

Article

The Analysis of Factors Influencing Incidence rates of Toddler Pneumonia in Purwakarta Districts Using Panel Data Spatial Regression

Article Info

Article history :

Received January 23, 2023
Revised February 15, 2023
Accepted February 20, 2023
Published June 30, 2023

Keywords :

Pneumonia, omitted variables, spatial dependencies, fixed effect spatial models

Nadiyah Nisrina¹, Budhi Handoko^{1*}, Yudhie Andriyana¹

¹Department of Statistics, Faculty of Mathematics and Natural Science (FMIPA), Universitas Padjadjaran, Sumedang, Indonesia

Abstract. Pneumonia is an acute respiratory infection that attacks the lungs and can cause inflammation of the air sacs due to the alveoli is filled with pus and fluid. This research aims at identifying factors influencing pneumonia and mapping its incidence rate for toddlers in the Purwakarta Regency. Many factors influence pneumonia, but due to the limitation of data or information, some factors cannot be included in the model and are called omitted variables. The incidence rate of toddler pneumonia in sub-districts of Purwakarta Regency is assumed to be related to one another or have a spatial dependency. Therefore, modeling pneumonia with the Fixed Effect Spatial Model can accommodate spatial aspects. The results show that MR2 measles immunization, low birth weight, exclusive breastfeeding, and clean and healthy living habits significantly affect the incidence rate of toddler pneumonia. Based on the mapping results, Wanayasa sub-district has a high incidence rate of toddler pneumonia, while some sub-districts such as Campaka, Pondoksalam, and Darangdan have low incidence rates.

This is an open access article under the [CC-BY](https://creativecommons.org/licenses/by/4.0/) license.



This is an open access article distributed under the Creative Commons 4.0 Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. ©2023 by author.

Corresponding Author :

Budhi Handoko

Department of Statistics, Faculty of Mathematics and Natural Science (FMIPA), Universitas Padjadjaran, Sumedang, Indonesia

Email : budhi.handoko@unpad.ac.id

1. Introduction

Pneumonia is a form of acute respiratory infection that attacks the lungs and can cause inflammation of the air sacs due to the alveoli being filled up with pus and fluid [1]. The symptoms of pneumonia are cough, dyspnea, sputum production, and chest pain. In addition, someone who suffers from

pneumonia might have additional symptoms such as fever, rigors, myalgia, and confusion [2-4]. Pneumonia can attack both adults and children. One dangerous type of pneumonia for adults is *Pneumocystis pneumonia* (PCP), as it causes morbidity and mortality in developed countries [5-7]. According to WHO data in 2020, pneumonia causes 1.4 million children under five to die yearly. According to [8-10], Java Island is the most significant contributor to the incidence of toddler pneumonia, and the highest incidence was found in West Java Province, with 104,866 toddlers and a prevalence of 4.62%. Compared to other provinces, the prevalence rate of toddler pneumonia in West Java Province is above the national limit, which should be less than 3.55%. Based on profile data from [11], Purwakarta Regency, one of the districts in West Java, has a high number of pneumonia cases. To understand the development of pneumonia cases yearly in Purwakarta District, we investigate a graph of the incidence rate of toddler pneumonia from 2016 to 2021.

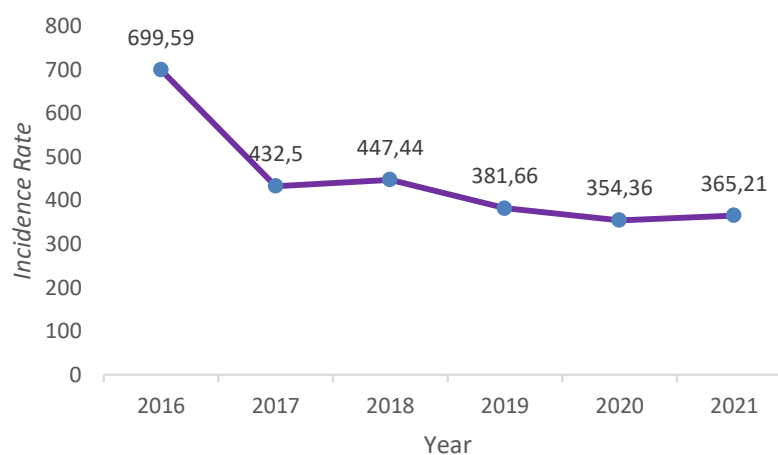


Figure 1. Incidence Rate of Toddler pneumonia in Purwakarta Regency from 2016 to 2021

Based on Figure 1, the incidence rates of toddler pneumonia in the Purwakarta District in the last six years tend to fluctuate, and their values are far above the target set by the work plan (Renja) of the Dinas Kesehatan Kabupaten Purwakarta in 2020, which should be less than 245 per 100,000 residents. To suppress the incidence rate and death from toddler pneumonia, the Purwakarta Health Office carried out various prevention efforts, including pneumonia's early detection. To do so, we propose a method for modeling factors that are suspected to affect the incidence rate of pneumonia which is called the risk factors. The risk factors for pneumonia can be classified into individual, behavior, and environmental factors. However, due to the limitation of data availability, not all of these risk factors are included in the analysis. Variables that should be included in the model but are omitted called omitted variables. However, taking into account these variables may lead to bias and inconsistent estimators [12]. To overcome this issue, we use panel data regression which considers omitted variables.

Pneumonia is an infectious disease transmitted from pneumonia sufferers who spread the virus by droplets into the air when coughing or sneezing. Consequently, adjacent areas may develop pneumonia as well. Given this situation, it is worth noting that the analysis should consider spatial dependence to determine the relationship of an element in an area with other adjacent areas. Some research that considers spatial dependence as it can affect the result of the analysis is conducted by [13-15]. In this paper, we fit a model involving factors that affect the incidence rate of toddler pneumonia using panel data spatial regression. In addition, to investigate areas prone to contracting toddler pneumonia, we create an incident rate map of districts in Purwakarta Regency.

2. Materials and Methods

2.1 Materials

The material used in this research is the data obtained from the Health Office of Purwakarta District. The observation unit involves 17 districts in Purwakarta Regency and covers three years from 2019 to 2021. The dataset contains some variables: the incidence rate as the dependent variable (Y), MR2 Measles Immunization (X1), Low Birth Weight (X2), Exclusive Breastfeeding (X3), Nutritional Status (X4), and Clean and Healthy Behavior (X5), are as independent variables (X).

2.2 Methods

The method used in this study is Spatial Panel Data Regression, which is the extension of Panel Data Regression. The spatial panel data regression has two types: spatial lag fixed effect model and spatial error fixed effect model. The explanation for each model is as follows.

2.2.1 Panel Data Regression

Panel data regression combines cross-section and time series data so that the number of observations is large. According to [16], the panel data regression model is expressed as follows

$$Y_{it} = \alpha + \mathbf{x}'_{it}\boldsymbol{\beta} + \varepsilon_{it},$$

where i is the cross-section unit where $i = 1, 2, \dots, N$; t denotes the number of periods, where $t = 1, 2, \dots, T$. \mathbf{x} is the independent variable matrix, $\boldsymbol{\beta}$ is the vector of the regression parameter coefficient and ε_{it} denotes an error. In this study, we select one of the two approaches to the panel data regression model, fixed effect model (model with individual influence) and random effect model (model with error effect). The fixed effect model is a model with different effects between individuals. The slope coefficient in the fixed effect model is constant over time, but the intercept varies between cross-section units. The general form of the fixed effect model according to [16] is as follows

$$Y_{it} = \alpha + \mathbf{x}'_{it}\boldsymbol{\beta} + \mu_i + \varepsilon_{it}$$

where μ_i is a specific effect of the individual and ε_{it} denotes errors. Difference intercept in the panel data regression analysis for the fixed effect case can be represented as incomplete information in making the model due to omitted variables.

On the other hand, the random effects model is a panel data model with differences in individual characteristics that are treated as part of the error component. Thus, individual characteristics are accommodated to the error of the model. The slope and intercept coefficients in the random effect model are constant. The general form of the random effect model is as follows

$$Y_{it} = \alpha + \mathbf{x}'_{it}\boldsymbol{\beta} + u_i + \varepsilon_{it}$$

where u_i is the cross-section error component. Spatial regression is an analysis that evaluates the relationship between variables that consider spatial effects or the presence of an area or location effect. The general model of spatial regression according to [17] is

$$\mathbf{y} = \alpha + \lambda \mathbf{W}_\lambda \mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \mathbf{W}_\beta \mathbf{X}_w \boldsymbol{\beta}_w + \boldsymbol{\phi}$$

$$\boldsymbol{\phi} = \rho \mathbf{W}_\rho \boldsymbol{\phi} + \boldsymbol{\varepsilon}$$

where \mathbf{y} is the dependent variable vector, \mathbf{X} is the independent variable matrix, λ is the spatial lag autoregression coefficient, ρ is the spatial error autoregression coefficient, $\boldsymbol{\beta}$ is the coefficient vector regression parameters, \mathbf{W} is the spatial weight matrix, $\boldsymbol{\phi}$ is the error vector containing autocorrelation, and $\boldsymbol{\varepsilon}$ is an autocorrelation-free error vector. The spatial models specification might unstable in terms of sample period. To test the stability of the model, a simple test has been proposed by [18].

There are three model approaches in spatial regression, Spatial Autoregressive (SAR), Spatial Error Model (SEM), and Spatial Lag X (SLX). SAR occurs when the dependent variable depends on the other neighboring dependent variables. SLX occurs when the dependent variable depends on the independent variable on neighboring units [12]. While SEM occurs when the dependent variable depends on the error results in the unit's neighbors.

Parameter model estimation for a spatial panel model uses the maximum likelihood estimator. An estimation technique has been developed based on an asymptotic approach or solving the time-varying covariate and spatially correlated errors [19-20]. A robust inference for the linear panel model regression was proposed by [21]. In the case of skewed spatial data, a Bayesian approach is employed to perform the analysis [22-23], while if the spatial outcomes are multivariate then one can use the method introduced by [24].

2.2.2 Spatial Panel Data Regression

Spatial panel data regression analyzes combined data across individuals and time series (panel data) involving spatial influences to determine the relationship between the dependent variable and several independent variables. Some research that uses panel data methods and their development can be found in [25-29]. There are two types of fixed effect models in Spatial Panel Data Regression: Spatial Lag Fixed Effect Model and Spatial error fixed effect model.

2.2.2.1 Spatial Lag Fixed Effect Model (SAR-FEM)

The spatial lag fixed effect occurs when the dependent variable depends on the observed independent variable and the dependent variable on the nearest unit. The spatial fixed effect model is stated as follows:

$$y_{it} = \alpha + \lambda \sum_{j=1}^N w_{ij} y_{jt} + \mathbf{x}'_{it} \boldsymbol{\beta} + \mu_i + \varepsilon_{it} \quad (1)$$

where \mathbf{y} is the dependent variable vector, \mathbf{X} is the independent variable matrix, λ is the spatial lag autoregression coefficient, $\boldsymbol{\beta}$ is the coefficient vector of regression parameters, w_{ij} is the spatial weight standardized, μ_i is a specific effect of individual and ε_{it} are error. The development of the models in (1) using quantile regression can be found in [30-31].

2.2.2.2 Spatial Error Fixed Effect Model (SEM-FEM)

Spatial errors in one region will depend on the results of errors in other regions on the panel data. Parameter estimation in this model uses the Maximum Likelihood Estimator (MLE). According to [17], the spatial error panel model is presented below

$$\begin{aligned} y_{it} &= \alpha + \mathbf{x}'_{it} \boldsymbol{\beta} + \mu_i + \phi_{it} \\ \phi_{it} &= \rho \sum_{j=1}^N w_{ij} \phi_{jt} + \varepsilon_{it} \end{aligned} \quad (2)$$

where \mathbf{y} is the dependent variable vector, \mathbf{X} is the independent variable matrix, ρ is the spatial error autoregression coefficient, $\boldsymbol{\beta}$ is the coefficient vector regression parameters, w_{ij} is the spatial standardized weight, μ_i is a specific effect of individual and ε_{it} is the error.

Using the concept presented previously, we present the steps for data analysis. There are seven steps beginning with the determination of the weight of the spatial matrix, and the final step is to test the model via goodness of fit test.

1. Create a spatial matrix weighting with queen continuity.
2. Perform a spatial dependency test using Moran ST.
3. Test the panel data spatial regression model using the Lagrange Multiplier test and Spatial Hausman test.
4. Estimate the parameter using the method Maximum Likelihood Estimator (MLE).
5. Perform Assumption Testing
 - (a) Spatiotemporal Non-Autocorrelation
 - (b) Homoscedasticity
 - (c) Non-Multicollinearity
 - (d) Normality Distribution
6. Conduct parameter significance test.

7. Perform goodness of fit test.

The steps for data analysis are also provided in Figure 2 below.

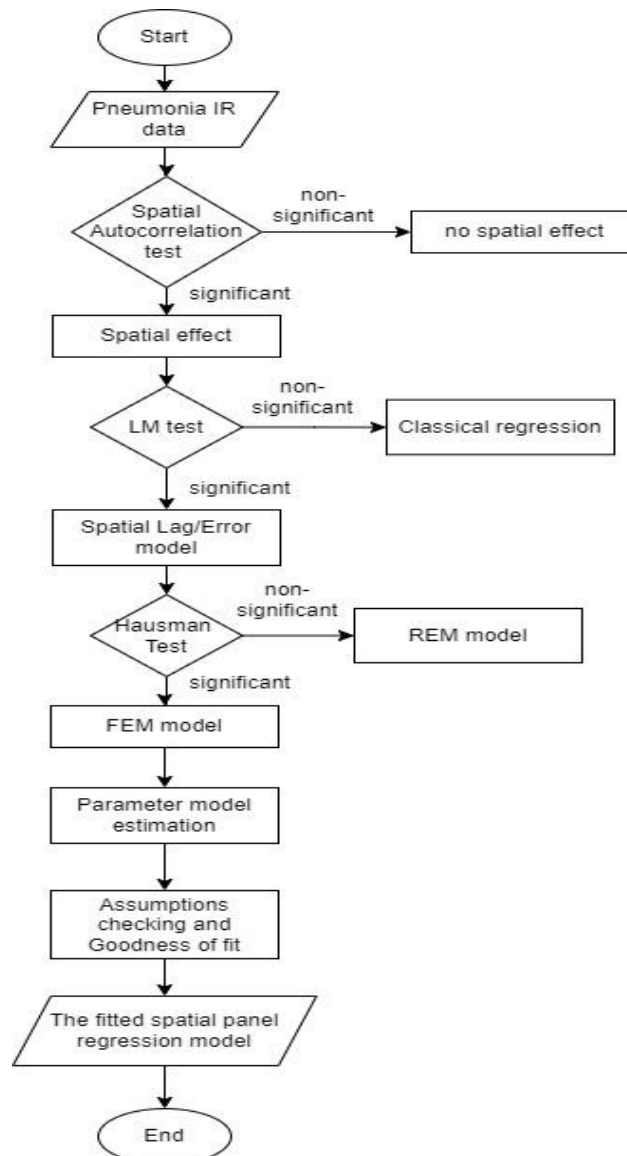


Figure 2. Flow chart of data analysis

3. Results and Discussions

There are some results obtained from this research: an incident rate map and several hypothesis testing such as spatial dependency test, Panel Data Spatial Regression model test, Hausman test, Spatiotemporal non-autocorrelation test, Homoscedasticity test, Normality Distribution Test, Parameter significance test, and goodness of fit test.

3.1 Map of the Spread of Pneumonia Incidence Rate in Purwakarta Regency

Figure 3 presents the result in the form of an incidence rates map of pneumonia in Purwakarta Regency in 2021. The incident rate for each sub-district indicated by different colors. The brighter the color, the smaller the incident rate. On the other hand, dark color regions indicate high incidence rate

of pneumonia which means high risk of toddlers contracting pneumonia. In 2021, the sub-district that is at the highest risk for toddler pneumonia is Wanayasa sub-district, while Campaka is at the lowest risk.

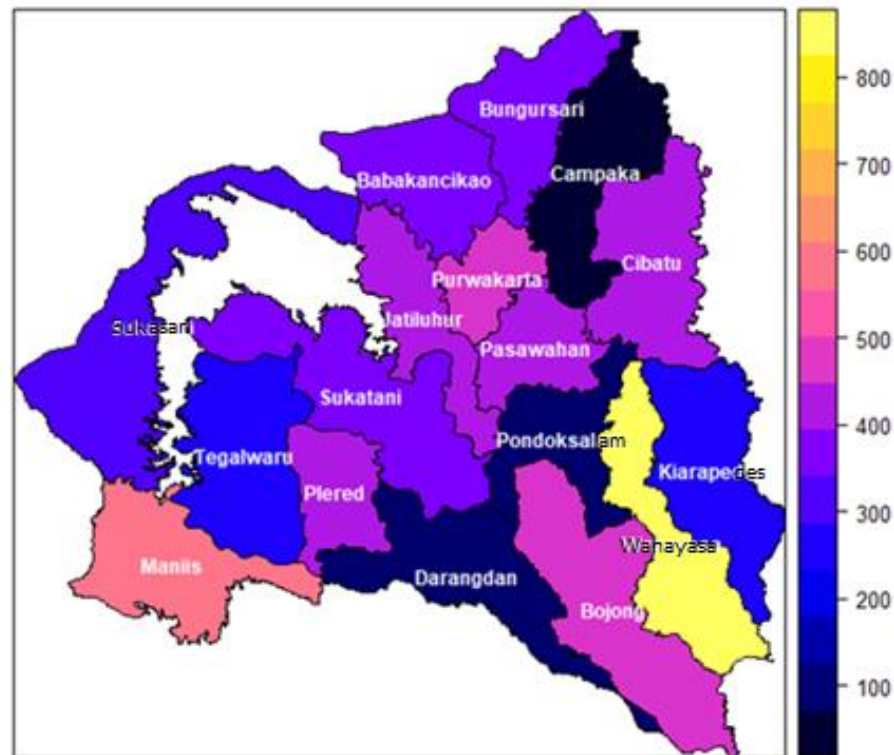


Figure 3. A map of the distribution of Toddler pneumonia In Purwakarta Regency in 2021

3.2 Spatial Weighting Matrix

Spatial weighting matrix is created to indicate the neighborhood sub-district. In our analysis, we created a spatial weighting matrix (C) with Queen Contiguity with dimensions of 17×17 as follows:

$$C = \begin{bmatrix} 0 & 0 & 1 & \dots & 0 \\ 0 & 0 & 0 & \dots & 1 \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & 0 & \dots & 0 \end{bmatrix}$$

The value in matrix C will be standardized to produce a spatial weight matrix W as below.

$$W = \begin{bmatrix} 0 & 0 & 1/4 & \dots & 0 \\ 0 & 0 & 0 & \dots & 1/3 \\ 1/3 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1/4 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 1/4 & 0 & \dots & 0 \end{bmatrix}$$

The spatial weighting matrix should be estimated in the case of unknown heteroscedasticity [32], spared spatial dependence structure [33], or when structural breaks exist [34]. In the first case, The M-estimation strategy is used and is extended for heteroscedasticity. While in the second case, the adaptive least absolute shrinkage and selection operator (LASSO) is used to select and estimate the individual nonzero connections of the spatial weight matrix. A two-stage LASSO is employed to estimate a full spatial weight matrix for the final case.

3.3 Hypothesis testing results

3.3.1 Spatial Dependency Test

The spatial dependency test was performed using software R.3.6.1 packages "Moran ST". The Moran's ST index is 0.5 with p-value of 0.01. The Moran ST index is in the range $0 < I \leq 1$, meaning that there is a positive spatial dependency between districts. The p-value of 0.01 is less than $\alpha = 0.05$ so that H_0 is rejected. In conclusion, with a 95% confidence level, there is a spatial dependency between sub-districts in Purwakarta Regency.

3.3.2 Panel Data Spatial Regression Model test

A Lagrange Multiplier (LM) test is performed to determine whether there is a spatial dependence in the lag or error. The results of LM test are shown in Table 1.

Table 1. LM test results

Model	LM	p-value	$\chi^2_{(0,05;1)}$
<i>Spatial Lag</i>	3.0356	0.0815	3.8415
<i>Spatial Error</i>	3.8529	0.0497	

Based on table 1, the statistics LM for the *Spatial Error Model* is $3.8529 > \chi^2_{(0,05;1)} = 3.8415$ and $p - value = 0.0497 < \alpha = 0.05$ then H_0 is rejected. That is, with a 95% confidence level, it can be concluded that there is a spatial dependence of the error in the regression model. On the other hand, the Spatial Lag Model is not significant. The next test is the Hausman that is used to determine the type of model to use either the spatial random or fixed effect model. The Hausman Statistic is 12.454 with p-value of 0.029. Therefore H_0 is rejected which means the resulting model follows the Fixed Effect Model.

3.3.3 Parameter estimation

Parameter estimation of the spatial error fixed effect model (SEM-FEM) was carried out using the maximum likelihood estimator (MLE). The purpose is to fit the model and to determine the contribution of factors suspected of influencing toddler pneumonia in Purwakarta Regency for three years (2019-2021). The results are shown in Table 2.

Table 2. SEM-FEM parameter estimation results

Parameter	Estimate	Standard Error	p-value
<i>Intercept</i>	743.32	147.36	4.55×10^{-07}
<i>Spatial Error coefficient</i>	0.432	0.1515	0.0044
<i>Slope Variable MR2 Measles Immunization</i>	-1.793	0.7745	0.0206
<i>Slope Variable Low Birth Weight</i>	30.167	9.7125	0.0019
<i>Slope Variable Exclusive Breastfeeding</i>	-2.808	0.9375	0.0027
<i>Slope Variable Nutritional Status</i>	3.405	5.4767	0.5341
<i>Slope Variable Clean and Healthy Behavior</i>	-4.011	1.7433	0.0214

Table 3. Spatial fixed effect for each district in Purwakarta regency

Indeks (i)	Subdistrict	μ_i
1	Babakancikao	-159.247
2	Bojong	-10.108
3	Bungursari	83.298
4	Campaka	-212.797
5	Cibatu	105.436
6	Darangdan	-246.867
7	Jatiluhur	63.388
8	Kiarapedes	-160.086
9	Maniis	-54.903
10	Pasawahan	-24.059
11	Plered	127.713
12	Pondoksalam	-189.115
13	Purwakarta	199.924
14	Sukasari	-255.715
15	Sukatani	-15.557
16	Tegal Waru	93.140
17	Wanayasa	655.555

Table 2 shows the estimate of the Fixed Effect Model parameters. According to Table 2, Measles Immunization, Low Birth Weight, Exclusive Breastfeeding, and Clean and Healthy Behavior influence the Incidence rate of toddler pneumonia while Nutritional status does not influence the Incidence rate. Based on the direction of the regression coefficient, it can be seen that the percentage of MR2 Measles Immunization, Exclusive Breastfeeding, and Cleanness and Healthy Behavior has a negative relationship. On the other hand, low birth weight has a positive relationship with the Incidence rate. Based on the results of parameter estimates in Table 2, the Spatial Error Fixed Effect Model (SAR-FEM) for toddler pneumonia can be expressed as follows.

$$\hat{y}_{it} = 743.32 + 0.432 \sum_{j=1}^{17} w_{ij} \hat{\phi}_{jt} - 1.793X_{1it} + 30.167X_{2it} - 2.808X_{3it} + 3.405X_{4it} - 4.011X_{5it} + \hat{\mu}_i \quad (3)$$

Table 3 presents the specific effect for each sub-district. Sub-district Wanayasa has the highest effect, while Sub-district Bojong has the smallest effect. The effect of each sub-district will be added to (3).

3.3.4 Spatiotemporal Non-Autocorrelation Test

Spatiotemporal non-autocorrelation testing can be done using the Moran ST index value. If the MoranST value is close to one, it indicates a solid positive Spatiotemporal autocorrelation from model error. The Moran spatiotemporal statistic is -0.152 with p-value of 0.911 meaning that the test is not significant. Therefore, using a 95% confidence level, it can be concluded that there is no Spatiotemporal autocorrelation in the model error.

3.3.5 Homoscedasticity Test

If the error variance is not constant, then the assumption of homoscedasticity is violated. Homoscedasticity testing can be done using the Breusch-Pagan test statistic. Obtained $BP = 0,663 < \chi^2_{(0,05;4)} = 9.49$ then H_0 is accepted. That is, with a 95% level of confidence it can be concluded that the error variance is homoscedastic.

3.3.6 Non-Multicollinearity Detection

Multicollinearity detection can be done by observing the matrix determinant value ($X'X$). If the determinant is positive, then there is no perfect multicollinearity. The determinant value of the matrix obtained is $\det(X'X) = 7.369 \times 10^{24}$. Thus, it can be concluded that there is no perfect multicollinearity. Furthermore, the Variance Inflation Factor (VIF) is obtained, which is used to see whether or not multicollinearity exists between predictor variables and the results are shown in Table 4.

Table 4. Multicollinearity detection

Predictor/Factor	VIF
MR2 Measles Immunization	1.051
Low Birth Weight	1.065
Exclusive Breastfeeding	1.089
Nutritional Status	1.294
Clean and Healthy Behavior	1.053

Based on Table 4, it can be seen that all predictors have no multicollinearity between the independent variables.

3.3.7 Normality Distribution Test

Normality assumption testing is carried out to see whether the errors resulting from the regression model are normally distributed. Testing the normality assumption in this study uses the Jarque-Bera (JB) test statistic. The JB statistic obtained from the analysis is 1.09 with p-value of 0.58 which means that errors are normally distributed.

3.3.8 Parameter Significance Test

To partially test the significance of the parameters in a model, namely the coefficient of spatial error autocorrelation (ρ) and the parameter coefficients of the independent variables, one can use the Wald test. The results of the test can be seen in the following table.

Table 5 Parameter Significance Test Results

Parameter	Estimate	Standard Error	Wald value	p-value
Spatial Error Coefficient (ρ)	0.432	0.152	2.850	0.002
β_{X1}	-1.793	0.775	2.315	0.010
β_{X2}	30.167	9.713	3.106	0.001
β_{X3}	-2.808	0.937	2.995	0.001
β_{X4}	3.405	5.477	0.622	0.267
β_{X5}	-4.011	1.743	2.301	0.011

Table 5 shows that the coefficient of measles immunization MR2 ($X1$), low birth weight ($X2$), exclusive breastfeeding ($X3$), and clean and healthy living behavior ($X5$) affect the incidence rate (IR) of toddler pneumonia with significant level 5%. Meanwhile, malnutrition status ($X4$) is insignificant to IR toddler pneumonia.

3.3.9 Goodness of Fit

The goodness of fit test of the incidence rate of pneumonia model yields a coefficient of determination (R^2) of 82.77%. This value indicates that the regression model can explain the immense diversity of pneumonia incidence rates which the independent variable of 82.77%. The remaining 17.23% is explained by other variables which are not included in the model.

4. Conclusions

Based on the analysis, the Spatial Error Fixed Effect Model is more appropriate for this case study rather than Spatial Error Random Effect Model. The predictors of the model such as measles immunization MR2 (X1), low birth weight (X2), exclusive breastfeeding (X3), and cleanness and healthy living behavior (X5) have significant effects on the incidence rate (IR). In addition, the map suggests that Campaka and Darangdan sub-districts have low incidence rates. In contrast, Wanayasa sub-district has the highest toddler pneumonia incidence rate.

References

- [1] B. Dadonaite and M. Roser. (2019). Pneumonia. *Our World in Data*. <https://ourworldindata.org/pneumonia>
- [2] S. N. Grief and J. K. Loza. (2028). Guidelines for the Evaluation and Treatment of Pneumonia, *Prim Care*, vol. 45, no. 3, pp. 485–503.
- [3] W. S. Lim. (2022). Pneumonia—Overview. *Encyclopedia of Respiratory Medicine*, Elsevier, pp. 185–197.
- [4] N. Rezaei and A. J. Rodriguez-Morales. (2022). Pneumonia. *IntechOpen*.
- [5] L. Huang *et al.* (2011). HIV-associated Pneumocystis pneumonia. *Proc Am Thorac Soc*, vol. 8, no. 3, pp. 294–300.
- [6] I. Govender, O. M. Maphasha, S. Rangiah, and C. Steyn. (2019). An overview of Pneumocystis jirovecii pneumonia for the African generalist practitioner. *Afr Health Sci*, vol. 19, no. 4, pp. 3200–3207.
- [7] G. Nevez, P. M. Hauser, and S. Le Gal. (2020). Pneumocystis jirovecii. *Trends in Microbiology*, vol. 28, no. 12, pp. 1034–1035.
- [8] Dinas Kesehatan Kabupaten Purwakarta. (2019). Profil Kesehatan Kabupaten Purwakarta 2019. *Dinas Kesehatan Kabupaten Purwakarta*, Purwakarta.
- [9] Dinas Kesehatan Kabupaten Purwakarta. (2020). Profil Kesehatan Kabupaten Purwakarta 2020. *Dinas Kesehatan Kabupaten Purwakarta*, Purwakarta.
- [10] Dinas Kesehatan Kabupaten Purwakarta. (2021). Profil Kesehatan Kabupaten Purwakarta 2021. *Dinas Kesehatan Kabupaten Purwakarta*, Purwakarta.
- [11] Dinas Kesehatan Provinsi Jawa Barat. (2020). Profil Kesehatan Jawa Barat Tahun 2020. *Dinas Kesehatan Provinsi Jawa barat*. www.diskes.jabarprov.go.id
- [12] J. P. Elhorst. (2014). Spatial Econometrics From Cross-Sectional Data to Spatial Panels. New York: *Springer*.
- [13] H. A. Klaiber, J. K. Abbott, and V. K. Smith. (2017). Some like it (Less) hot: Extracting trade-off measures for physically coupled amenities. *Journal of the Association of Environmental and Resource Economists*, vol. 4, no. 4, pp. 1053–1079.
- [14] M. Leiva, F. Vasquez-Lavín, and R. D. Ponce Oliva. (2020). Do immigrants increase crime? Spatial analysis in a middle-income country. *World Development*, vol. 126.
- [15] A.-T. Renner. (2020). Inefficiencies in a healthcare system with a regulatory split of power: a spatial panel data analysis of avoidable hospitalisations in Austria. *European Journal of Health Economics*, vol. 21, no. 1, pp. 85–104.
- [16] E. Biørn. (2017). *Econometrics of Panel Data*. Oxford: *Oxford University Press*.
- [17] H. Kelejian and G. Piras. (2017). *Spatial Econometrics*. *Elsevier Science*.
- [18] H. Kelejian and G. Piras. (2022). A simple test for stability of a spatial model. *Spatial Economic Analysis*, vol. 17, no. 2, pp. 245–261.
- [19] C. Jiang, D. L. Vecchia, E. Ronchetti, and O. Scaillet. (2021). Saddlepoint Approximations for Spatial Panel Data Models. *Journal of the American Statistical Association*.

-
- [20] L. Li and Z. Yang. (2021). Spatial dynamic panel data models with correlated random effects. *Journal of Econometrics*, vol. 221, no. 2, pp. 424–454.
- [21] U. K. Müller and M. W. Watson. (2022). Spatial Correlation Robust Inference in Linear Regression and Panel Models. *Journal of Business and Economic Statistics*.
- [22] S. Leorato and M. Mezzetti. (2021). A Bayesian Factor Model for Spatial Panel Data with a Separable Covariance Approach. *Bayesian Analysis*, vol. 16, no. 2. pp. 489–519.
- [23] H. Zareifard and M. Jafari Khaledi. (2021). A heterogeneous Bayesian regression model for skewed spatial data. *Spatial Statistics*, vol. 46.
- [24] H. Zareifard, M. Jafari Khaledi, and O. Dahdouh. (2019). Multivariate spatial modelling through a convolution-based skewed process. *Stochastic Environmental Research and Risk Assessment*, vol. 33, no. 3, pp. 657–671.
- [25] Sudartianto, Firman, Y. Suparman, and I. Ginanjar. (2021). A Fixed Effect Panel Spatial Error Model in Identifying Factors of Poverty in West Java Province. *Journal of Physics: Conference Series*, vol. 1776, no. 1.
- [26] M. T. Majeed and M. Mazhar. (2021). An empirical analysis of output volatility and environmental degradation: A spatial panel data approach. *Environmental and Sustainability Indicators*, vol. 10.
- [27] Y. Liu, X. Chen, and Y. Zhang. (2022). Analysis of Business Environment and Medical Insurance Coverage Rates in the Destination of China’s Migrant Population: Based on Geographically and Temporally Weighted Regression Model for Panel Data. *Mathematical Problems in Engineering*, vol. 2022.
- [28] N. Paramita, M. Masjkur, and Indahwati. (2021). Spatial Regression Model with Optimum Spatial Weighting Matrix on GRDP Data of Sulawesi Island. *Journal of Physics: Conference Series*, vol. 1863, no. 1.
- [29] S. Viriyathorn, M. Phaiyaron, P. Rueangsom, and R. Suphanchaimat. (2022). Spatial panel data analysis on the relationship between provincial economic status and enrolment in the social security scheme amongst migrant workers in Thailand, 2015–2018. *International Journal of Environmental Research and Public Health*, vol. 19, no. 1.
- [30] X. Dai, Z. Yan, M. Tian, and M. Tang. (2020). Quantile regression for general spatial panel data models with fixed effects. *Journal of Applied Statistics*, vol. 47, no. 1, pp. 45–60.
- [31] Y. Zhang, J. Jiang, and Y. Feng. (2021). Penalized quantile regression for spatial panel data with fixed effects. *Communications in Statistics - Theory and Methods*.
- [32] L. Li and Z. Yang. (2020). Estimation of fixed effects spatial dynamic panel data models with small T and unknown heteroskedasticity. *Regional Science and Urban Economics*, vol. 8.
- [33] M. S. Merk and P. Otto. (2020). Estimation of the spatial weighting matrix for regular lattice data—An adaptive lasso approach with cross-sectional resampling *Environmetrics*, vol. 33, no. 1.
- [34] P. Otto and R. Steinert. (2022). Estimation of the Spatial Weighting Matrix for Spatiotemporal Data under the Presence of Structural Breaks. *Journal of Computational and Graphical Statistics*.
-