine learning ensemble techniques, Sci. Rep. 13 (2023), 6591. https://doi.org/10.1038/s41598-023-33796-7.