

Spatial Modelling of Covid-19 Confirmed Cases in Kalimantan, Indonesia: How Neighborhood Matters?

Muhammad Luthfi Setiarno PUTERA^{1,*},
Laili WAHYUNITA¹ and Febrianawati YUSUP²

¹Faculty of Sharia, State Islamic Institute of Palangka Raya, Palangka Raya 73111, Indonesia

²Faculty of Education and Teaching Training, Antasari State Islamic University of Banjarmasin, Banjarmasin 70235, Indonesia

(*Corresponding author's e-mail: m.luthfi@iain-palangkaraya.ac.id)

Received: 4 March 2021, Revised: 1 May 2021, Accepted: 9 May 2021

Abstract

The Covid-19 outbreak has hit all countries across the globe, including Indonesia, in which the impact is detrimental and costly. We investigated 14 determinants that could spatially influence Covid-19 cases in Central Kalimantan and South Kalimantan provinces in mid-2020 by using the Geographically Weighted Negative Binomial Regression (GWNBR) and Mixed Geographically Weighted Negative Binomial Regression (MGWNBR). This study conducted iterative Limited-memory Broyden-Fletcher-Goldfarb-Shanno with boundaries (L-BFGS-B) to utilize the numerical parameter estimation of MGWNBR. MGWNBR identified that the adjacent regions tend to group in 8 clusters containing the same significant determinants. Through MGWNBR, the comorbid prevalences (acute respiratory infection, pneumonia, and diabetes) were positively associated with the Covid-19 increasing cases in most regions. The unemployment rate and the number of health care facilities were negatively related to the increase of Covid-19 cases in some regions. MGWNBR was better than GWNBR in terms of AIC, deviance, and pseudo R-sq. The residual map also suggested that MGWNBR produced a more accurate projection than GWNBR.

Keywords: Comorbid, Covid-19, GWNBR, Health care facilities, L-BFGS-B, MGWNBR

Introduction

The Covid-19 outbreak has been the global main issue considering the adverse effects on many sectors, ranging from health, economy to social interaction. Such an outbreak was initially identified in Wuhan, China in late 2019 [1]. Many believe that Covid-19, caused by SARS-CoV-2, is originated from bats and can easily transmit between humans. Common symptoms of Covid-19 patients are relentless cough, fever, fatigue, and loss of smell or taste [2].

In Indonesia, the Covid-19 1st case was reported on March 2, 2020. At the end of 2020, Covid-19 nationally had infected over 750 thousand patients [3]. Kalimantan, the island with the 2nd largest area in Indonesia, accounted for as much as 7 % of the total Covid-19 cases [4]. Specifically, several provinces in Kalimantan were in the public spotlight, such as Central Kalimantan and South Kalimantan provinces. The positivity rate of both provinces was more than 5 %, the maximum rate targeted by the national authority. The outbreak was not effectively controlled by the local authorities, characterized by Covid-19 confirmed cases of 9,005 in Central Kalimantan and 14,692 in South Kalimantan [5-6]. It was ignited by low public awareness in implementing health protocols. Thus, the Covid-19 cases in South Kalimantan became the 6th highest across Indonesia provinces as of June 2020 [7].

Covid-19 determinants relate to locations where neighboring areas can influence each other. Significant determinants of Covid-19 may vary between regions [8]. This allows for local models with

different estimator values that can be used to identify Covid-19 determinants in certain regions. One method used for spatial cases is Geographically Weighted Regression (GWR). The difference between GWR and linear regression lies in the estimator parameters that apply globally for linear regression, while the opposite applies locally for GWR [9].

In its development, it is common that significant local estimators can be combined with significant global estimators since not all parameters have different values between regions that some are grouped [10]. This has led our research to the usage of a mixed geographically weighted regression model. Since the Covid-19 cases are count data, the proposed spatial model is Mixed Geographically Weighted Negative Binomial Regression (MGWNBR).

The over dispersion in the count data is commonly found in epidemiological cases due to different incidence rates in each region. When over dispersion occurs, i.e., variance greater than mean, MGWNBR is better than Mixed Geographically Weighted Poisson Regression (MGWPR) [11]. Unlike MGWPR, MGWNBR has a dispersion parameter. Such parameter is aimed to capture spatial data variability.

The MGWNBR model in this study was designed using adaptive bisquare kernel weight. This weight is suitable for the spatial model in Kalimantan whose regions border irregular patterns and tend to form groups. The parameter estimation of the mixed spatial model needs an algorithm which allows the iteration fits the parameter in a fast and effective way. Unlike past research, this study used Limited-memory Broyden-Fletcher-Goldfarb-Shanno with boundaries (L-BFGS-B) iteration in estimating MGWNBR parameters. This aims to increase the accuracy of spatial models in search of determinants affecting Covid-19 in Kalimantan. Hence, the estimation processing can be achieved efficiently.

This research provides a new iterative approach using L-BFGS-B in estimating the parameters of the MGWNBR model. Given the limited number of similar studies in Indonesia, in-depth investigations are necessary to seek the determinants which affect Covid-19 spatially. Through this research, it is expected that there will be a residual map that accurately describes the performance of spatial models in investigating Covid-19 cases.

Materials and methods

Parameter estimation of MGWNBR

Mixed Geographically Weighted Negative Binomial Regression (MGWNBR) was developed based on a model similar to GWNBR, though the essential difference lies in the combination of global and local parameters in MGWNBR instead of local-only as in GWNBR. The MGWNBR data structure consisted of 1 dependent variable and several independent variables, equipped with latitude and longitude coordinate information of each involved region [11]. In this study, the estimation of the MGWNBR was conducted iteratively by using the L-BFGS-B algorithm. The weight function used in this work was an adaptive bisquare kernel based on bandwidth value selected through Generalized Cross-Validation (GCV). The process of estimation and results from mapping were executed by R version 3.5.2 and ArcMap version 10.2.1, respectively.

The following is the MGWNBR model for the i -th region, $i = 1, 2, 3, \dots, n$,

$$y_i = \exp \left(\sum_{j=0}^{p^*} \beta_j(u_i, v_i) x_{ij} + \sum_{m=p^*+1}^p \gamma_m x_{im} + \varepsilon_i \right) \quad (1)$$

where the 1st term represents the local parameters, while the 2nd term represents the global parameter. Given that Covid-19 cases are count data that does not allow for negative values of y_i , a connecting function between the estimated value and predictor is required to be linear, shown in Eq. (2),

$$\begin{aligned} \eta_i &= \ln(y_i) \\ &= \beta_0(u_i, v_i) + \beta_1(u_i, v_i) x_{1i} + \beta_2(u_i, v_i) x_{2i} + \dots \\ &\quad + \beta_{p^*}(u_i, v_i) x_{p^*i} + \gamma_m x_{mi} + \gamma_{m+1} x_{(m+1)i} + \dots \\ &\quad + \gamma_p x_{pi} + \varepsilon_i \end{aligned} \quad (2)$$

Then, by partitioning vector \mathbf{x} into vectors for global and local terms, Eq. (2) can be converted in matrix form to,

$$\ln \mathbf{y} = (\boldsymbol{\beta}(u, v) \otimes \mathbf{x}^T) \mathbf{1} + \mathbf{x}^{*T} \boldsymbol{\gamma} + \boldsymbol{\varepsilon}. \quad (3)$$

Given $E(Y_i | X_{1i} = x_{1i}, X_{2i} = x_{2i}, \dots, X_{pi} = x_{pi}) = \mu_i$ and $E(\varepsilon_i | X_i) = 0$, then the expected value of y_i is

$$\mu_i = \sum_{j=0}^{p^*} \beta_j(u_i, v_i) x_{ij} + \sum_{m=p^*+1}^p \gamma_m x_{im}. \text{ If } \mu_i \text{ is converted to a matrix, we will obtain,}$$

$$\mu_i = \exp(\mathbf{x}_i^T \boldsymbol{\beta}(u_i, v_i) + \mathbf{x}_i^{*T} \boldsymbol{\gamma}) \quad (4)$$

where 3 unknown parameters will be estimated: $\boldsymbol{\beta}(u_i, v_i)$, $\boldsymbol{\gamma}$, and dispersion parameter r_i .

The subsequent step is to obtain the likelihood function of the MGWNBR model at the i -th region which can be written as follows.

$$\begin{aligned} L(\boldsymbol{\beta}(u_i, v_i), \boldsymbol{\gamma}, r_i) &= \prod_{i=1}^n \frac{\Gamma\left(y_i + \frac{1}{r_i}\right)}{\Gamma\left(\frac{1}{r_i}\right) y_i!} \left(\frac{\mu_i}{\mu_i + \frac{1}{r_i}}\right)^{y_i} \left(\frac{\frac{1}{r_i}}{\mu_i + \frac{1}{r_i}}\right)^{\frac{1}{r_i}} \\ &= \prod_{i=1}^n \left\{ \left(\prod_{q=0}^{y_i-1} (1 + qr_i)\right) \left(\frac{1}{y_i!}\right) \left(\frac{\mu_i}{\mu_i r_i + 1}\right)^{y_i} \left(\frac{1}{\mu_i r_i + 1}\right)^{\frac{1}{r_i}} \right\} \end{aligned} \quad (5)$$

Eq. (5) is then converted into ln likelihood such that it is easier to calculate the maximum likelihood in MGWNBR parameter estimation,

$$\begin{aligned} l(\boldsymbol{\beta}(u_i, v_i), \boldsymbol{\gamma}, r_i) &= \ln L(\boldsymbol{\beta}(u_i, v_i), \boldsymbol{\gamma}, r_i) \\ &= \ln \prod_{i=1}^n \left\{ \left(\prod_{q=0}^{y_i-1} (1 + qr_i)\right) \left(\frac{1}{y_i!}\right) \left(\frac{\mu_i}{\mu_i r_i + 1}\right)^{y_i} \left(\frac{1}{\mu_i r_i + 1}\right)^{\frac{1}{r_i}} \right\} \\ &= \sum_{i=1}^n \left\{ \left(\sum_{q=0}^{y_i-1} (1 + qr_i)\right) + \ln\left(\frac{1}{y_i!}\right) + \ln\left(\frac{\mu_i}{\mu_i r_i + 1}\right)^{y_i} + \ln\left(\frac{1}{\mu_i r_i + 1}\right)^{\frac{1}{r_i}} \right\} \\ &= \sum_{i=1}^n \left\{ \left(\sum_{q=0}^{y_i-1} (1 + qr_i)\right) - \ln(y_i!) + y_i \ln \mu_i - \left(y_i + \frac{1}{r_i}\right) \ln(\mu_i r_i + 1) \right\} \end{aligned} \quad (6)$$

with $\mu_i = \exp(\beta_0(u_i, v_i) + \beta_1(u_i, v_i)x_{1i} + \dots + \beta_{p^*}(u_i, v_i)x_{p^*i} + \gamma_m x_{mi} + \dots + \gamma_p x_{pi})$. Eq. (6) is then simplified by considering:

$$\begin{aligned} a &= \left(\sum_{q=0}^{y_i-1} (1 + qr_i)\right) - \ln(y_i!), \\ b &= y_i \ln \mu_i, \end{aligned}$$

$$c = \left(y_i + \frac{1}{r_i} \right) \ln(\mu_i r_i + 1). \text{ Thus, it will produce that } l(\boldsymbol{\beta}(u_i, v_i), \boldsymbol{\gamma}, r_i) = \sum_{i=1}^n (a + b - c).$$

The difference in this study lies in the use of different bandwidths between regions. The adaptive bisquare kernel in Eq. (7) is selected since if the distance between 2 locations d_{ij} is higher than threshold h , the bandwidth value will be 0 [12].

$$w_j(u_i, v_i) = \left(1 - \left(\frac{d_{ij}}{h_i} \right)^2 \right)^2. \tag{7}$$

Then, with the role of weighting in Eq. (7), the ln likelihood of GWNBR becomes,

$$l^*(\boldsymbol{\beta}(u_i, v_i), \boldsymbol{\gamma}, r_i) = \sum_{i=1}^n w_{ij}(u_i, v_i)(a + b - c). \tag{8}$$

The next step is to seek the 1st derivative of Eq. (8) for each parameter to form the gradient vector. The following is the 1st derivative of $\beta_0(u_i, v_i)$;

$$\frac{\partial l^*(\boldsymbol{\beta}(u_i, v_i), \boldsymbol{\gamma}, r_i)}{\partial \beta_0(u_i, v_i)} = \sum_{i=1}^n w_{ij}(u_i, v_i) \left(\frac{\partial a}{\partial \beta_0(u_i, v_i)} + \frac{\partial b}{\partial \beta_0(u_i, v_i)} - \frac{\partial c}{\partial \beta_0(u_i, v_i)} \right). \tag{9}$$

It is obtained that;

$$\begin{aligned} \frac{\partial a}{\partial \beta_0(u_i, v_i)} &= \frac{\partial}{\partial \beta_0(u_i, v_i)} \left(\sum_{q=0}^{y_i-1} (1 + qr_i) - \ln(y_i!) \right) = 0, \\ \frac{\partial b}{\partial \beta_0(u_i, v_i)} &= \frac{\partial}{\partial \beta_0(u_i, v_i)} (y_i \ln \mu_i) = y_i, \\ \frac{\partial c}{\partial \beta_0(u_i, v_i)} &= \frac{\partial}{\partial \beta_0(u_i, v_i)} \left(\left(y_i + \frac{1}{r_i} \right) \ln(\mu_i r_i + 1) \right) \\ &= \left(y_i + \frac{1}{r_i} \right) \frac{\mu_i r_i}{\mu_i r_i + 1}. \end{aligned}$$

Equation (9) is then simplified to,

$$\begin{aligned} \frac{\partial l^*(\boldsymbol{\beta}(u_i, v_i), \boldsymbol{\gamma}, r_i)}{\partial \beta_0(u_i, v_i)} &= \sum_{i=1}^n w_{ij}(u_i, v_i) \left(\frac{\partial a}{\partial \beta_0(u_i, v_i)} + \frac{\partial b}{\partial \beta_0(u_i, v_i)} - \frac{\partial c}{\partial \beta_0(u_i, v_i)} \right) \\ &= \sum_{i=1}^n w_{ij}(u_i, v_i) \left(0 + y_i - \left(y_i + \frac{1}{r_i} \right) \frac{\mu_i r_i}{\mu_i r_i + 1} \right) \\ &= \sum_{i=1}^n w_{ij}(u_i, v_i) \left(\frac{y_i - \mu_i}{\mu_i r_i + 1} \right). \end{aligned} \tag{10}$$

With similar steps as in Eq. (10), the 1st derivative of $l^*(\beta(u_i, v_i), \gamma, r_i)$ with respect to other $\beta(u_i, v_i)$ can be calculated.

The next step is to find the 1st derivative of Eq. (8) with respect to global parameter γ , for example γ_m ,

$$\frac{\partial l^*(\beta(u_i, v_i), \gamma, r_i)}{\partial \gamma_m} = \sum_{i=1}^n w_{ij}(u_i, v_i) \left(\frac{\partial a}{\partial \gamma_m} + \frac{\partial b}{\partial \gamma_m} - \frac{\partial c}{\partial \gamma_m} \right). \quad (11)$$

It is shown that,

$$\begin{aligned} \frac{\partial a}{\partial \gamma_m} &= \frac{\partial}{\partial \gamma_m} \left(\sum_{q=0}^{y_i-1} (1 + qr_i) \right) - \ln(y_i!) = 0, \\ \frac{\partial b}{\partial \gamma_m} &= \frac{\partial}{\partial \gamma_m} (y_i \ln \mu_i) = y_i x_{mi}, \\ \frac{\partial c}{\partial \gamma_m} &= \frac{\partial}{\partial \gamma_m} \left(\left(y_i + \frac{1}{r_i} \right) \ln(\mu_i r_i + 1) \right) \\ &= \left(y_i + \frac{1}{r_i} \right) \frac{\mu_i r_i x_{mi}}{\mu_i r_i + 1}. \end{aligned}$$

Equation (11) then will be shown as,

$$\begin{aligned} \frac{\partial l^*(\beta(u_i, v_i), \gamma, r_i)}{\partial \gamma_m} &= \sum_{i=1}^n w_{ij}(u_i, v_i) \left(\frac{\partial a}{\partial \gamma_m} + \frac{\partial b}{\partial \gamma_m} - \frac{\partial c}{\partial \gamma_m} \right) \\ &= \sum_{i=1}^n w_{ij}(u_i, v_i) \left(0 + y_i x_{mi} - \left(y_i + \frac{1}{r_i} \right) \frac{\mu_i r_i x_{mi}}{\mu_i r_i + 1} \right) \\ &= \sum_{i=1}^n w_{ij}(u_i, v_i) \left(\frac{x_{mi}(y_i - \mu_i)}{\mu_i r_i + 1} \right). \end{aligned} \quad (12)$$

Alike step on Eq. (12), the 1st derivative of $l^*(\beta(u_i, v_i), \gamma, r_i)$ with respect to other γ can be calculated.

Then, the step taken is to determine the 1st derivative of the dispersion parameter r_i ,

$$\frac{\partial l^*(\beta(u_i, v_i), \gamma, r_i)}{\partial r_i} = \sum_{i=1}^n w_{ij}(u_i, v_i) \left(\frac{\partial a}{\partial r_i} + \frac{\partial b}{\partial r_i} - \frac{\partial c}{\partial r_i} \right). \quad (13)$$

It is shown that,

$$\begin{aligned} \frac{\partial a}{\partial r_i} &= \frac{\partial}{\partial r_i} \left(\sum_{q=0}^{y_i-1} (1 + qr_i) \right) - \ln(y_i!) = \sum_{q=0}^{y_i-1} \left(\frac{q}{1 + qr_i} \right), \\ \frac{\partial b}{\partial r_i} &= \frac{\partial}{\partial r_i} (y_i \ln \mu_i) = 0, \end{aligned}$$

$$\begin{aligned} \frac{\partial c}{\partial r_i} &= \frac{\partial}{\partial r_i} \left(\left(y_i + \frac{1}{r_i} \right) \ln(\mu_i r_i + 1) \right) \\ &= -\frac{1}{r_i^2} \ln(\mu_i r_i + 1) + \left(y_i + \frac{1}{r_i} \right) \frac{\mu_i}{\mu_i r_i + 1}. \end{aligned}$$

Thus, Eq. (13) will be,

$$\begin{aligned} \frac{\partial l^*(\boldsymbol{\beta}(u_i, v_i), \boldsymbol{\gamma}, r_i)}{\partial r_i} &= \sum_{i=1}^n w_{ij}(u_i, v_i) \left(\frac{\partial a}{\partial r_i} + \frac{\partial b}{\partial r_i} - \frac{\partial c}{\partial r_i} \right) \\ &= \sum_{i=1}^n w_{ij}(u_i, v_i) \left(\sum_{q=0}^{y_i-1} \frac{q}{1+qr_i} \right) + \frac{1}{r_i^2} \ln(\mu_i r_i + 1) - \left(y_i + \frac{1}{r_i} \right) \frac{\mu_i}{\mu_i r_i + 1}. \end{aligned} \tag{14}$$

All partial derivatives, as shown in Eqs. (10), (12), and (14), can be arranged into gradient vectors as follows,

$$\nabla g = \begin{bmatrix} \frac{\partial l^*(\boldsymbol{\beta}(u_i, v_i), \boldsymbol{\gamma}, r_i)}{\partial \beta_0(u_i, v_i)} \\ \frac{\partial l^*(\boldsymbol{\beta}(u_i, v_i), \boldsymbol{\gamma}, r_i)}{\partial \beta_1(u_i, v_i)} \\ \vdots \\ \frac{\partial l^*(\boldsymbol{\beta}(u_i, v_i), \boldsymbol{\gamma}, r_i)}{\partial \gamma_m} \\ \vdots \\ \frac{\partial l^*(\boldsymbol{\beta}(u_i, v_i), \boldsymbol{\gamma}, r_i)}{\partial r_i} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n w_{ij}(u_i, v_i) \left(\frac{y_i - \mu_i}{\mu_i r_i + 1} \right) \\ \sum_{i=1}^n w_{ij}(u_i, v_i) \left(\frac{x_{i1}(y_i - \mu_i)}{\mu_i r_i + 1} \right) \\ \vdots \\ \sum_{i=1}^n w_{ij}(u_i, v_i) \left(\frac{x_{im}(y_i - \mu_i)}{\mu_i r_i + 1} \right) \\ \vdots \\ \sum_{i=1}^n w_{ij}(u_i, v_i) s \end{bmatrix} \tag{15}$$

with $s = \left(\sum_{q=0}^{y_i-1} \frac{q}{1+qr_i} \right) + \frac{1}{r_i^2} \ln(\mu_i r_i + 1) - \left(y_i + \frac{1}{r_i} \right) \frac{\mu_i}{\mu_i r_i + 1}$.

Numerical parameter estimation using the iterative L-BFGS-B method requires a Hessian matrix containing the 2nd derivatives of ln likelihood in Eq. (6) [13]. The following is the Hessian matrix used in this study with $\mathbf{z} = (\boldsymbol{\beta}(u_i, v_i), \boldsymbol{\gamma}, r_i)$.

$$H(\mathbf{z}) = \begin{pmatrix} \frac{\partial l^*(\mathbf{z})}{\partial \beta_0 \partial \beta_0} & \frac{\partial l^*(\mathbf{z})}{\partial \beta_1 \partial \beta_0} & \dots & \frac{\partial l^*(\mathbf{z})}{\partial \beta_{p^*} \partial \beta_0} & \frac{\partial l^*(\mathbf{z})}{\partial \gamma_m \partial \beta_0} & \dots & \frac{\partial l^*(\mathbf{z})}{\partial \gamma_p \partial \beta_0} & \frac{\partial l^*(\mathbf{z})}{\partial r_i \partial \beta_0} \\ \frac{\partial l^*(\mathbf{z})}{\partial \beta_0 \partial \beta_1} & \frac{\partial l^*(\mathbf{z})}{\partial \beta_1 \partial \beta_1} & \dots & \frac{\partial l^*(\mathbf{z})}{\partial \beta_{p^*} \partial \beta_1} & \frac{\partial l^*(\mathbf{z})}{\partial \gamma_m \partial \beta_1} & \dots & \frac{\partial l^*(\mathbf{z})}{\partial \gamma_p \partial \beta_1} & \frac{\partial l^*(\mathbf{z})}{\partial r_i \partial \beta_1} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{\partial l^*(\mathbf{z})}{\partial \beta_0 \partial \beta_{p^*}} & \frac{\partial l^*(\mathbf{z})}{\partial \beta_1 \partial \beta_{p^*}} & \dots & \frac{\partial l^*(\mathbf{z})}{\partial \beta_{p^*} \partial \beta_{p^*}} & \frac{\partial l^*(\mathbf{z})}{\partial \gamma_m \partial \beta_{p^*}} & \dots & \frac{\partial l^*(\mathbf{z})}{\partial \gamma_p \partial \beta_{p^*}} & \frac{\partial l^*(\mathbf{z})}{\partial r_i \partial \beta_{p^*}} \\ \frac{\partial l^*(\mathbf{z})}{\partial \beta_0 \partial \gamma_m} & \frac{\partial l^*(\mathbf{z})}{\partial \beta_1 \partial \gamma_m} & \dots & \frac{\partial l^*(\mathbf{z})}{\partial \beta_{p^*} \partial \gamma_m} & \frac{\partial l^*(\mathbf{z})}{\partial \gamma_m \partial \gamma_m} & \dots & \frac{\partial l^*(\mathbf{z})}{\partial \gamma_p \partial \gamma_m} & \frac{\partial l^*(\mathbf{z})}{\partial r_i \partial \gamma_m} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{\partial l^*(\mathbf{z})}{\partial \beta_0 \partial \gamma_p} & \frac{\partial l^*(\mathbf{z})}{\partial \beta_1 \partial \gamma_p} & \dots & \frac{\partial l^*(\mathbf{z})}{\partial \beta_{p^*} \partial \gamma_p} & \frac{\partial l^*(\mathbf{z})}{\partial \gamma_m \partial \gamma_p} & \dots & \frac{\partial l^*(\mathbf{z})}{\partial \gamma_p \partial \gamma_p} & \frac{\partial l^*(\mathbf{z})}{\partial r_i \partial \gamma_p} \\ \frac{\partial l^*(\mathbf{z})}{\partial \beta_0 \partial r_i} & \frac{\partial l^*(\mathbf{z})}{\partial \beta_1 \partial r_i} & \dots & \frac{\partial l^*(\mathbf{z})}{\partial \beta_{p^*} \partial r_i} & \frac{\partial l^*(\mathbf{z})}{\partial \gamma_m \partial r_i} & \dots & \frac{\partial l^*(\mathbf{z})}{\partial \gamma_p \partial r_i} & \frac{\partial l^*(\mathbf{z})}{\partial r_i \partial r_i} \end{pmatrix} \quad (16)$$

The matrix $H(\mathbf{z})$ contains the results of the 2nd derivative of the MGWNBR In-likelihood, among others:

1. For $\frac{\partial l^*(\mathbf{z})}{\partial \beta_0 \partial \beta_0}$, it is shown that

$$\begin{aligned} \frac{\partial l^*(\mathbf{z})}{\partial \beta_0^2(u_i, v_i)} &= \frac{\partial}{\partial \beta_0(u_i, v_i)} \sum_{i=1}^n w_{ij}(u_i, v_i) \left(\frac{y_i - \mu_i}{\mu_i r_i + 1} \right) \\ &= \sum_{i=1}^n w_{ij}(u_i, v_i) \left(\frac{-\mu_i(\mu_i r_i + 1) - (y_i - \mu_i)\mu_i r_i}{(\mu_i r_i + 1)^2} \right) \\ &= \sum_{i=1}^n w_{ij}(u_i, v_i) \left(\frac{-\mu_i(y_i r_i + 1)}{(\mu_i r_i + 1)^2} \right). \end{aligned}$$

2. For $\frac{\partial l^*(\mathbf{z})}{\partial r_i \partial \beta_0}$, it is shown that

$$\begin{aligned} \frac{\partial l^*(\mathbf{z})}{\partial r_i \partial \beta_0(u_i, v_i)} &= \frac{\partial}{\partial r_i} \sum_{i=1}^n w_{ij}(u_i, v_i) \left(\frac{y_i - \mu_i}{\mu_i r_i + 1} \right) \\ &= \sum_{i=1}^n w_{ij}(u_i, v_i) \left(\frac{0 - (y_i - \mu_i)\mu_i}{(\mu_i r_i + 1)^2} \right) \\ &= \sum_{i=1}^n w_{ij}(u_i, v_i) \left(\frac{-\mu_i(y_i - \mu_i)}{(\mu_i r_i + 1)^2} \right). \end{aligned}$$

With a similar process of derivative, other derivatives that makeup $H(\mathbf{z})$ will be obtained.

L-BFGS-B algorithm

The L-BFGS-B (Limited-memory Broyden-Fletcher-Goldfarb-Shanno with boundaries) is an algorithm intended for non-linear optimization problems, whether constrained or not [14]. In this study, the L-BFGS-B algorithm is used to avoid an infinite estimator value, in addition to the main objective of achieving efficient computation. In this case, it is necessary to limit the estimator value based on Eq. (8) as the foundation for forming the Hessian matrix.

These are the stages of the L-BFGS-B per 1 iteration [15]:

1. Calculate Cauchy point estimate,

$$2. \Phi(\lambda) = m(\mathbf{z}(\lambda)) = m(P(\mathbf{z}_k - \lambda \mathbf{g}_k)), \tag{17}$$

where $m(\mathbf{z})$ in Eq. (16) is calculated based on;

$$m(\mathbf{z}) = f(\mathbf{z}_k) + \nabla \mathbf{g}_k^T (\mathbf{z} - \mathbf{z}_k) + (\mathbf{z} - \mathbf{z}_k)^T \mathbf{H}_k (\mathbf{z} - \mathbf{z}_k) / 2. \tag{18}$$

\mathbf{H}_k is the Hessian matrix in Eq. (16). The goal is to obtain \mathbf{z} which minimizes Eq. (18) such that $\|\mathbf{z} - \mathbf{z}_k\| \leq \Delta$ where Δ is the radius of the trust region.

3. Identify the active set $A(\mathbf{z})$ and in-active set $I(\mathbf{z})$ where $A(\mathbf{z})$ is obtained based on the gradient projection method. The projection on P is carried out with respect to Ω for certain \mathbf{z} until each z_i meets,

$$4. P(z_i) = \begin{cases} l_i & \text{if } z_i \leq l_i, \\ z_i & \text{if } l_i \leq z_i \leq u_i, \\ u_i & \text{if } z_i \geq u_i, \end{cases} \tag{19}$$

where l_i and u_i are upper bound and lower bound, respectively.

5. Minimize Eq. (18) for unconstrained variables.

6. Apply the linear search method to obtain minimum $f(\mathbf{z})$.

7. Update the Hessian \mathbf{H}_k based on \mathbf{H}_0 \hat{m} times such that it obtains \mathbf{H}_{k+1} with $\hat{m} = \min(k+1, m)$. The higher the constant m , the more consistent the \mathbf{H}_0 is used to update \mathbf{H}_{k+1} . The constant m for L-BFGS-B is usually 5 at most as convergence is generally reached at this point.

8. Calculate the updated \mathbf{H}_{k+1} using the following;

9.

$$\mathbf{H}_{k+1} = \mathbf{V}_k^T \mathbf{H}_k \mathbf{V}_k + \rho_k \mathbf{s}_k \mathbf{s}_k^T \tag{20}$$

where $\rho_k = \frac{1}{\mathbf{y}_k^T \mathbf{s}_k}$; $\mathbf{V}_k = \mathbf{I} - \rho_k \mathbf{y}_k \mathbf{s}_k^T$; $\{\mathbf{s}_i, \mathbf{y}_i\} = \{\mathbf{z}_{i+1} - \mathbf{z}_i, \nabla g(\mathbf{z}_{i+1}) - \nabla g(\mathbf{z}_i)\}, i = k - \hat{m} + 1, \dots, k$.

10. Return to step 1 and then repeat the iteration for $k + 1$ until convergence is achieved that satisfies $\|\nabla \mathbf{g}_{k+1} - \nabla \mathbf{g}_k\| < \varepsilon$ where ε is a very small number.

Research method

The regency-level data set was assembled from Statistics Indonesia, Primary Health Research of Indonesia’s Ministry of Health, and Indonesian Covid-19 Data of Ministry of Research and Technology. The total observation included 27 cities/regencies located in Central Kalimantan and South Kalimantan, shown in **Figure 1**.

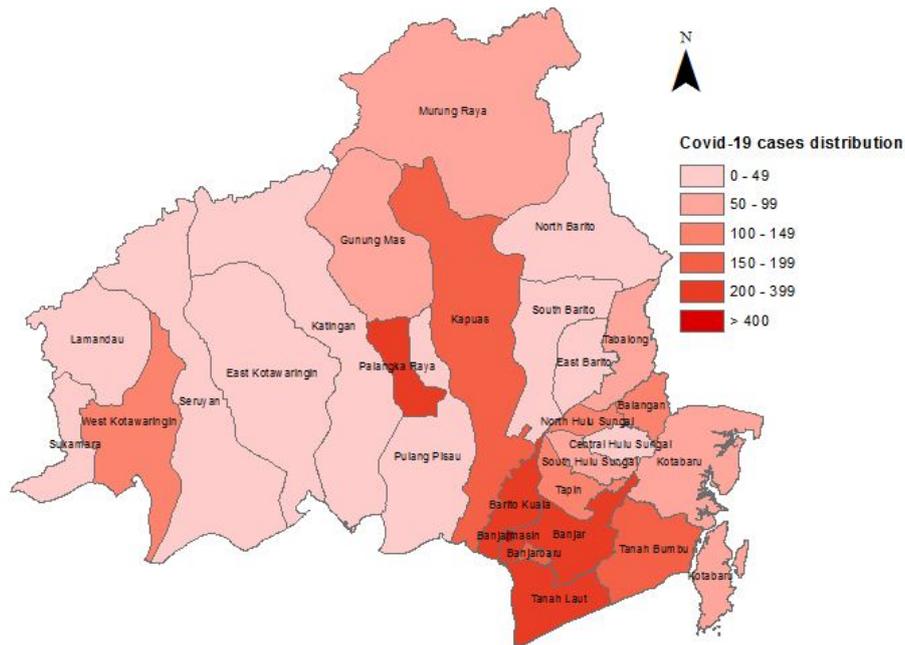


Figure 1 The Covid-19 cases distribution in Central and South Kalimantan.

This study involved 1 response variable (y) and 14 determinants (x). The focus of time was when the Covid-19 spread in both regions starting from March 2020 to June 2020. The response variable was the Covid-19 cases calculated cumulatively until the end of June 2020. The variables described in this study is shown in **Table 1**.

Table 1 The description of variables.

Variables	Description
Y	Number of Covid-19 confirmed cases
x_1	Acute respiratory infection (ARI) prevalence
x_2	Pneumonia prevalence
x_3	Diabetes prevalence
x_4	Hypertension prevalence
x_5	Proportion of correct handwashing behavior, in %
x_6	Number of health care facilities
x_7	Number of doctors
x_8	Number of nurses
x_9	Sex ratio, in %
x_{10}	Labor force participation rate
x_{11}	Unemployment rate
x_{12}	Income per capita, in IDR
x_{13}	Number of populations, transformed by log
x_{14}	Percentage of population ≥ 65 years old

This study presents a comparison between the count data global regression and the spatial regression. In the early stage, description analysis and examination of the strength of the relationship were carried out by performing Pearson correlation analysis. To obtain the significant determinants free of multicollinearity, the y and x modeling was carried out through Ordinary Least Squares (OLS) regression. Variables that revealed multicollinearity (seen from $VIF > 10$) were eliminated from the model and the regression was repeated until we obtained x which was all considered significant at $\alpha = 5\%$.

It was, then, continued with Poisson regression and negative binomial regression based on x variables which were multicollinearity-free and significant at $\alpha = 5\%$. After that, to test the existence of spatial effect on Covid-19 cases, the significance of spatial dependence was tested by Moran's I, and that of spatial heterogeneity was tested by Breusch-Pagan, both at $\alpha = 5\%$.

The significance of both global models was subsequently compared to that of the GWNBR and MGWNBR models. The parameter estimation of the MGWNBR model was examined by using the L-BFGS algorithm, as in Eqs. (17) - (20). The weight function used was as depicted in Eq. (7). The performance was, then, compared with respect to the deviance, AIC, and pseudo R-sq. The residual map was also observed to provide a better insight into the spatial model performance.

Results and discussion

The description of y shows that the 2 provinces have a significant difference in the number of cases where the Covid-19 confirmation in Central Kalimantan and South Kalimantan is 904 and 3,108 cases, respectively, as of June 2020. The region with the highest number of cases in Central Kalimantan was Palangka Raya City with 354 cases. While in South Kalimantan, the region with the highest cases was Banjarmasin City with 1,380 cases. The relationship between the number of Covid-19 confirmed cases (y) and 14 determinants, as indicated in **Table 2**, led to a significant positive relationship between y and x at $\alpha = 5\%$, including diabetes prevalence, number of health care facilities, and population density.

Table 2 Summary of Pearson correlation between y and x .

Variables	Correlation	Variables	Correlation
ARI prevalence	-0.33	Num. of nurses	0.88**
Pneumonia prevalence	0.07	Sex ratio	-0.37
Diabetes prevalence	0.55**	Labor force participation rate	-0.33
Hypertension prevalence	0.31	Unemployment rate	0.54**
% of correct handwashing behavior	0.25	Income per capita	0.56**
Num. of health care facilities	0.58**	Log-population	0.60**
Num. of doctors	0.94**	% of the population ≥ 65 years old	0.01

*significant at p -value < 0.05 , ** significant at p -value < 0.01

The result of OLS regression in **Table 3** indicates that several determinants experience multicollinearity. To overcome this, the number of nurses and the number of doctors were eliminated. After the OLS regression is carried out again without multicollinearity variables, there are 2 parameters significant at $\alpha = 5\%$. Those significant parameters are pneumonia prevalence and unemployment rate. However, it is necessary to have a comparison to determine the goodness of the global (Poisson and negative binomial), local (GWNBR), and MGWNBR regression models.

Table 3 Summary of OLS regression.

Predictors	Before eliminating multicollinearity variables			After eliminating multicollinearity variables		
	Estimate	p-value	VIF	Estimate	p-value	VIF
Intercept	1,155	0.28	-	71	0.98	
ARI prevalence	11.44	0.11	2.98	15.2	0.36	2.67
Pneumonia prevalence	55.80	0.18	2.60	220.4	0.02*	1.70
Diabetes prevalence	64.20	0.45	6.46	289	0.06	3.07
Hypertension prevalence	6.22	0.10	2.76	13.15	0.14	2.58
% of correct handwashing behavior	0.72	0.64	2.85	0.01	0.99	2.84
Num. of health care facilities	-0.98	0.78	6.59	11.47	0.15	5.11
Num. of doctors	2.59	0.00**	25.48	-	-	-
Num. of nurses	0.17	0.31	14.76	-	-	-
Sex ratio	-10.94	0.06	4.50	-13.90	0.30	4.30
Labor force rate	6.27	0.36	2.70	10.40	0.54	2.67
Unemployment rate	18.70	0.36	4.88	109.90	0.02*	3.48
Income per capita	-0.04	0.05	3.80	0.01	0.87	3.43
Log-population	-133.00	0.36	7.83	-240.00	0.44	5.86
% of population ≥ 65 years old	-2.70	0.88	4.67	3.30	0.94	4.44

* p -value < 0.05, ** p -value < 0.01

The global regression results, as shown in **Table 4**, show that negative binomial regression is better than Poisson regression. It can be seen from the lower AIC of the former. Besides, its deviance is also smaller. The ratio of the deviance of the negative binomial model divided by the degrees of freedom yields 1.94. It is much lower than that of the Poisson model of 21.18. This proves that the negative binomial model can overcome the over dispersion in the Poisson model, though the ratio still shows that the overdispersion has not been completely resolved.

Table 4 Summary of Poisson and negative binomial regression.

Parameters	Poisson regression		Negative binomial regression	
	Estimate	p-value	Estimate	p-value
Intercept	4.74	0.00**	9.41	0.12
ARI prevalence	0.06	0.00**	0.09	0.04*
Pneumonia prevalence	0.70	0.00**	0.77	0.00**
Diabetes prevalence	0.99	0.00**	0.84	0.03*
Hypertension prevalence	0.02	0.00**	0.00	0.99
% of correct handwashing behavior	-0.02	0.00**	-0.01	0.48

Parameters	Poisson regression		Negative binomial regression	
	Estimate	p-value	Estimate	p-value
Num. of health care facilities	-0.02	0.00**	-0.01	0.78
Sex ratio	-0.13	0.00**	-0.16	0.00**
Labor force rate	-0.08	0.00**	-0.08	0.10
Unemployment rate	-0.24	0.00**	-0.16	0.21
Income per capita	0.00	0.00**	0.00	0.04*
Log-population	3.24	0.00**	2.66	0.00**
% of population \geq 65 years old	-0.28	0.00**	-0.19	0.11
		AIC = 484.63, Deviance = 296.46	AIC = 285.64, Deviance = 27.19	

* p -value < 0.05, ** p -value < 0.01, degrees of freedom (df) = 14

South Kalimantan, which is directly adjacent to Central Kalimantan, could cause several regions to have a stronger influence on the Covid-19 transmission due to the mobility of citizens who are difficult to be monitored. It might be influenced by spatial effects divided into 2 aspects, spatial dependence, and spatial heterogeneity. The spatial dependency yielded Moran's I of 0.082. The p -value of 0.00004 indicates that the spatial dependency was significant at $\alpha = 5\%$. It shows that the Covid-19 cases in 1 location depend on cases in nearby locations.

The Breusch-Pagan value, the spatial effect variation indicator, of 19.46 compared with the 12 degrees of freedom yields a p -value of 0.07. This indicates that the spatial variation among regions is not different at the 5% significance level, though it is significant at a moderate 10% level. It implied the spatial effect did not vary in some regions. Based on these findings, the model that is thought to be appropriate for determining the factors influencing Covid-19 is MGWNBR. This is based on the statistical tests that led to the application of the mixed model in the Covid-19 confirmed cases in both provinces.

The MGWNBR model involves the same 12 determinants used in negative binomial regression (after eliminating multicollinearity variables). That number is also applied to the local GWNBR model. Through modeling conducted by the iterative L-BFGS-B, it was identified that there are 4 local parameters and 7 global parameters having a significant effect on the Covid-19 cases modeled by MGWNBR, while hypertension prevalence is not significant in the entire region. The local parameters are ARI prevalence, number of health care facilities, proportion of correct handwashing behavior, and labor force participation rate. While the global parameters are pneumonia prevalence, diabetes prevalence, unemployment rate, sex ratio, income per capita, logged number of populations, and percentage of 65 years old population and above.

Table 5 shows the distribution of regions in the 2 provinces, named clusters, which divide the regions by the same significant parameters.

Table 5 The clusters of Covid-19 cases by regions upon MGWNBR.

Clusters	Region (Regency / city)	Significant determinants
1	West Kotawaringin, East Kotawaringin, Seruyan, Lamandau, Sukamara	All MGWNBR parameters excluding a proportion of correct handwashing behavior, number of health care facilities, and labor force participation rate
2	Pulang Pisau, Kapuas	All MGWNBR parameters excluding ARI prevalence and number of health care facilities
3	South Barito, East Barito, Gunung Mas	All MGWNBR parameters excluding a proportion of correct handwashing behavior and number of health care facilities
4	North Barito, Murung Raya	All MGWNBR parameters excluding number of health care facilities
5	Palangka Raya City, Katingan	All MGWNBR parameters excluding ARI prevalence, a proportion of correct hand washing behavior, and number of health care facilities
6	Tanah Laut	All MGWNBR parameters
7	Banjarmasin City, Banjarbaru City, Balangan Banjar, Kotabaru, Tanah Bumbu, Tapin, Barito Kuala, North Hulu Sungai, South Hulu Sungai, Central Hulu Sungai	All MGWNBR parameters excluding a proportion of correct handwashing behavior
8	Tabalong	All MGWNBR parameters excluding ARI prevalence and proportion of correct handwashing behavior

Table 5 shows that 27 regencies/cities are divided into 8 clusters. It is indicated that the adjacent areas tend to group into 8 clusters. For example, cluster 5 consists of Palangka Raya City and Katingan Regency which are geographically neighboring. Hence, both are indicated to have a spatial influence on the transmission of Covid-19 cases. As an illustration, **Figure 2** displays the distribution of coefficient of local parameter, ARI prevalence.

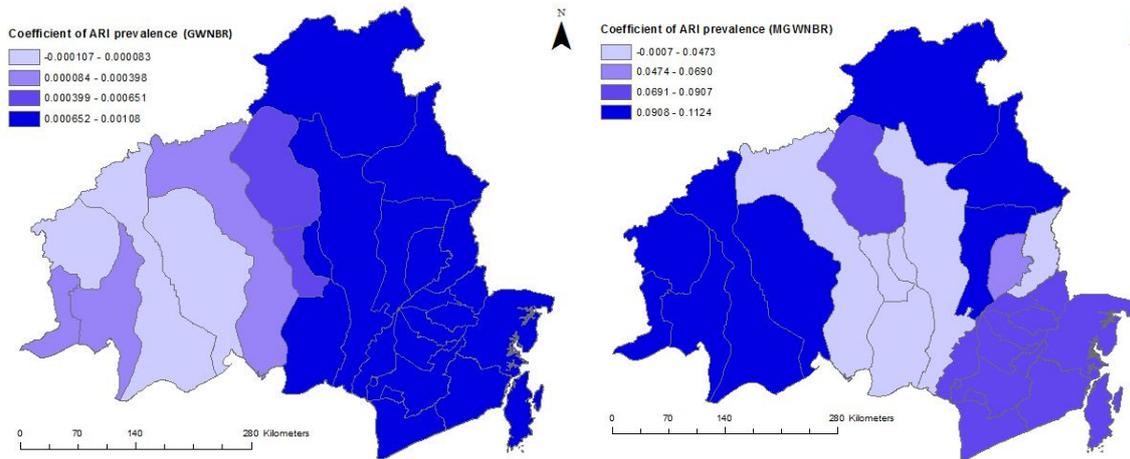


Figure 2 Coefficient distribution of ARI prevalence in describing Covid-19 cases.

Figure 2 indicates the heavy difference of coefficient of ARI prevalence yielded by both models. Such coefficient estimated by GWNBR represents a similar pattern in the whole region of South Kalimantan and northern region of Central Kalimantan, shown by dark blue pattern. The coefficient yielded by MGWNBR somehow exhibits the larger effect of the ARI prevalence on Covid-19 cases. It could be seen from the coefficient range of ARI prevalence of MGWNBR. It is clearly shown that neighboring areas in both provinces have a similar pattern, and the awareness should be focused on the western and northern region of Central Kalimantan since they represent the darker color based on MGWNBR map.

Table 6 provides a summary of the goodness of fit of several models in exploring the Covid-19 case in 27 regencies/cities in Kalimantan as of June 2020.

Table 6 Goodness-of-fit.

Criterion	OLS	Poisson regression	Negative binomial regression	GWNBR	MGWNBR
AIC	362.05	484.63	285.64	333.77	214.07
Deviance	373367	296.46	27.19	963.18	35.82
Pseudo R-sq	79.54	95.35	89.61	81.02	92.02

Table 6 indicates that the MGWNBR model shown satisfying performance in modeling the factors spatially influencing the Covid-19 case in Kalimantan. This is evident from the least AIC. It is also MGWNBR's deviance less than others' including GWNBR's, except negative binomial model's. Based on the R-sq, MGWNBR is better than GWNBR and negative binomial regression in projecting the Covid-19 cases in Central and South Kalimantan.

As an illustration, Eq. (21) shows the estimation of the MGWNBR model for the Covid-19 case in West Kotawaringin regency with a dispersion r is 6.34.

$$\hat{\mu} = \exp(9.409 + 0.1x_1 + 0.769x_2 + 0.838x_3 - 0.03x_5 - 0.068x_6 - 0.211x_9 - 0.013x_{10} - 0.193x_{11} + 0.001x_{12} + 2.659x_{13} - 0.165x_{14}) \quad (21)$$

From Eq. (21), 3 comorbid prevalences, such as ARI, pneumonia, and diabetes have a positive effect on the number of Covid-19 cases in West Kotawaringin. If other variables are assumed to be constant, a 1 % increase in pneumonia prevalence will increase the number of Covid-19 cases by $\exp(0.769) \approx 2$ cases. Besides pneumonia, increasing ARI prevalence, diabetes prevalence, income per capita, or number of populations should impact on increasing Covid-19 cases in West Kotawaringin

On the contrary, the proportion of correct handwashing behavior negatively affects the number of Covid-19 cases in West Kotawaringin. Assuming other variables are constant, a 1 % decrease in this proportion will increase the number of Covid-19 cases by $\exp(-0.03) \approx 1$ case. Aside from handwashing behavior, an increasing number of health care facilities, sex ratio, labor force participation, unemployment rate, or percentage of 65 years old and above will also impact on decreasing Covid-19 cases.

Another thing is that areas with high unemployment rates might have a negative relationship with the number of Covid-19 cases. In West Kotawaringin, this is indicated by a coefficient value of -0.193 for the unemployment rate parameter. This indicates that a 1 % increase in the unemployment rate will decrease the number of Covid-19 cases by $\exp(-0.193) \approx 1$ case, assuming others are constant. It means that workers might have a higher possibility to transmit the virus while working in the site or factory. Unemployed people in this case somehow might be relevant with reduced Covid-19 cases since they were forced to stay at home.

Considering the importance of residual on models' performance, **Figure 3** illustrates the residual distribution in 27 regencies/cities for the GWNBR and MGWNBR models.

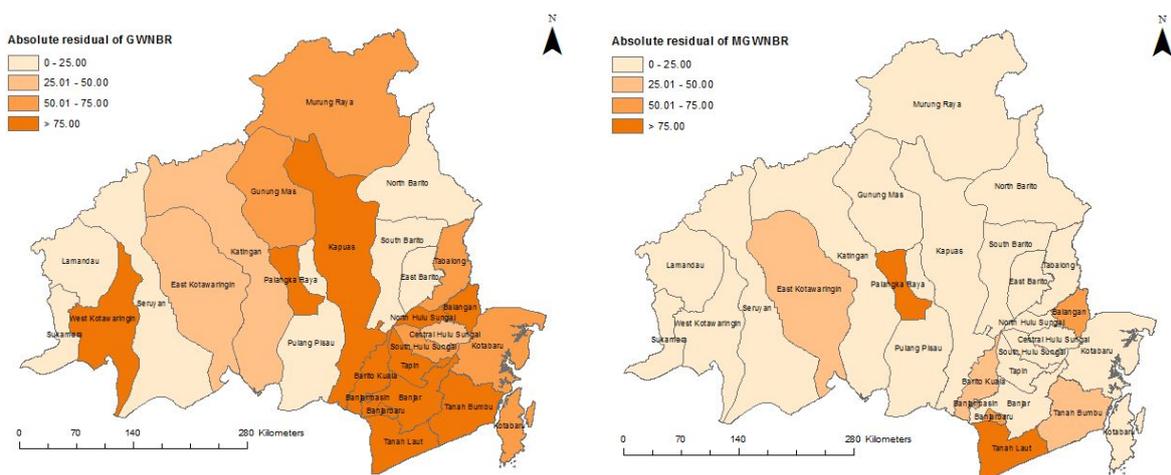


Figure 3 Absolute residual distribution of spatial models.

The light color indicates low residual, while the darker color indicates the opposite, either under-or over-estimation. Both maps in **Figure 3** reveal the strong consistency from most neighboring areas in a similar range of residual. The performance of MGWNBR shows a better projection than that of GWNBR. It seems that the number of the region with the light color in the MGWNBR residual map is much more than that of in GWNBR. It indicates that MGWNBR is satisfying in projecting Covid-19 cases in Central and South Kalimantan rather than GWNBR is doing.

Conclusions

This study implemented several statistical models to explore the effect of sociodemographic and economic determinants on Covid-19 cases in 2 provinces in Kalimantan, involving Central Kalimantan and South Kalimantan. Even though the global models (Poisson and negative binomial regression) showed adequate performance, but both did not adequately address the spatial variation in Covid-19 cases. To explore the effect of the determinants to Covid-19 cases, GWNBR and MGWNBR were proposed to identify the spatial variation between neighboring regions. Based on the findings, we have established as follows.

1) The MGWNBR suggested that 4 determinants have different effects on the Covid-19 cases across regions (locally significant) and 7 parameters have global effects on the Covid-19 cases. The local parameters comprise ARI prevalence, number of health care facilities, proportion of correct handwashing behavior, and labor force participation rate. While the global parameters are pneumonia prevalence, diabetes prevalence, unemployment rate, sex ratio, income per capita, logged number of populations, and percentage of 65 years old population and above. It indicated that some regions could face a challenging task in overcoming the outbreak rather than the others when those 4 local determinants did not meet the expected figure.

2) The proposed MGWNBR identified that the adjacent regions tend to group in 8 clusters containing the same significant determinants.

3) Proposed L-BFGS-B iteration in estimating MGWNBR parameters led to a significant difference of coefficient estimation for some determinants compared to the usual Newton-Raphson as GWNBR did.

4) MGWNBR is better than local GWNBR in terms of AIC, deviance, and R-sq. Thus, the more appropriate spatial model in identifying the factors influencing Covid-19 cases is MGWNBR. The map of residual also suggested that MGWNBR produced a more accurate projection than GWNBR.

5) The comorbid prevalences were positively associated to the Covid-19 increasing cases in most regions. Such comorbids were acute respiratory infection, pneumonia, and diabetes. Besides, the unemployment rate can be negatively related to an increase in Covid-19 cases. Unemployed people in this case somehow might be relevant with reduced Covid-19 cases since they were forced to stay at home. The negative relationship was also found on the number of health care facilities. More health care facilities should reduce Covid-19 cases in some regions.

Those findings draw some possibilities in further studies on extending the univariate to bivariate spatial model that relies on discrete distribution, such as Poisson and negative binomial. The variables that would be chosen as the pair for Covid-19 cases are incidence rate and Covid-19 death cases, among others. The analysis on Covid-19 confirmed cases which involves dynamic model, such as autoregressive integrated moving average (ARIMA), is suggested as well. A more advanced approach that allows for spatial fixed-effect model also might be a outstanding contribution to the Covid-19 research.

Acknowledgement

This study was supported by the Collaborative Research Project of State Islamic Institute of Palangka Raya and Antasari State Islamic University of Banjarmasin. We are indebted to Statistics Indonesia of Central Kalimantan and South Kalimantan, Indonesia's Ministry of Health, and Indonesia's Ministry of Research and Technology for data sources support used in this study and to the anonymous referees for suggestions and constructive comments which improved our study.

References

- [1] D Kang, H Choi, JH Kim and J Choi. Spatial epidemic dynamics of the COVID-19 outbreak in China. *Int. J. Infect. Dis.* 2020; **94**, 96-102.
- [2] LM Andersen, SR Harden, MM Sugg, JD Runkle and TE Lundquist. Analyzing the spatial determinants of local Covid-19 transmission in the United States. *Sci. Total Environ.* 2021; **754**, 142396.
- [3] A Hasan and Y Nasution. A compartmental epidemic model incorporating probable cases to model COVID-19 outbreak in regions with limited testing capacity. *Isa T.* 2021; **S0019-0578**, 00031-8.
- [4] MN Huda, Sifriyani and Fitriani. Real time epidemic modeling using Richards model: Application for the Covid-19 outbreak in East Kalimantan, Indonesia. *J. Phys. Conf. Ser.* 2021; **1751**, 012025.
- [5] Satuan Tugas Covid-19 Kalimantan Tengah, Available at: <https://mmc.kalteng.go.id/berita/read/31654/perkembangan-terbaru-kasus-covid-19-di-kalteng-semuh-165-orang-dan-konfirmasi-baru-142-orang>, accessed December 2020.
- [6] Dinas Kesehatan Provinsi Kalimantan Selatan, Available at: <http://dinkes.kalselprov.go.id/berita/informasi-terbaru-covid-19-di-kalimantan-selatan-22-desember-2020.html>, accessed December 2020.
- [7] WHO Coronavirus Disease 2019 (Covid-19) Indonesia Situation Report - 13, Available at: <https://www.who.int/indonesia/news/novel-coronavirus/situation-reports>, accessed December 2020.
- [8] F Sun, SA Matthews, TC Yang and MH Hu. A spatial analysis of the COVID-19 period prevalence in US counties through June 28, 2020: Where geography matters? *Ann. Epidemiol.* 2020; **52**, 54-9.
- [9] E Marhamah and I Jaya. Modeling positive COVID-19 cases in Bandung City by means geographically weighted regression. *Commun. Math. Biol. Neurosci.* 2020; **2020**, 77.
- [10] S Mansour, A Al-Kindi, A Al-Said, A Al-Said and P Atkinson. Sociodemographic determinants of COVID-19 incidence rates in Oman: Geospatial modelling using multiscale geographically weighted regression (MGWR). *Sustain. Cities Soc.* 2020; **65**, 102627.
- [11] Z Dien, A Saefuddin and A Djuraidah, Modelling the Number of Cases of Dengue Hemorrhagic Fever with Mixed Geographically Negative Binomial Regression in West Java Province. *Int. J. Sci. Eng. Technol.* 2019; **6**, 71-7.
- [12] A Mollalo, B Vahedi and KM Rivera, GIS-based spatial modeling of COVID-19 incidence rate in the continental United States. *Sci. Total Environ.* 2020; **728**, 138884.
- [13] AL Kapetanović and D Poljak. 2020. Modeling the epidemic outbreak and dynamics of COVID-19 in Croatia. In: Proceedings of the 5th International Conference on Smart and Sustainable Technologies (SpliTech), Bol, Croatia. 2020, p. 1-5.
- [14] JH Claver, NI Séraphin, F Simeon, W Guy, AG Nicolas, ER Malgoire and T Bertrand. Predictive modelling of the COVID-19 epidemic in Cameroon with innovative models. *J. Adv. Appl. Math.* 2020; **5**, 117-37.
- [15] MLS Putera, Kausalitas Bivariat antara Nilai Tukar Rupiah dan Indeks Harga Saham Gabungan. *Statistika* 2019; **19**, 103-17.