Spatial Modeling of Covid-19 Alertness Indeks (IKK) Post-Eid Holidays in Central Java

Rezky Yayang Yakhamid [a, 1*]

[a] Politeknik Statistika STIS, Jakarta, Indonesia.
[1] 211810559@stis.ac.id

Informasi artikel

Sejarah artikel
Diterima : 5 Sept 2021
Revisi : 29 Sept 2021
Dipublikasikan : 30 Sept 2021

Kata kunci:
Indeks Kewaspadaan Covid-19
Jawa Tengah
Model Spasial Error
Regresi

A B S T R A K

Pandemi Covid-19 menyebar ke seluruh dunia termasuk Indonesia. Peningkatan kasus Covid-19 di Indonesia dipengaruhi oleh libur panjang. Libur lebaran yang disertai dengan pelanggaran mudik dan juga protokol kesehatan diduga menyebabkan penyebaran Covid-19 pasca lebaran. Penelitian ini bertujuan untuk mengetahui faktor-faktor yang signifikan mempengaruhi Indeks Kewaspadaan Covid-19 Kabupaten (IKK) di Provinsi Jawa Tengah pasca libur lebaran. Metode analisis data yang digunakan adalah analisis spasial dengan menggunakan regresi spasial. Hasil penelitian menunjukkan regresi spasial error (SEM) lebih baik daripada model lainnya. Variabel klasifikasi daerah, persentase dokter per 10.000 penduduk, dan persentase penduduk miskin, dan tingkat pengangguran terbuka berpengaruh signifikan terhadap IKK..

Keywords:
Covid-19 Alertness Index
Central Java
Spatial Error Model
Regression

A B S T R A C T

The Covid-19 pandemic has spread throughout the world, including Indonesia. The increase in Covid-19 cases in Indonesia is influenced by long holidays. Eid holidays accompanied by violations of going home and also health protocols are suspected of causing the spread of Covid-19 after Eid. This study aims to determine the factors that significantly affect the Regency/City Covid-19 Alertness Index in Central Java Province after the Eid holiday. The data analysis method used is a spatial analysis using spatial regression. The results showed that the spatial error regression (SEM) was better than the other models. The regional classification variables, the percentage of doctors per 10,000 population, and the percentage of the poor, and the open unemployment rate have a significant effect on the Covid-19 Alertness Index

Introduction

Coronavirus disease (Covid-19) or better known as the Covid-19 pandemic is thought to have first appeared in Wuhan, China in December 2019. This pandemic is caused by a new type of coronavirus known as SARS-CoV-2. This virus attacks the human respiratory system so that the initial symptoms in people affected by Covid-19 are fever and difficulty breathing. (WHO, 2020).

Coronaviruses that are exposed to humans cause respiratory tract infections (Ministry of Health, 2020). Around 20 percent of people with Covid-19 symptoms become seriously ill, critically ill, and require intensive care such as oxygen to help breathe (WHO, 2020).

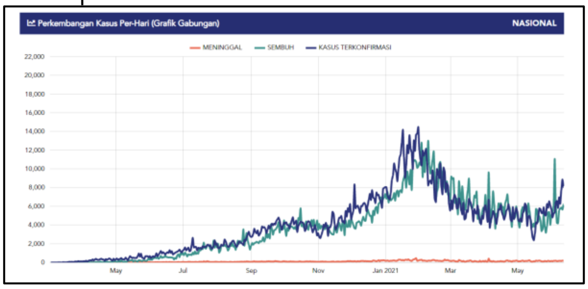
Patients with advanced age as well as people with congenital diseases of diabetes, high blood pressure, heart disease, lung disease, and cancer are usually more susceptible to getting more severe illnesses than people in general (Ministry of Health, 2020). The Covid-19 pandemic can be transmitted to anyone and can cause severe illness and even death regardless of age (WHO, 2020).

The covid-19 disease can be transmitted through droplets from the nose or mouth when a person with Covid-19 sneezes, coughs, or talks (Ministry of Health, 2020). The transmission of Covid-19 can be said to be so fast that three months since the first case in December 2019, the Covid-19 pandemic has spread to all countries on the earth's surface. Considering the level of spread and severity of the disease caused, WHO declared Covid-19 a global pandemic in March 2020 (WHO, 2020).

The Covid-19 pandemic has spread throughout the country, including Indonesia. Indonesia recorded its first case on March 2, 2020. Two Indonesian citizens, a mother, and a child who lives in Depok were found to be positive for the SARS-Cov-2 virus. According to the Minister of Health, Terawan, at the time, both were exposed to the virus from a Japanese citizen living in Malaysia (Indonesia.go.id, 2020).

After the first reporting of the Covid-19 case in Indonesia, it didn't take long for the addition of new Covid-19 cases to continue. Less than two months after the first report, on April 10, 2020, positive cases of the corona were recorded in all provinces in Indonesia (CNN Indonesia, 2020).

The rapid increase in the number of Covid-19 cases in Indonesia was caused by several things, such as violations of health protocols, the lack of 3T (Testing, Tracing, and Treatment), and long holidays. The following is a graph of the development of Covid-19 cases in Indonesia.



Source : (Covid19.go.id, 2021)

Figure 1. Graph of Daily Cases of Covid-19 in Indonesia

The data presents the development of positive confirmed cases, recovered cases, and Covid-19 death cases in Indonesia obtained from the Covid-19 Task Force's website (Covid19.go.id, 2021).

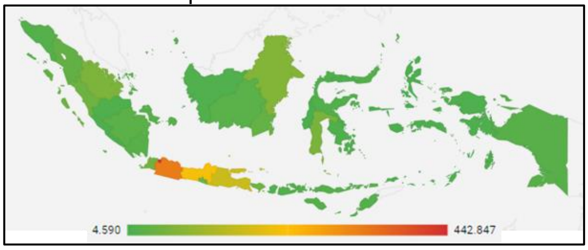
The graph shows that a spike in Covid-19 cases occurred a few weeks after the long Christmas holidays (24-27 December 2020) and the new year (31 December 2020 – 3 January 2021). This spike became the highest spike in Covid-19 cases in Indonesia, recording more than 10 thousand new cases in the range of January 8 to February 8, 2021 (Covid19.go.id, 2021).

The increase in cases after the Christmas and New Year holidays was due to the movement of people during holidays to travel both home and on vacation. Movement and crowds can have an impact on increasing Covid-19 cases. This is exacerbated by violations of health protocols which still often occur during long holidays (Indonesia.go.id, 2021).

The spike in Covid-19 cases always occurs after a long holiday. Four long holidays in 2020 and 2021 led to a 37 percent to 119 percent increase in new cases. The trend of increasing cases is also accompanied by an increase in the number of deaths. Therefore, every time there is a long holiday, the government implements a movement restriction policy such as a ban on going home (Kompaspedia, 2021).

The 2021 Eid holiday also had the impact of a spike in Covid-19 cases after that. The high mobility of the community and enthusiasm for going home even though it has been banned is one of the causes. Data as of June 1, 2021 (two weeks after Eid al-Fitr), shows an increase in new active cases by 56.6 percent and an increase in death cases due to Covid-19 by 3.52 percent (Kompaspedia, 2021).

The five provinces with the highest increase calculated from the comparison of the two weeks before and after Eid al-Fitr were Central Java (up by 103%), Riau Islands (up by 103%), Riau (up by 69%), DKI Jakarta (up by 49, 5%), and West Java (up by 25%). The following is a map diagram of the cumulative distribution of Covid-19 cases in Indonesia in 34 provinces.



Source : (KawalCOVID19, 2021)

Figure 2. Map of the Distribution of Cumulative Cases of Covid-19 per Province

From the diagram, it can be seen that the highest distribution of Covid-19 is on the island of Java, including DKI Jakarta, West Java, Central Java, and East Java. The very dangerous Covid-19 pandemic and it's very fast spread as well as the spike in Covid-19 cases after the Eid holiday, of course, really need to be investigated further, especially for Central Java Province itself which recorded the highest increase in cases after the Eid holiday.

The research, which was conducted with the Central Java Province locus and utilizing spatial data of districts/cities in Central Java, aims to obtain a spatial model of Covid-19 and determine the factors that influence the surge in Covid-19 cases in Central Java. With this research, it is hoped that the government, policymakers, and the community can act appropriately so that it can inhibit the spread of Covid-19, especially in Central Java.

Literature review

The daily positive number of Covid-19 is not enough to describe the spread of Covid-19 cases as well as the basis for Covid-19 policies. This is because the positive number of Covid-19, both cumulative and daily, is directly proportional to how active the regional Covid-19 Task Force is in preventing the spread of Covid-19. There are three aspects in preventing the Covid-19 outbreak, namely Testing, Tracing, and Treatment, or often referred to as 3T (KawalCOVID19, 2020).

One of the abilities of the regional Covid-19 Task Force in preventing the spread of Covid-19 is by looking at the tracing capabilities of an area. The measurement of the tracing ability of an area can be done by calculating the Trace Isolation Ratio (RLI). RLI is defined as the ratio of the number of people being tracked and isolated (OTG and ODP confirmed positive for Covid-19) to the number of cases confirmed positive through PCR testing in the area. The higher the RLI value, the greater the number of OTG and ODP netted through tracing compared to the number of confirmed cases of Covid-19 in the area (KawalCOVID19, 2020).

A new measure that can more accurately describe the spread of Covid-19 in an area is the Covid-19 Alertness Index (IKK) compiled by KawalCOVID19. The alert index is assumed to be a measure of potential transmission in a community that is not detected in an area based on Covid-19 data. The higher the Alertness Index, the higher the possibility of community transmission in an area (KawalCOVID19, 2020).

The Covid-19 District Alert Index (IKK) is a more precise measure of measuring the spread of Covid-19 because it does not only consider the number of positive cases but is also compiled from other indicators. The addition of the number of active positive cases, the number of PDP, and the total number of deaths increased the IKK, while the addition of the number of ODP, the total recovered and the testing ratio decreased the IKK (KawalCOVID19, 2020).

Elderly patients over the age of 60 years and patients with congenital diseases such as high blood pressure, heart and lung disease, diabetes, obesity, and cancer have a greater risk of getting more severe effects from Covid-19 (WHO, 2020). In addition, the Covid-19 disease can quickly spread through the droplets of people who have been infected with the virus. Therefore, the application of the 3M health protocol (maintaining distance, wearing masks, and washing hands) greatly affects the inhibition of the spread of Covid-19 (Ministry of Health, 2020). Massive Testing, Tracing, and Treatment (3T) are also three important aspects in preventing the spread of the Covid-19 outbreak (KawalCOVID19, 2020).

Method

Scope and Data Source

This study has a unit of analysis for all districts and cities in Central Java Province which consists of 35 districts and cities. The Central Java research locus was chosen because Central Java ranks first with the highest addition of Covid-19 cases after the Eid holiday, which is 103%.

The Covid-19 Alertness Index (IKK) data was obtained from the kawalcovid19.id page. The site contains spatial data on the Covid-19 IKK in each district/city in each district/city in Indonesia along with its indicators. Several explanatory variables such as isolation tracking ratio (RLI) and tracing ratio are also sourced from the KawalCOVID19 page.

Sources of data for the explanatory variables used are secondary data from various surveys and censuses conducted by Statistics Indonesia (BPS). The spatial data for each district/city is generally taken from the publication of Central Java Province in Figures 2021. The classification of areas categorized into cities and districts is taken from the stipulation of applicable laws which classify level II autonomous regions into municipalities and districts.

Research variable

The response variable (Y) in this study is the Covid-19 district alert index (IKK) in each district/city in Central Java as of June 10, 2021. The Covid-19 IKK was compiled by KawalCOVID19 which was compiled from several indicators related to Covid-19 such as the number of active positive cases, the number of PDP, the total number of deaths due to Covid-19, the number of ODP, the total recovered, and the testing ratio. The cutoff selection on June 10 was chosen because that date was the right moment a few weeks after the Eid holiday so that the spike in Covid-19 cases was expected to be analyzed.

The explanatory variable used in this study is based on several relevant studies and has a significant influence on Covid-19 cases. The explanatory variables used in this study include the percentage of the elderly population over 60 years based on the results of the 2020 Population Census (elderly), regional classification (class), population density (density) based on the results of the 2020 Population Census, number of nurses per 10,000 population (nurses), number of doctors per 10,000 population (doctors), life expectancy (life_exp), household sanitation and clean water, human development indeks (HDI), proverty population percentage (proverty), and Open Unemployment Rate (TPT).

Analysis Method

The analytical method used in this research is descriptive analysis and inferential analysis. Descriptive analysis is a method of analysis by collecting and presenting data so that it can provide information that can be used (Walpole, 1988). Descriptive analysis was used to determine the distribution of the Covid-19 alertness index in Central Java Province. Descriptive analysis is presented in visualization through map diagrams.

In addition, descriptive analysis can also be used in the early identification of the variables to be used. In identifying the relationship between explanatory variables and response variables, choropleth bivariate map visualization was used. The visualization will see the relationship between two variables by looking at the color in each district/city. The software used for descriptive analysis is QGIS.

The inferential analysis is an analytical method by analyzing part of the data so that conclusions can be drawn about the population (Walpole, 1988). The inferential analysis in this study is in the form of modeling to determine the factors that significantly affect the Covid-19 district alert index in Central Java. The method used in the modeling is regression analysis by adding spatial effects. The software used for inferential analysis is Geoda.

Multiple Linear Regression by OLS

Regression analysis is concerned with studying the dependence of one dependent variable (response variable) on one or more explanatory variables. The purpose of doing regression analysis is to estimate or predict the arithmetic mean (mean) of the response variable from the known explanatory variable values (Gujarati, 2003). The formula of the linear regression model can be written as follows.

$$\boldsymbol{\hat{Y}} = \boldsymbol{X}\boldsymbol{\hat{\beta}} + \boldsymbol{\varepsilon} \quad (1)$$

with

$\boldsymbol{\hat{Y}}$ = vector of response variables of size n x 1

$\boldsymbol{X}$ = matrix of explanatory variables of size n x k

$\boldsymbol{\hat{\beta}}$ = parameter estimation vector of size k x 1

$\boldsymbol{\varepsilon}$ = model residual vector of size n x 1

Estimation of regression parameters (β) using the least-squares method (ordinary least square, OLS). In the OLS method, a good estimator is an estimator that can meet the BLUE criteria (Best Linear Unbiased Estimator). To be able to produce a BLUE estimator using the OLS method, there are several assumptions so that the resulting parameter estimates are not biased and the resulting model fits the data (Gujarati, 2003). Classical assumptions that must be met include normal distribution of residuals, no multicollinearity, no autocorrelation, and no heteroscedasticity (Rawlings, Pantula, & Dickey, 2001).

Linear regression assumes that the residuals are normally distributed with a mean of zero and a variance of σ^2 . If normality is violated, then the confidence interval will be biased and lead to incorrect decisions. To determine the normality of the residual model can use a graph or a statistical test (Rawlings, Pantula, & Dickey, 2001). The normality test was carried out to check whether the residuals were normally distributed or not. Normality test in Geoda can be done by Jarque Berra test.

In addition, in linear regression, there is also the assumption of the absence of multicollinearity, heteroscedasticity, and autocorrelation. The non-multicollinearity assumption is violated if the independent variables used are related. This assumption can be checked through Variance Inflation Factor (VIF) and multicollinearity condition number. The assumption of homoscedasticity is an assumption that requires the value of the variance to be constant. Violation of the assumption of homoscedasticity in Geoda can be detected through the Breusch-Pagan and Koenker-Bassett statistical tests. The non-autocorrelation assumption will be violated if the observation errors are correlated with each other. If it is violated, then the value of the variance estimator will be biased. Autocorrelation is often violated in time series data and spatial data (Rawlings, Pantula, & Dickey, 2001).

Spatial Regression

Spatial data analysis or data analysis related to regionalism cannot be separated from the analysis of the relationship between regions themselves. If the regression analysis on spatial data is continued with the OLS model, the resulting model may be biased and inconsistent.

In spatial regression, the main requirement for its use is the existence of spatial relationships. To test the presence or absence of spatial autocorrelation, the Moran's I test and the Kelejian-Robinson test can be done statistically. The general model that can be written in the modeling with spatial regression is as follows (Anselin & Bera, 1998).

$$\boldsymbol{y} = \rho \boldsymbol{W_1 y} + \boldsymbol{X}\boldsymbol{\hat{\beta}} + \boldsymbol{\varepsilon}$$
$$\boldsymbol{\varepsilon} = \lambda \boldsymbol{W_2 \varepsilon} + \boldsymbol{\xi} \quad (2)$$

with

y : dependent variable vector
X : independent variable matrix
W : standardized spatial weighting matrix
$\hat{\boldsymbol{\beta}}$: coefficient vector of explanatory variable parameter
ρ : coefficient of spatial lag parameter on the dependent variable
$\boldsymbol{\varepsilon}$: error vector which is assumed to contain autocorrelation
λ : spatial error parameter coefficient on error $\boldsymbol{\varepsilon}$
ξ : error vector

The spatial model is a general model or the basic model of spatial regression analysis involving all spatial parameters. However, spatial analysis can follow only one model, namely the spatial lag model or the spatial error model (Anselin & Bera, 1998).

Spatial Lag Model (SAR)

The spatial lag model can also be called the spatial autoregressive model (SAR). This model combines a linear regression model with spatial lag on the response variable. This model is used when the effect of spatial linkage on the dependent variable is significant and the effect of spatial linkage error on error is not significant.

Initial diagnostics to determine the significance of spatial lag on response variables in Geoda software can use the Lagrange Multiplier test (LM test). In equation (3.2), the spatial lag model can occur if $\rho \neq 0$ and $\lambda = 0$ so that the equation becomes as follows.

$$\boldsymbol{y} = \rho \boldsymbol{W} \boldsymbol{y} + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \qquad (3)$$

In these equations, ρ**Wy** is a component of the spatial lag in the response variable while **Xβ** is a component of the classical linear regression. The estimation of model parameters can be done using the maximum likelihood method (maximum likelihood). The significance test of spatial lag coefficient (ρ) is performed by the likelihood ratio (LR) test (Anselin and Bera, 1998).

Spatial Error Model (SEM)

A spatial error model (SEM) is used if there is a spatial relationship in the error. Spatial relationships in errors are often not measurable/observable in the model in certain cases. This model is used when the effect of spatial correlation on the error is significant and the effect of spatial linkage error on the dependent variable is not significant.

Initial diagnostics to determine the significance of spatial lag on response variables in Godea software can use Moran's I test and Lagrange Multiplier (LM) test. In equation (3.2), a model of spatial lag can occur if $\rho = 0$ and $\lambda \neq 0$ so that the equation becomes as follows.

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \lambda \boldsymbol{W} \boldsymbol{\varepsilon} + \boldsymbol{\xi} \qquad (4)$$

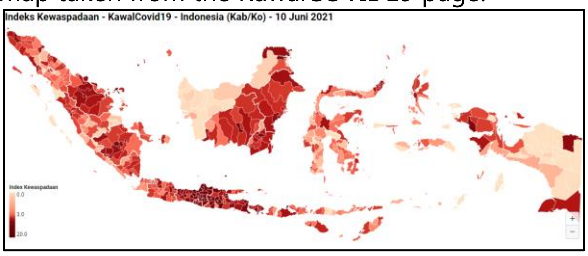
In this equation, λ**Wε** is the spatial lag component error while **Xβ** is a component of the classical linear regression. The estimation of model parameters can be done using the maximum likelihood method (maximum likelihood). The significance test of spatial error coefficient (λ) is done with the likelihood ratio (LR) test (Anselin and Bera, 1998).

Best model selection

After the formation of several spatial regression models, the next step is to select the model. The selection of the best model can be done based on the smallest Akaike's Information Criterion (AIC) value among the models that have been formed. In addition, the model selection can also consider the Bayesian Identification Criterion (BIC) or also known as the smallest Schwartz Criterion (SC), the smallest log-likelihood, and the largest R-square and adjusted R-square values (Lagona, 2011).

Results and Discussion

Indonesia is an archipelagic country whose territory is divided into several provinces and districts/cities. Although it has a fairly large area, the distribution of Indonesia's population is not evenly distributed between islands. This also has an impact on the spread of Covid-19 cases, which are dominated by certain islands. The following is a diagram of the Covid-19 district alert index (IKK) map taken from the KawalCOVID19 page.

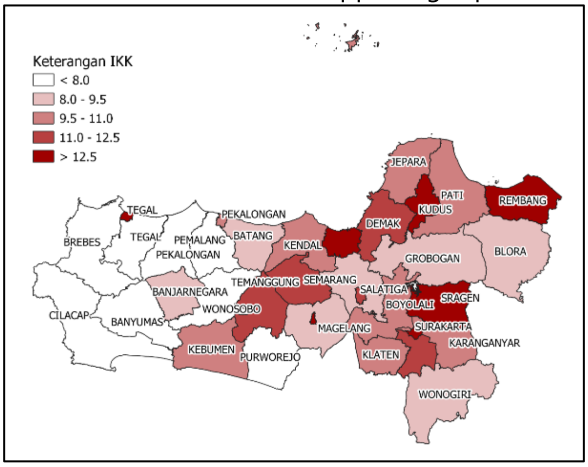


Source : KawalCOVID19 (2021)

Picture 3. Diagram of the Covid-19 Alertness Index Map Per Regency/City

In the map diagram, it can be seen that the Covid-19 alertness index with a high value marked by a solid red color is predominantly in western Indonesia, such as the islands of Sumatra, Java, and Bali, and the island of Kalimantan. The province of Central Java which is the locus in this study has a darker red color than other provinces on the island of Java.

The following is a diagram of the Covid-19 alertness index map at the district/city level in Central Java Province. The following map diagram, in addition to describing the distribution of Covid-19 cases, can also be used to identify areas with a fairly dangerous Covid-19 status because the district Covid-19 Alertness Index (IKK) has taken into account several other supporting aspects.



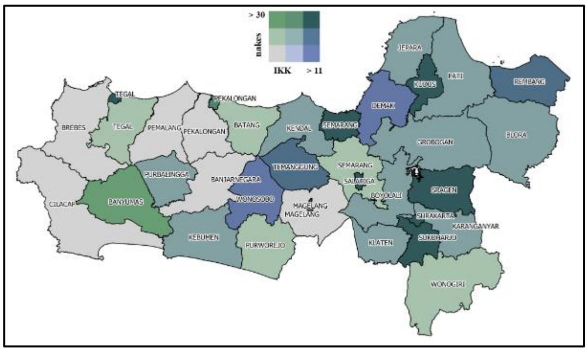
Source : KawalCovid (2021)

Picture 4. Diagram of the IKK Map in Central Java Province

In the map diagram, it can be seen that the Covid-19 IKK with high scores is spread in the eastern part of Central Java. If you look closely, there are two district/city clusters with a high Covid-19 IKK, namely the Semarang-Kudus cluster and its surroundings with distribution centers in Semarang City and Kudus Regency and the Solo Raya cluster (Surakarta and its surroundings) with distribution centers in Surakarta City and the surrounding areas. Sragen Regency.

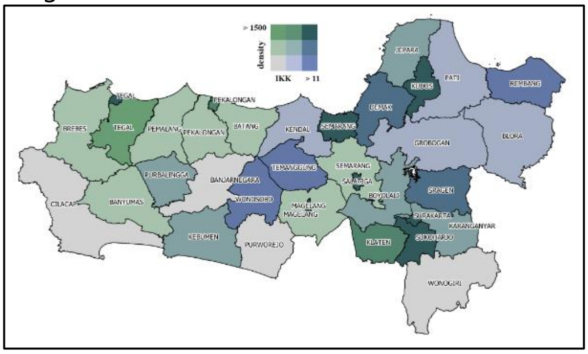
In addition, in the western part of Central Java, there is a lower IKK. This is quite interesting because if you look closely you will see clusters of districts/cities in Central Java due to differences in the Covid-19 IKK, namely the western region of Central Java with a low IKK and the eastern region of Central Java with a high IKK.

The following is a bivariate choropleth map diagram used to identify the relationship of several selected explanatory variables to the Covid-19 alertness index response variable.



Picture 5. Number of health workers by IKK

In some urban areas, a large number of health workers per 10,000 population can also imply a high Covid-19 alert index. This can be seen in several municipal areas such as Tegal City, Pekalongan City, Semarang City, Salatiga City, Magelang City, and Surakarta City. Several districts whose areas are close to the municipality also show a close relationship between the two variables. This relationship is caused by the number of health workers who are prioritized in areas with high Covid-19 cases such as big cities so that health workers are more at risk of being exposed to the dangers of Covid-19.



Picture 6. Population density by IKK

Generally, municipal areas have a higher population density than district areas. The map diagram shows that several municipal areas in Central Java, which also have a high population density, have a higher Covid-19 alert index than other regions. The diagram also shows that some areas that are far from the municipality have a low population density and a low Covid-19 alert index. Some of these districts are Cilacap, Banjarnegara, Purworejo, and Wonogiri. Generally, areas that are close to the municipality have a darker color than areas that are far away. This shows that the closer to the municipality area, the higher the population density and the Covid-19 alert index in the area.

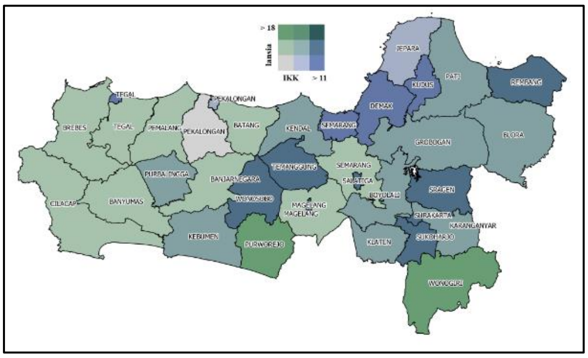


Figure 7. Percentage of the elderly population by IKK

One of the variables thought to affect the Covid-19 alertness index is the elderly. The more elderly people in an area, the higher the risk of contracting severe illness and death due to Covid-19 will also be higher. In Central Java Province as shown in the map diagram, several districts and cities that have a high percentage of elderly people and a high Covid-19 alertness index include Surakarta City, Salatiga City, Sragen, Sukoharjo Temanggung, Wonosobo, and Rembang. Some of these areas are generally located in the eastern part of Central Java.

After knowing the initial identification of the Covid-19 alertness index distribution and its relation to several explanatory variables that will be used in modeling through the map diagram, the discussion can be continued by doing spatial regression modeling. Before modeling using spatial regression, spatial weighting is first made. The spatial weighting used is queen contiguity.

After carrying out the spatial weighting, the next step is the initial identification of the occurrence of the spatial autocorrelation effect with Moran's I. The results of the Moran's I coefficient for the Covid-19 alertness index and their significance testing are as follows.

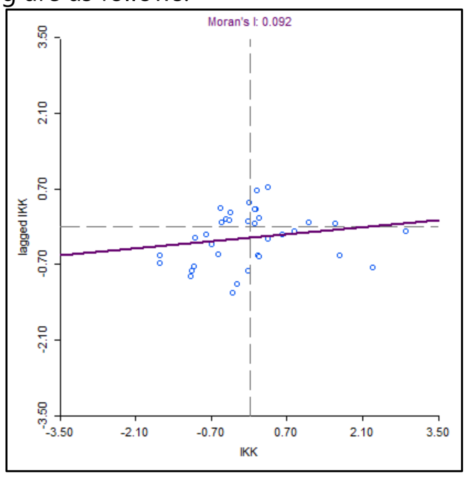


Figure 8. The plot of Moran's I

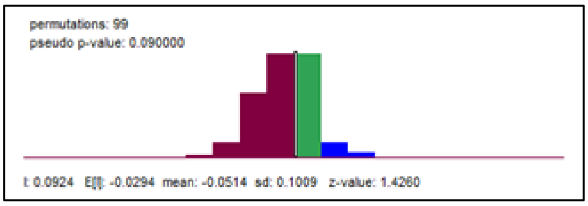


Figure 9. Moran's I Index

Based on the results of these calculations, the value of Moran's I was 0.092. A positive value indicates that the Covid-19 alertness index value in neighboring regions tends to be similar. Based on the results of statistical testing, the p-value obtained shows a significant value of 0.09 at 1% alpha. Based on these results, the modeling can proceed to the spatial dependency approach.

After conducting initial identification using Moran's I to detect the presence of spatial effects, the identification of the presence or absence of spatial lag effects, spatial error, spatial heterogeneity, and testing of classical assumptions in the regression model. The results are presented in the following table.

Tabel 1. Pengujian Efek Spasial dan Asumsi Klasik

Test	Test statistic	Score	p-value	Decision
(1)	(2)	(3)	(4)	(5)
Classical assumption test				
Multi-collinearity	Multicollinear-ity condition number	356,52	-	-
Normality	Jarque-Bera	0.7945	0.672	Fail reject Ho
Heterosce-dasticity	Breusch-Pagan	11.744	0.466	Fail reject Ho
	Koenker-Basset	11.926	0.451	Fail reject Ho
Spatial effect testing				
Lag	Lagrange Multiplier	0.8077	0.368	Fail reject Ho
	LM Robust	0.1462	0.702	Fail reject Ho
Error	Moran's I	2.0944	0.036	reject Ho
	Lagrange Multiplier	0.7168	0.397	Fail reject Ho
	LM Robust	0.0553	0.814	Fail reject Ho
SARMA	Lagrange Multiplier	0.8630	0.649	Fail reject Ho

From the results of the classical assumption test, the multicollinearity condition number shows a value of 356.5225, indicating that there is no multicollinearity in the model because the value is not in the range of 10 to 30. The statistical results of the Jarque-Bera test which fail to reject Ho indicate that the residuals meet the normal distribution. In addition, the statistical results of the Breusch-Pagan and Koenker-Basset tests which fail to reject Ho also show that the assumption of no heteroscedasticity in the model is met.

In addition, from the results of testing the effect of spatial lag, LM and LM robust test statistics show that they fail to reject Ho, which means that the spatial lag effect in the model is not significant. Therefore, the spatial lag model cannot be used. In addition, from the results of testing the effect of spatial error, Moran's I test statistic shows the results of rejecting Ho, which means that the effect of spatial error significantly affects the model. Although the other two spatial error tests are not significant, spatial error modeling can be used in this study.

Although the effect of spatial lag is not significant in the test, this study will still use the spatial lag model to be able to compare it with classical regression models and spatial error regression models. The following is the result of parameter estimation. The following three models compared are the classical regression model (OLS), the spatial lag regression model (spatial autoregression model / SAR), and the spatial error regression model (spatial error model / SEM).

Table 2. Parameters Estimation of Regression

Variabel	Parameter Estimation		
	OLS	SAR	SEM
(1)	(2)	(3)	(4)
constant	-70.1081**	-64.9183***	-69.1495***
elderly	-0.974944***	-0.975443***	-0.898221***
life_exp	0.76856**	0.705614***	0.76378***
class	7.83708**	8.22035***	8.13419***
density	-0.002167***	-0.002042***	-0.002034***
nurses	0.308254***	0.286518***	0.275018***
doctor	-3.13877***	-2.85444***	-2.76767***
tot_doctor	0.00839*	0.00941***	0.008903***
sanitation	0.112309***	0.114736***	0.112082***
clean_water	0.0895999***	0.0976587***	0.0880526***
HDI	0.301021**	0.233791**	0.266525**
proverty	0.272633	0.349814**	0.305263**
TPT	-0.679285**	-0.668603***	-0.60333**
$\boldsymbol{W}_{ikk}$	.	0.242379	.
λ	.	.	0.32515*

* significant at level 10%
** significant at level 5%
*** significant at level 1%

After estimating the parameters, the next step is to select the best model based on several existing criteria. The following table presents several modeling criteria such as R-square, Log-likelihood, AIC, and SC. These measurements can be very helpful in choosing the best model out of the three preformed models.

Table 3. Selection of Spatial Regression Model

Criteria	Criteria Value		
	OLS	SAR	SEM
(1)	(2)	(3)	(4)
R-squared	0.813112	0.822263	0.824245
Log likelihood	-64.7495	-64.1209	-64.138121
AIC	155.499	156.242	154.276
SC	175.718	178.017	174.496
Spatial effect testing			
Breusch-Pagan (heterogenity)	.	0.26516 (fail reject Ho)	11.3473 (fail reject Ho)
LR Test (spatial lag/error)	.	0.26219 (fail reject Ho)	1.2227 (fail reject Ho)

From the calculation of the model selection criteria, it can be seen that the spatial error model (SEM) has the smallest AIC and SC values. Although the resulting LR test shows a failure to reject Ho, which means there is no spatial error effect, the resulting lambda value shows a significant parameter estimate with a test level of 10%. Therefore, the SEM model can still be used to replace the classical regression model.

From table 2, the modeling of the spatial error model can be written with the following mathematical formula.

$$\boldsymbol{y} = \boldsymbol{X\beta} + \lambda \boldsymbol{W\varepsilon} + \boldsymbol{\xi}$$

$$y_i = -69.14 - 0.89(elderly) + 0.76(life\_exp) + 8.13(class) - 0.002(density) + 0.27(nurses) - 2.76(doctor) + 0.008(total\_doctor) + 0.11(sanitation) + 0.08(clean\_water) + 0.26(HDI) + 0.30(proverty) - 0.60(TPT) + 0.325W_i\varepsilon_i + \xi_i \quad (10)$$

All explanatory variables studied have a significant effect with a 5% test level on the Covid-19 alertness index. However, not all variables have an effect that is following the existing theory. Some variables that are not following the theory include the percentage of the elderly population, life expectancy, population density, the percentage of health workers, the number of doctors, the percentage of households with proper sanitation, the percentage of households with access to clean water, and the human development index. This is because the parameter estimation results are inversely proportional to the existing theory and logic. For example, the higher the percentage of the elderly population, the higher the Covid-19 alert index. However, in the parameter estimation, it was found that the coefficient value was negative, which means that the more elderly people, the lower the Covid-19 alertness index.

Several variables that have a direction of influence on the Covid-19 alertness index following theory and logic include regional classification (district/city), percentage of doctors per 10,000 population, percentage of poor population, and open unemployment rate (TPT).

Municipal areas have a Covid-19 alert index of 8.13 higher than district areas. This is because municipalities that are identical to urban areas have economic centers rather than regencies that are identical to rural areas. The status of the municipality as an economic center makes the municipality area have high population mobility, both residents in the area itself and residents who are close to the municipality.

The higher the number of doctors per 10,000 population, the Covid-19 alertness index will decrease by 2.76 points for every 1 doctor point increase per 10,000 population. This is because the role of doctors is very important in dealing with Covid-19.

The percentage of poor people has a positive influence on the Covid-19 alert index. Every one percent increase in the poor will increase the CPI by 0.305 points. This is because the poor tend to have difficulty implementing basic health protocols such as wearing masks, using hand sanitizers, and washing hands with soap due to their inability. One way to prevent the spread of Covid-19 for the poor is by distributing free masks and hand sanitizers as well as providing handwashing places in each local environmental unit and public places.

The Open Unemployment Rate (TPT) has a negative effect on the Covid-19 alert index of 0.603. This means that every one-point increase in TPT will reduce the IKK by 0.603. This is because the more unemployment there will be fewer people working so that the mobility of the population from home to work will decrease. The decrease in population mobility can play a role in preventing the transmission of Covid-19.

The model obtained can also be used to predict the Covid-19 vigilance index with the values of the explanatory variables targeted by local governments. Although many variables do not match the theory, there are several variables from the results of this study that can certainly be used by related parties such as district/city local governments to prevent the spread of Covid-19.

Conclusion

From the results and discussion, several conclusions can be drawn from this study, including: (1) there is a spatial error effect on the model so that the Spatial Error Model (SEM) with predetermined variables can be used to replace the classical regression model; (2) variable percentage of elderly, AHH, district classification, population density, health workers per 10,000 population, doctors per 10,000 population, number of doctors, percentage of households with proper sanitation, percentage of households with access to clean water, HDI, percentage of poor people, and TPT significantly affect the Covid-19 alert index; (3) several variables that have a direction of influence on the Covid-19 alertness index following theory and logic include regional classification (district/city), percentage of doctors per 10,000 population, percentage of poor population, and open unemployment rate (TPT).

As for some suggestions after the conclusions obtained in this study include: (1) further research is needed by expanding observations for example at the Java Island locus so that it can better capture causality relationships between variables following theory and logic, (2) deepen observations to the sub-district level because Spatial linkages in the spread of Covid-19 are very difficult to detect at the district/city level.

Thank-you note

Thanks to Wida Widiastuti, S.Si, M.Sc as a lecturer in Geographic Information Systems, STIS Polytechnic of Statistics who has provided knowledge of spatial analysis so that this paper can be compiled.

References

Anselin, L., & Bera, A. K. (1998). Spatial Dependence in linear Regression Models with an Introduction to Spatial Econometrics. Dalam A. Ullah, & D. E. Giles, Handbook of Applied Economic Statistics (hal. 237). CRC Press.

CNN Indonesia. (2020, April 10). Kasus Positif Corona di Indonesia Kini Ada di 34 Provinsi. Dipetik Juni 12, 2021, dari CNN Indonesia: https://www.cnnindonesia.com/nasional/20200410162438-20-492439/kasus-positif-corona-di-indonesia-kini-ada-di-34-provinsi

Covid19.go.id. (2021, Juli). Peta Sebaran Covid-19. Dipetik Juni 12, 2021, dari Covid19.go.id: https://covid19.go.id/peta-sebaran-covid19

Gujarati, D. N. (2003). Basics Econometrics Fourth Edition. New York: McGraw-Hill.

Indonesia.go.id. (2020, Maret 2). Kasus Covid-19 Pertama, Masyarakat Jangan Panik. Dipetik Juni 12, 2021, dari INDONESIA.GO.ID - Portal

Informasi Indonesia: https://indonesia.go.id/narasi/indonesia-dalam-angka/ekonomi/kasus-covid-19-pertama-masyarakat-jangan-panik

Indonesia.go.id. (2021, Januari 11). Prokes Kendor, Kasus Melonjak. Dipetik Juni 12, 2021, dari INDONESIA.GO.ID - Portal Informasi Indonesia: https://indonesia.go.id/narasi/indonesia-dalam-angka/sosial/prokes-kendor-kasus-melonjak

KawalCOVID19. (2020, Juli 04). Serial Data Virus Korona 1: 3T (Trace, Test, Treat). Dipetik Juni 12, 2021, dari KawalCOVID19: https://kawalcovid19.id/content/1288/serial-data-virus-korona-1-3t-trace-test-treat

KawalCOVID19. (2020, Juli 04). Serial Data Virus Korona 2: Rasio Lacak – Isolasi (RLI) dan Korelasinya dengan Kematian Kumulatif. Dipetik Juni 12, 2021, dari KawalCOVID19: https://kawalcovid19.id/content/1294/serial-data-virus-korona-2-rasio-lacak-isolasi-rli-dan-korelasinya-dengan-kematian-kumulatif

KawalCOVID19. (2020, Juli 09). Serial Data Virus Korona 3: Indeks Kewaspadaan. Dipetik Juni 12, 2021, dari KawalCOVID19: https://kawalcovid19.id/content/1330/serial-data-virus-korona-3-indeks-kewaspadaan

KawalCOVID19. (2021, Juni). Peta Indeks Kewaspadaan COVID-19. Dipetik Juni 12, 2021, dari KawalCOVID19: https://kcov.id/PetaKewaspadaan

KawalCOVID19. (2021, Juli). Peta Persebaran untuk Data per Provinsi. Dipetik Juli 12, 2021, dari KawalCOVID19: https://kcov.id/peta-provinsi

Kemkes. (2020, Juli 7). QnA : Pertanyaan dan Jawaban Terkait COVID-19. Dipetik Juli 12, 2021, dari Infeksi Emerging Kementerian Kesehatan RI: https://infeksiemerging.kemkes.go.id/uncategorized/qna-pertanyaan-dan-jawaban-terkait-covid-19

Kompaspedia. (2021, Juni 7). Lonjakan Kasus Covid-19 Pasca Libur Lebaran 2021. Dipetik Juni 12, 2021, dari Kompaspedia - Kompas: https://kompaspedia.kompas.id/baca/paparan-topik/lonjakan-kasus-covid-19-pasca-libur-lebaran-2021

Lagona, F. (2011). Model Selection in Markov Random Fields for High Spatial Resolution Hyperspectral Data. Dalam M. M. Fischer, & A. Getis, Handbook of Applied Spatial Analysis, Software Tools, Methods and Application. Springer Science & Business Media.

Rawlings, J. O., Pantula, S. G., & Dickey, D. A. (2001). Applied regression analysis: a research tool. Springer Science & Business Media.

Walpole, R. E. (1988). Pengantar Statistika Edisi ke-3. (B. Sumantri, Penerj.) Jakarta: Gramedia Pustaka Utama.

WHO. (2020, Oktober 12). Coronavirus disease (COVID-19). Dipetik Juli 12, 2021, dari World Health Organization: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/question-and-answers-hub/q-a-detail/coronavirus-disease-covid-19

WHO. (2020, April 27). WHO Timeline - COVID-19. Dipetik Juni 12, 2021, dari World Health Organization: https://www.who.int/news/item/27-04-2020-who-timeline---covid-19