Spatial analysis of COVID-19 outbreak to assess the effectiveness of social restriction policy in dealing with the pandemic in Jakarta

Didit Okta Pribadi [a,b,*], Khalid Saifullah [b], Andi Syah Putra [b], Muhammad Nurdin [b], La Ode Syamsul Iman [b], Ernan Rustiadi [b]

[a] Research Center for Plant Conservation and Botanic Gardens, Indonesian Institute of Sciences, Bogor 16003, Indonesia
[b] Center for Regional Systems Analysis, Planning and Development (CRESTPENT/P4W), IPB University, Bogor 16144, Indonesia

ARTICLE INFO

ABSTRACT

Coronavirus disease 2019 (COVID-19) has been spread globally and brought health and socioeconomic issues. Jakarta tried to accommodate health and economic interests through the Large-Scale Social Restriction (LSSR) policy that should be assessed. This study aims to (1) visualize the spatial patterns of confirmed Covid-19 cases and the locations of potential risk of transmission, and (2) determine the spatial processes underlying the spatial patterns of Covid-19 cases. The emerging hot spot analysis and space-time scan statistic were employed to analyze the dynamic of infected cases and transmission risk. A Geographical Weighted Regression (GWR) model was developed to define factors that influence the spatial transmission. The result shows that spatial transmission keeps continuing, despite a decline in the aggregate pandemic curve during LSSR implementation. This was likely affected by settlements types and population density distribution, and transportation networks. Spatial analysis supports the aggregate pandemic curve to increase the pandemic surveillance effectiveness.

1. Introduction

Since the coronavirus disease 2019 (COVID-19) outbreak has been started in Wuhan, China in January 2020, it quickly spread to 216 countries by 16th May 2020 leading to not only health problems but also socioeconomic issues (WHO, 2020). The most important aspect of this pandemic is the speed of transmission through people mobility and interaction. Therefore, people mobility and physical interaction have been restricted to control the pandemic, that in turn, this situation has given great impacts on social and economic activities. However, as policies to reduce mobility and interaction are different between countries, thus, these affect the effectiveness of efforts to slow down the global transmission.

Some countries have implemented a national lockdown due to the soaring number of cases such as India, Italy, and Malaysia, while other countries have done a partial lockdown in some epicenter regions to contain the virus such as Vietnam, China, US, etc. According to Ren (2020) and Homburg (2020), the effectiveness of a lockdown policy depends on the time frame where its application at the beginning of contagion is more effective such as in China and Vietnam compared to

Italy which was too late to lockdown. Still, Ren (2020) argued that the lockdown policy is the last option when the increasing number of cases is beyond control. Other countries such as Hongkong, South Korea and Taiwan choose a different way than a lockdown policy by quickly doing mass testing followed by contact tracing, quarantining, and treating the infected individuals (Ren, 2020). Furthermore, their citizens are disciplined in following health protocols. Therefore, following their steps is uneasy as many countries have a lack of capabilities in medical supplies as well as of disciplined society.

In developing countries, the lockdown policy has become problematic as it reduces economic activities as well as the purchasing power of the society, thus, the government should provide social aid amid declining national income due to deceleration of economic growth. Even at the beginning of the pandemic, lockdown policy did not become an option thus many governments were late to respond to this unprecedented challenge. Indonesia faced the same problem where the government was a bit delayed in responding to the pandemic (Djalante et al., 2020). Since the first announcement of 2 confirmed cases on 2nd March 2020, it took almost one month for the Central Government to establish a policy for restricting social mobility and interaction on 31st

March. Still, the regulation itself was formally enacted on 3rd April 2020 through the Ministry of Health Regulation Number 9/2020. By this regulation, local governments including Jakarta's government have a legal instrument to restrict people mobility and interaction adjusting to the socio-economic situation and the dynamics of confirmed Covid-19 cases. Unfortunately, by the time when the regulation was established, the number of infected Covid-19 has jumped into 1528, thus controlling the pandemic has become uneasy, especially for the epicenter like Jakarta.

The policy is not aimed to have a total lockdown, but try to find a compromise between health and economic interests. This policy called Large Scale Social Restriction (LSSR) or *Pembatasan Sosial Berskala Besar* (PSBB) that allows local government at Provincial or Municipality level to close schools, offices, industries, as well as to restrict public transportation and activities. However, some strategic sectors are still permitted to operate, including health, food, energy, communication, finance, logistics, retail for daily needs, and other sectors such as NGO and social organization that involved in handling COVID-19 issues. Still, these strategic sectors should apply protocols to prevent COVID-19 transmission such as physical distancing, wearing a mask, washing hands regularly, etc.

One of the regions that firstly applied LSSR was Jakarta Province among others. This decision was made based on data of number of confirmed Covid-19 cases that always monitored (i.e. this can be seen on the website: https://corona.jakarta.go.id) and the recommendation from the Jakarta's Task-Force for Covid-19 which consists of experts from universities, research centers, and hospitals. The Governor of Jakarta announced that LSSR was begun on 10th April till 23rd April 2020 for the first phase, based on the Governor Regulation Number 33/2020. This should be quickly implemented as Jakarta has become the largest epicenter of infected cases in Indonesia. Besides, Jakarta is the city core of Jabodetabek Metropolitan Area, which has almost 35 million inhabitants (Cox, 2019), thus, it is extremely high risk.

The implementation of this policy got a lot of critiques as many people were still doing their normal activities, even most companies were still operating. The Indonesian Task-Force for COVID-19 Rapid Response stated that there were 543 companies in Jakarta that broke the LSSR policy in the first phase (Kompas, 2020a). Besides, most of the informal sectors still worked as usual and created a crowd (Kompas, 2020b). This sector was difficult to be controlled as one-third of Jakarta's labor are informal workers who are depending on daily income (Rustiadi et al., 2015). This situation was exacerbated by poor distribution of social aid due to inaccurate database; thus, many targeted people were late or even not getting the aid. Due to this situation, coupled with the number of confirmed cases that kept increasing, the LSSR policy was continued to the second phase started from 24th April to 22nd May 2020.

Still, conflict of interests between health and economic purposes was getting stronger, and this is typical in developing countries. As the number of daily infected cases had a bit slowed down at the end of phase 1 to the beginning of phase 2 (21st April to 26th April 2020), suddenly the implementation of LSSR tended to have a relaxation. This situation was criticized by the Indonesian Medical Association and academician from the Indonesian University who stated that it was too rash for having a conclusion that the pandemic curve has started to flatten or even decline (Detik, 2020). This critique was finally proven as the number of daily confirmed cases increased again and experienced a fluctuation afterwards (Provincial Government of Jakarta, 2020a).

Jakarta's case shows the high intensity of economic pressure on the LSSR policy. In dealing with this situation, focusing only on the aggregate pandemic curve as a base for decision making might lead to misleading responses. We need to identify the distribution of confirmed cases and the dynamic of different local pandemic curves across Jakarta, particularly at the smallest unit of governance called village (*kelurahan*). Then we can make clusters of area that need a specific approach depending on the spatiotemporal pattern of infected cases and the risk of

infection in certain periods. Some areas need a very strict restriction, while others might need a bit relaxed. Thus, a compromise policy could be made based on more precise data, instead of going to endless debates between tightening or relaxing the LSSR policy. This situation is more or less similar in developing countries where longer periods and larger areas of social restriction increasingly threaten economic sustainability.

The spatial analysis approach using GIS technology can be employed to increase government precision in doing surveillance (Mollalo et al., 2020; Sarwar et al., 2020). According to Zhou et al. (2020), spatial data and analysis are important to define four aspects consisting of (1) spatial transmission of the pandemic, (2) spatial prevention and control of the pandemic, (3) spatial allocation of resources, and (4) spatial detection of social sentiments. This study will focus on evaluating the LSSR policy based on the spatiotemporal perspective, especially the transmission of the pandemic. Particularly, this study aims to (1) visualize the spatial patterns of confirmed Covid-19 cases and the locations of potential risk of transmission, and (2) determine the spatial processes (driving factors) underlying the spatial patterns of Covid-19 cases.

2. Materials and methods

2.1. Study area

This study was conducted in Jakarta Province as the biggest epicenter of positive cases of COVID-19 in Indonesia. Jakarta is the capital of Indonesia inhabited by more than 11 million people in 2019 (Provincial Government of Jakarta, 2020b) and becomes the city core of Jabodetabek Metropolitan Area as the second-largest metropolitan in the world after Tokyo Metropolitan Area according to its population number (RIHN 2014; Cox 2019). Therefore, the Provincial Government of Jakarta has become alert in facing the pandemic as the consequences will be high when its spread could not be controlled. The Jakarta has terrestrial and island regions (see Fig. 1), but we only focused on the terrestrial area as the number of confirmed cases in the islands was very low. Our study area consists of 261 villages (*kelurahan*) which are divided into 5 sub-districts.

2.2. Data collection

The data on the accumulated number of confirmed cases per village per day were obtained from the website https://corona.jakarta.go.id/id/peta-persebaran. We used data from 25th March to 28th April 2020, divided into two periods. The first period was from 25th March to 12th April 2020, representing the situation before the implementation of the LSSR policy. Although the LSSR policy was established on 10th April 2020, we set 2 days delayed as a required period from taking the test until releasing the result (Herdiyeni et al., 2020). The second period was from 25th March to 28th April 2020, representing the overall situation until the LSSR policy was implemented. This approach followed Desjardins et al. (2020) when comparing the dynamics of the pandemic of COVID-19 in different periods.

Other data used were population number by village obtained from Indonesia's National Statistics Office (Badan Pusat Statistik, BPS) 2019. The data were employed to calculate the risk of transmission in every village. Villages which have more people facing a greater risk of transmission.

Furthermore, we collected shapefile data on building and road maps of Jakarta from the Open Street Map (OSM) and a train station map from the Detail Spatial Planning of Jakarta 2010-2030. These data used as proxies for depicting people distribution, interaction, and movement considered as main factors that determine Covid-19 transmission. Maps of building types coupled with population density (i.e. calculated from data on village population number divided by village's area) were used to describe people distribution. Whereas road and train station maps were used to describe people interaction and movement as both are transportation networks frequently used by Jakarta's urbanites.
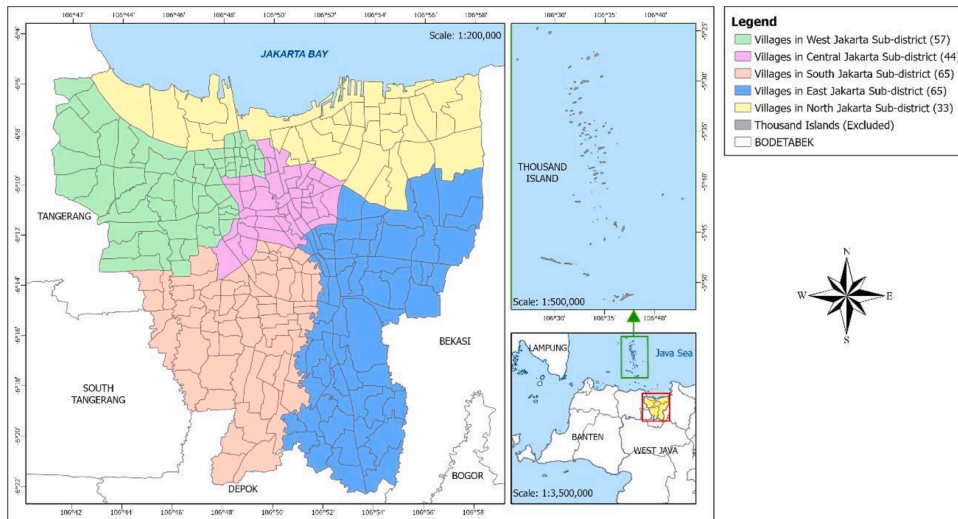
Fig. 1. Map of Jakarta Province as the study area.

2.3. Evaluating the effectiveness of LSSR policy in controlling the pandemic

There were two analysis that would be used to evaluate the situation before and after the implementation of the LSSR policy. First, the emerging hot spot analysis to define the change of hot spot clusters of cases (i.e. every village would have a trend of hot spot scores over time). Second, space-time scan-statistic to define the dynamic of transmission risk clusters across villages. Combining both methods is important as they provide different information. Hot spot clusters of cases might have low intensity but the risks could be high as it occurred in densely population. Contrarily, hot spot clusters of cases might be intense but the risk could be low as it occurred in less densely population. Therefore, those two results would be overlaid to evaluate the effectiveness of the LSSR policy. Furthermore, recommendations could be formulated by defining village clusters that should be prioritized to contain the pandemic transmission. The two approaches can be further explained as follows.

2.3.1. The emerging hot spot analysis

This analysis introduced the time variable into the hot spot analysis. The ordinary hot spot analysis is based on the equation of Getis-Ord Gi* Statistic as follows.

$$G_i^* = \frac{\sum_{j=1}^n w_{i,j} x_j - \overline{X} \sum_{j=1}^n w_{i,j}}{S \sqrt{\frac{\left[n \sum_{j=1}^n w_{i,j}^2 - \left(\sum_{j=1}^n w_{i,j}\right)^2\right]}{n-1}}} \tag{1}$$

Where x_j is the attribute value for village j, $w_{i,j}$ is the spatial weight between villages i and j, n is equal to the total number of villages, while:

$$\overline{X} = \frac{\sum_{j=1}^n x_j}{n} \tag{2}$$

$$S = \sqrt{\frac{\sum_{j=1}^n x_j^2}{n} - \left(\overline{X}\right)^2} \tag{3}$$

The G_i^* statistic produces *hot spot z-score* and *p-value* of a region showing clusters with high and low value spatially.

The time variable was included in the spatial weight, thus, we have 3-dimensional spatiotemporal weight where x (latitude) and y (longitude) axes represent the centroid of the region (village) while z axes represents time (see Fig. 2). These three axes form a unit called bin as a spatiotemporal weight variable, and every bin has a close relationship
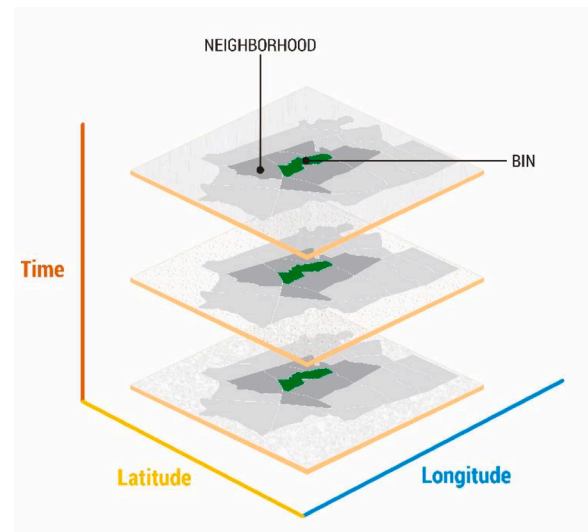

Fig. 2. Space-time cube for calculating Getis-Ord Gi* Statistic.

with its neighboring bin spatially and temporally. By replacing spatial weight ($w_{i,j}$) with bin, the Eq. 1 can be written as follows:

$$G_i^* = \frac{\sum_{j=1}^n b_{i,j,t} x_j - \overline{X} \sum_{j=1}^n b_{i,j,t}}{S \sqrt{\frac{\left[n \sum_{j=1}^n b_{i,j,t}^2 - \left(\sum_{j=1}^n b_{i,j,t}\right)^2\right]}{n-1}}} \tag{4}$$

Where $b_{i,j,t}$ (bin) is the spatial relationship between villages i and j at a certain period of related time.

The neighboring bin based on spatial contiguity (x and y axes) could be calculated by: (1) fixed distance, (2) K nearest neighbors, (3) contiguity edges only, and (4) contiguity edges and corner. In this study, we used a fixed distance as it is a better approach for defining spatial weight within an area which has variable polygon sizes such as villages in Jakarta Province. The fixed distance was defined based on spatial distribution of point data (i.e. centroids of villages) and calculated by using a kernel density search radius. Within this method, the mean center of input point data was determined, followed by calculating the distance between the mean center and all points (village centroids). Then, median of the distances and standard distances were calculated. The radius was then defined by Eq. 5, and we got a radius of 2.6 km.

$$Search radius = 0,9 * min\left(SD, \sqrt{\frac{1}{\ln(2)}} * D_m\right) * n^{-0,2} \tag{5}$$

Where: D_m is median distances from the mean center, n is the number of points, SD is standard distance. Note that min means smaller value would be used whichever SD or $\sqrt{\frac{1}{\ln(2)}}$.

While the neighboring bin based on temporal contiguity (z axes) was defined two days as a period needed for testing until having a result. It means that the data that we have today depend on the situation two days ago.

By using Eq. 4, every bin will have *hot spot z-score* and *p-value*. Statistically significant positive *hot spot z-scores* indicate hot spot, whereas statistically significant negative *hot spot z-scores* indicate cold spot. Then, the trend of hot spot and cold spot over time (based on *hot spot z-score* and *p-value*) was evaluated by using Mann-Kendall trend test. Mann-Kendall statistic do a pair comparison between two periods. If the first time period is smaller than the second, the result is +1, contrarily, if the second is smaller than the first, the result is -1. While, if the both values are similar, the result is 0. Afterwards, all results of the comparisons were summed, then compared with the expected result which is 0 (zero), indicating that there is no trend within a certain period. This trend was statistically tested by using *trend z-scores* and *p-value* which were calculated based on the sum of a trend compared to the expected sum (zero). Based on the test, every village can be classified into 17 categories. One of those categories is **no pattern detected**, while the rest of the 16 categories are classified into hot spot and cold spot groups. For either hot spot or cold spot classification, there are 8 categories as explained in Table 1. This emerging hot spot analysis was done by using ArcGIS Pro 2.5 software.

2.3.2. Space-time scan statistic

This approach has been applied by Desjardins et al. (2020) and Hohl et al. (2020) for COVID-19 case in the US. Scan statistic was aimed to detect and evaluate the cluster of risk in space-time setting thus, we can identify clusters which were active in a certain period. We used SaTscan[TM] software (Kulldorff, 2018) that defines the clusters by a cylindrical window that moving across the region where the base of the cylinder represents space and the height of the cylinder represents time. The center of the cylinder is the centroid of the village.

Like Desjardins et al. (2020) and Hohl et al. (2020), we used a Poisson discrete model which means that the number of COVID-19 infected cases in each village was Poisson-distributed. We used data of population by village which were assumed unchanged during the study period. While the data on the daily number of infected cases were obtained by subtracting the current day count (n_t) with the previous day count (n_{t-1}). We used the maximum spatial window of 2.6 km radius (e.g.

Table 1
Category of hot spots or cold spots classification.

Category	Explanation
New	the most recent time step interval is hot/cold spot cluster of confirmed cases for the first time
Consecutive	a single uninterrupted run of hot/cold spot cluster of confirmed cases time step intervals, comprised of less than 90% of all intervals
Intensifying	at least 90% of the time step intervals are hot/cold cluster of confirmed cases and becoming hotter / colder over time
Persistent	at least 90% of the time step intervals are hot/cold cluster of confirmed cases, with no trend up and down
Diminishing	at least 90% of the time step intervals are hot/cold cluster of confirmed cases, and becoming less hot/cold over time
Sporadic	some of the time step intervals are hot/cold cluster of confirmed cases
Oscillating	some of the time step intervals are hot spot cluster of confirmed cases, some are cold spot cluster of confirmed cases or contrarily
Historical	at least 90% of the time step intervals are hot/cold cluster of confirmed cases, but the most recent time step interval is not

Source: Esri (2020)

similar to the radius used in the emerging hot spot analysis), while the maximum temporal window by default was 50% of the study period. Each cluster should contain at least 5 cases as the default of the model which was also used by Desjardins et al. (2020) and Holh et al. (2020). We used a temporal trend adjustment based on a log-linear trend automatically calculated approach, meaning that the trend of increasing or decreasing infected cases was calculated during the study period and considered in the cluster formation.

The null hypothesis of the Poisson discrete model is the expected number of cases across areas are proportional to population size. While the alternative hypothesis is some areas have several cases exceed the expected number in the null hypothesis. The likelihood ratio test was done to test the hypothesis. Under the Poisson assumption, the likelihood function can be written as follows:

$$\left(\frac{c}{E[c]}\right)^c \left(\frac{C-c}{C-E[c]}\right)^{C-c} I() \tag{6}$$

where C is the total number of cases, c is the number of cases within the scanning window, while E[c] is the expected number of cases within the scanning window under the null hypothesis. As we set scan statistic to find clusters with high rates only, I() value is equal to 1 when the window has a greater number of cases than expected under the null hypothesis.

The most important result of scan statistic is a relative risk showing the risk value within the cluster divided by the risk value outside the cluster. In this study, we used clusters that have p-value ≤ 0.05. The risk value itself was calculated by dividing the observed number by the expected number. In general, this can be written in the equation as follows:

$$RR = \frac{c/E[c]}{(C-c)/(E[C]-E[c])} = \frac{c/E[c]}{(C-c)/(C-E[c])} \tag{7}$$

where c is the number of cases within-cluster, C is the total number of cases, E(C) is the expected total number of cases, thus, in this case $E[C] = C$.

2.4. Defining driving factors that affect the distribution of infected cases between villages

This analysis is important to identify driving factors, including their specific location, that likely affect the spatial transmission of infected cases across villages. We employed Geographically Weighted Regression (GWR) since the preliminary test showed that our model had a heteroskedasticity problem based on Breusch-Pagan Test (p-value < 0.05) and Koenker-Basset test (p-value < 0.1). In this analysis, we used the data for the whole period from 25th March to 28th April 2020.

GWR has been used by Mollalo et al. (2020) to analyze the spatial transmission of COVID-19 in the US. They focused on socioeconomic variables of the population between the counties that likely affect the transmission. However, our study is different from Mollalo et al. (2020) as we more focus on urban structure variables that affecting people distribution, interaction, and movement. Variables used in the GWR model are shown in Table 2.

GWR produced local coefficients which are more reliable than global coefficients when the regression has a heteroskedasticity problem. GWR was firstly introduced by McMillen (1996), McMillen and McDonald (1997), and Brundson et al. (1996), and the model can be written as follows:

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^{n} \beta_k(u_i, v_i)x_{i,k} + \varepsilon_i \tag{8}$$

where y_i is dependent variable at the village i; $\beta_0(u_i, v_i)$ is the intercept at the village i; $x_{i,k}$ is the kth independent variables at the village i; $\beta_k(u_i, v_i)$ is the local regression coefficient for kth explanatory variables at the village i, while (u_i, y_i) is the coordinate of the centroid of the village i.

Table 2

Variables used in the GWR model.

Variables	Explanation	Source
Dependent		
Mean of cases per village	The average value of daily infected cases per village from 25th March to 28th April 2020 (the dependent variable)	Calculated from the data of daily cases per village obtained from the website
Independent		
Number of building with size $< 90\ m^2$ per village	Building with size $< 90\ m^2$ represents low-middle class settlements	Obtained from Open Street Map (OSM) data
Number of building with size $90 - 300\ m^2$ per village	Building with size $90-300\ m^2$ represents high-middle class settlements	Obtained from OSM data
Number of building with size $> 1000\ m^2$ per village	Buildings with size $> 1000\ m^2$ have the potential to make a crowd	Obtained from OSM data
Population density	The number of populations divided by the village's area	Calculated from population data per village obtained from BPS 2019.
Number of road intersection per village	We only considered the intersection of 3 road lines and above	Obtained from OSM data
Distance from Train Stations	We used this variable as the train is the main public transportation frequently used in Jakarta	Calculated from the train station map obtained from the Detail Spatial Planning of Jakarta 2010-2030

Then the local coefficients were calculated based on this following equation:

$$\widehat{\beta}(u_i, v_i) = (X'W(u_i, y_i)X)^{-1} X'W(u_i, y_i)y \qquad (7)$$

where $\widehat{\beta}(u_i, v_i)$ is the estimate of coefficient values at the location i, while $W(u_i, y_i)$ is a n by n spatial weighting matrix for n number of location i. In this study, the spatial weighting matrix was defined by the Kernel function with the type of adaptive Gaussian. This type was chosen as in the case of an emerging disease like COVID-19, the outcome across villages might be unbalanced. The adaptive Gaussian function is:

$$W_{i,j} = \exp\left(-d_{ij}^2 \Big/ \theta_{ik}^2\right) \qquad (8)$$

where W_{ij} is the weight value of observation at village j to estimate coefficients at a village i, d_{ij} is the Euclidian distance between i and j, $\theta_{i(k)}$ is an adaptive bandwidth size based on the kth nearest neighbor distance. The GWR model was calculated by using the GWR4 software (Nakaya et al., 2016).

3. Results

3.1. The dynamic of infected cases based on the emerging hot spot analysis

As mentioned in the introduction, the increasing number of new infected cases in Jakarta had been slowed down at the end of the first phase until the beginning of the second phase of the LSSR policy. This declining trend can be seen in the dash line in Fig. 3, although afterwards the number of cases has tended to increase due to a relaxation in the second phase of the LSSR policy. Although the Governor still wanted to enforce the LSSR policy, the economic pressure was getting higher, particularly after the decline.

This decline was confirmed by the result of the emerging hot spot analysis. Fig. 4(a) shows the emerging hot spot from 25th March to 12th April 2020 (i.e. before the implementation of the LSSR policy). It can be seen that, from higher to lower intensity of hot spot categories, there were 3 villages of the intensifying hot spot, 66 villages of the consecutive hot spot, 23 villages of the new hot spot, and 34 villages of the oscillating hot spot, respectively. Other villages were either cold spot (21 villages) or no pattern detected (114 villages). While Fig. 4(b) shows the emerging hot spot from 25th March to 28 April 2020 (i.e. including the period of the LSSR policy). It can be seen that intensifying hot spot were already gone, while numbers of consecutive and new hot spot were reduced into 13 villages and 42 villages, respectively. However, the number of oscillating hot spot increased to 146, while the number of no pattern detected declined to 59.

These results indicate that the intensity of hot spot had declined after the implementation of the LSSR policy. However, it should be noticed that the spatial transmission was kept going as the number of no pattern detected had been dropped from 114 villages to 59 villages. Furthermore, the number of oscillating hot spot increased and became dominant. Regarding the intensity, the oscillating hot spot has less intensity as it sometimes turns into hot spot or cold spot during the period. However, as the people and government were not alert and tried to relax the LSSR policy, then the number of infected cases increased again and spread wider than before. This situation highlights that decision making in tightening or relaxing the LSSR policy should not only rely on the aggregate pandemic curve.

3.2. The dynamic of COVID-19 transmission risk based on the space-time scan statistic

Besides the dynamic of clusters of infected cases by village, it is also important to identify the transmission risk by village. Fig. 5 coupled with Tables 3 and 4 show the results of space-time scan statistic of the
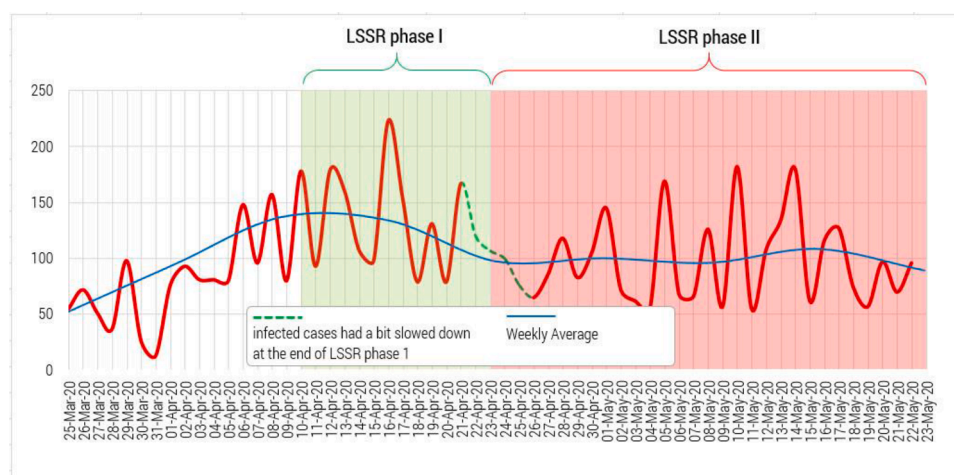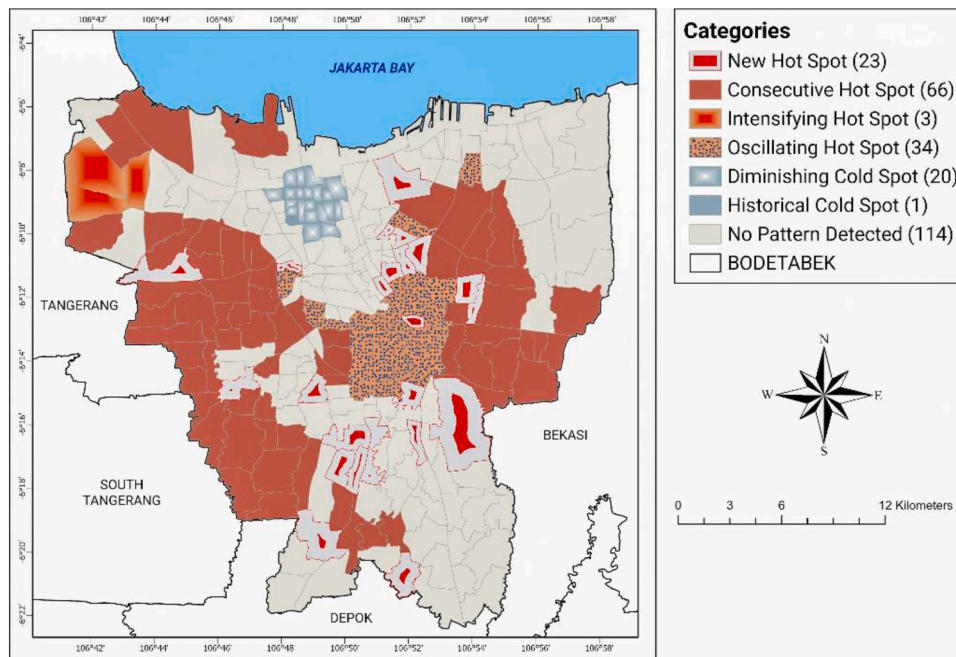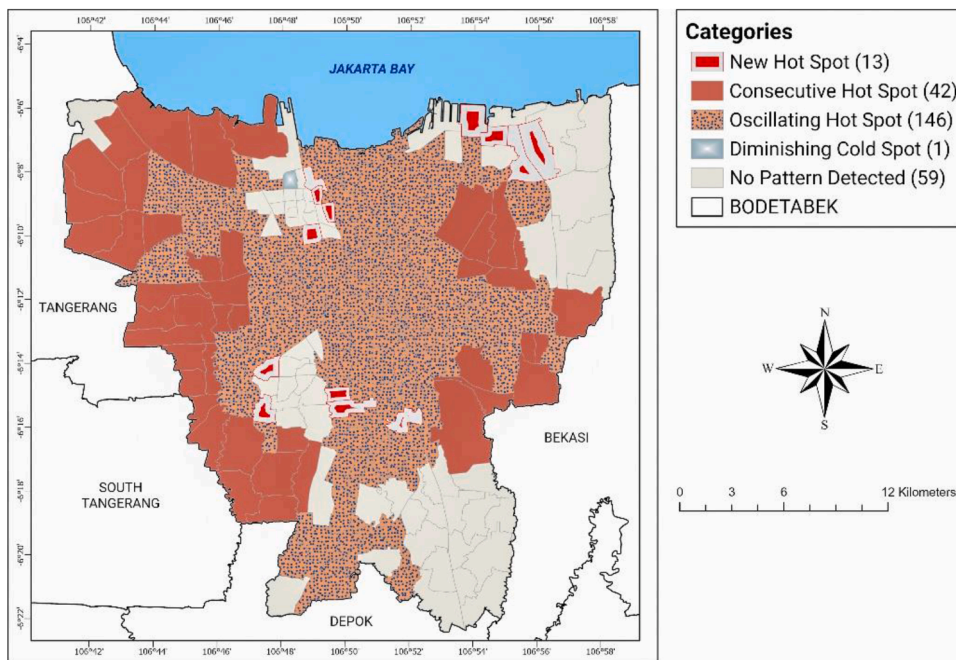

Fig. 3. Daily number and weekly averages of new infected cases in Jakarta Province.
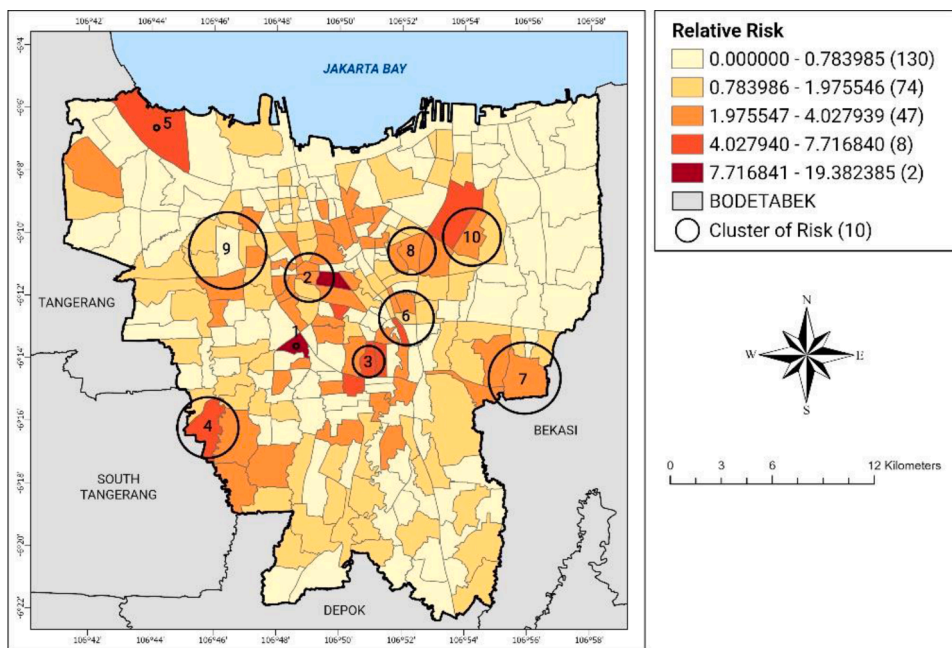
(a)

(b)

Fig. 4. The result of the emerging hot spot (a) from 25th March to 12th April 2020 and (b) from 25th March to 28th April 2020 (note: number of villages is written in the bracket).

two periods. Here, clusters of risk were identified based on the dynamic of infected cases and population by village under the Poisson probability distribution. Table 3 and Fig. 5(a) show that during the first period (before the implementation of the LSSR policy), there were 10 significant clusters (p-value ≤ 0.05) which have the minimum RR value of 4.617, and the maximum RR value of 231.22.

While, Table 4 and Fig. 5(b) show that the RR value was lower than the previous period, where the minimum RR of clusters was 2.007, and the maximum was 199.60. This indicated that the risk had declined after the implementation of the LSSR policy. Again, this confirms that an

increasing number of infected cases had been slowed down. However, it should be noticed that the number of clusters increased from 10 to 14 clusters, which means more areas or villages have been put under the risk of COVID-19 transmission. These results were in line with the results of the emerging hot spot analysis, indicating that spatial transmission still occurred to wider areas/villages, despite the declining number of new infected cases in the aggregate pandemic curve.
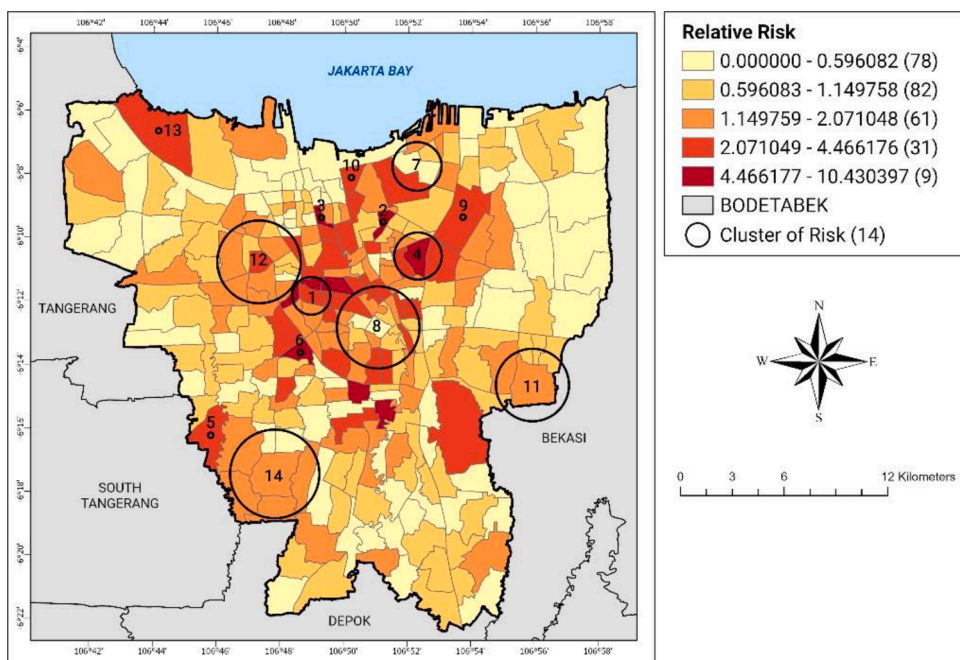
Fig. 5. Spatial distribution of emerging space-time clusters of COVID-19 risk at the village-level (a) from 25th March to 12th April 2020 and (b) from 25th March to 28th April 2020 (note: number of villages is written in the bracket).

3.3. The overlay between the emerging hot spot clusters and transmission risk clusters

Combining the dynamic of hot spot clusters of cases and the dynamic of transmission risk clusters is crucial. It will guide the government to do more precise spatial prevention and control, as well as the spatial allocation of resources. Therefore, infected case surveillance can be done more efficiently and effectively. We can make priorities which areas/villages that should be strongly quarantined and which areas/villages where the population may continue their activities but should apply health protocols.

Regarding hot spot clusters of cases the combination was based on intensity of the hot spot trend from high to low consisting of: persistent, intensifying, consecutive, new, diminishing, sporadic, oscillating and historical, respectively. While according to the transmission risk the combination was based on high risk (i.e. when the villages were included within the risk clusters) and low risk (i.e. when the villages were outside the risk clusters). Thus, the combination of high-risk and hot spot clusters of cases has higher priority than the combination of low-risk and hot spot clusters of cases.

Fig. 6(a) shows 42 villages that should be prioritized. There were 20 villages in the category of high risk – consecutive hot spot, 4 villages in

Tabel 3

Emerging space-time clusters of COVID-19 from March 25th-April 12th, 2020 at the village-level (RR = relative risk).

Cluster	Duration (days)	p	Observed	Expected	RR	Villages	Villages with RR > 1
1	Mar 27th - Mar 28th	0,000	8	0,035	231,228	1	1
2	April 11th - April 12th	0,000	17	1,646	10,507	5	4
3	Mar 29th - April 6th	0,000	22	3,237	6,941	3	3
4	April 4th - April 11th	0,000	28	6,216	4,617	3	3
5	April 6th - April 8th	0,000	7	0,228	30,965	1	1
6	April 5th - April 5th	0,000	14	1,714	8,281	11	7
7	Mar 31st - April 5th	0,000	20	3,831	5,316	2	2
8	April 9th - April 12th	0,001	21	4,363	4,904	5	5
9	Mar 30th - April 31st	0,005	14	2,176	6,520	7	6
10	Mar 27th - Mar 30th	0,012	12	1,739	6,981	3	3

Tabel 4

Emerging space-time clusters of COVID-19 from March 25th-April 12th, 2020 at the village-level (RR = relative risk).

Cluster	Duration (days)	P	Observed	Expected	RR	Villages	Villages with RR > 1
1	April 12th - April 28th	0,000	154	16,809	9,665	3	3
2	April 16th - April 17th	0,000	37	0,614	61,097	1	1
3	April 13th - April 19th	0,000	35	1,049	33,796	1	1
4	April 9th - April 24th	0,000	82	18,746	4,482	4	4
5	April 9th - April 18th	0,000	31	3,663	8,551	1	1
6	March 27th - March 28th	0,000	8	0,040	199,600	1	1
7	April 16th - April 22th	0,000	43	11,501	3,784	3	1
8	April 5th - April 21th	0,000	133	67,963	2,007	21	17
9	April 2nd - April 18th	0,000	25	4,302	5,857	1	1
10	April 14th - April 26th	0,000	39	10,904	3,615	1	1
11	March 31st - April 9th	0,000	30	7,476	4,047	2	2
12	April 24th - April 24th	0,000	24	4,925	4,908	13	10
13	April 6th - April 8th	0,001	7	0,232	30,255	1	1
14	March 27 - April 11th	0,022	50	22,009	2,296	7	5

the category of high risk – new hot spot, and 18 villages in the category of high risk – oscillating hot spot. The next priority was given to the combination of low risk and hot spot. There were three villages with the intensifying hot spot category but had a low risk. Afterwards, villages that become the next priorities were 53 villages in the category of low risk – consecutive hot spot, 19 villages in the category of low risk – new hot spot, and 21 villages in the category of low risk – oscillating hot spot, respectively.

Then Fig. 6(b) shows that the number of high priority villages increased to 73 consisting of 17 villages in the category of high risk - consecutive hot spot and 56 villages in the category of high risk - oscillating hot spot. Still, the intensity was lower than the period before the LSSR policy implementation as the combination of high risk- oscillating hot spot declined from 20 to 17, and there were no villages in the category of high risk- new hot spot. The increase only occurred in the category of high risk - oscillating hot spot from 18 to 56 villages. Again, this situation indicated that the spatial transmission continued, despite the intensity declined due to the LSSR policy. The next priority was given to 31 villages of low risk - consecutive hot spot, 13 villages of low risk - new hot spot, and 104 villages of low risk - oscillating hot spot, respectively.

3.4. Driving factors of spatial transmission of infected cases

As the spatial transmission of infected cases continued in Jakarta, thus it is important to identify driving factors that likely push the transmission. These driving factors might drive social mobility and interaction as key factors of spatial transmission. Here, the GWR model was employed as it produced better results than the global regression. The global regression has R^2 value of 32.31% and AIC value of 1459.6. Whereas the GWR was able to increase the R^2 value to 40.37 % and has a lower AIC value of 1453.89. The R^2 value of the model is relatively low, other variables might be included to improve the model. However, the ANOVA result shows that the GWR model could improve the global regression model, and the improvement was statistically significant (see

Table 5).

The local coefficients of each independent variable which were significant at p-value < 0.1, were mapped. The p-value < 0.1 was used as it resulted better maps than the p-value < 0.05. The maps have more contiguous forms thus could display a better zonation.

The first three explanatory variables were building size classification that depicting settlement types and their potential to make a crowd. These variables denote people distribution. Fig. 7(a) shows the coefficients of building size < 90 m² that represent low-middle class settlements. The results show that the low-middle class settlements close to the border of Jakarta decreased infected cases (e.g. having negative signs), whereas the low-middle class settlements in the centre of Jakarta increased infected cases (e.g. having positive signs).

While Fig. 7(b) shows the coefficients of building size 90-300 m² representing high-middle class settlements. The results show that high-middle class settlements could push increasing infected cases almost in the entire of Jakarta region, especially in the Southern and Eastern part of Jakarta. These areas are dominated by large urban housing development, which are contiguous from large urban housing in Southern Tangerang, Bogor, and Depok in the Southern fringes of Jakarta to Bekasi in the Eastern fringes of Jakarta (Winarso and Firman, 2002). Seemingly, high-middle class society has higher mobility and interaction; thus, it could increase the spatial transmission of infected cases. This confirms findings by the Eijkman Institute that infected cases in Indonesia were dominated by people in the group of high-middle income (dailynewsindonesia, 2020). However, recently some cases were found in the low-middle income class as the people still work outside when the high-middle income class of the population do work from home.

Another building size variable is the larger buildings with size > 1000 m² that have the potential to make a crowd. Fig. 8 shows that large buildings in Northern Jakarta have decreased infected cases (e.g. having negative signs). Contrarily, the large buildings in the Southern Jakarta stimulated increasing infected cases (e.g. having positive signs). Seemingly, the large buildings in the North which are dominated by industries, offices, and malls have applied health protocols as well as work
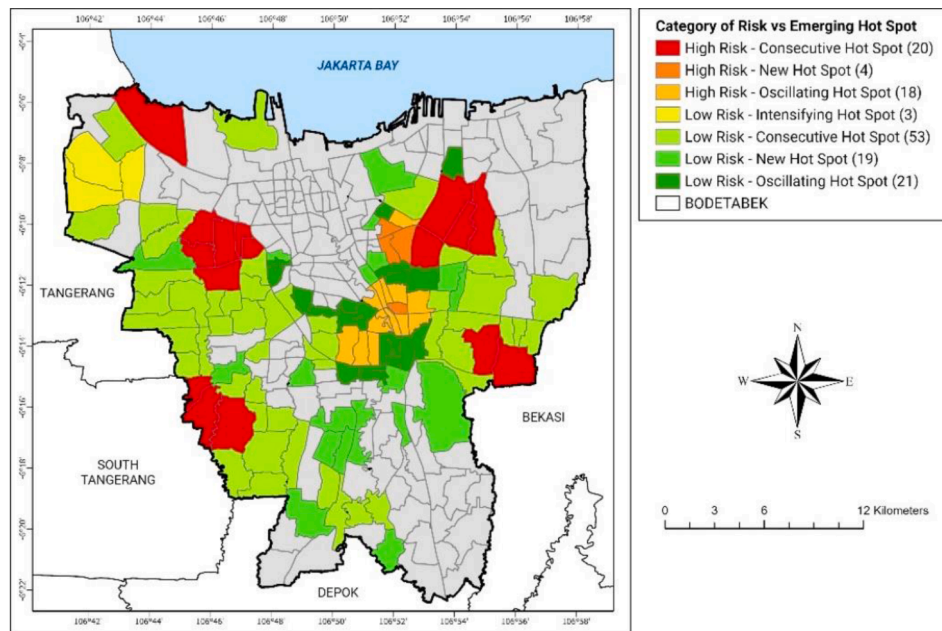
(a)
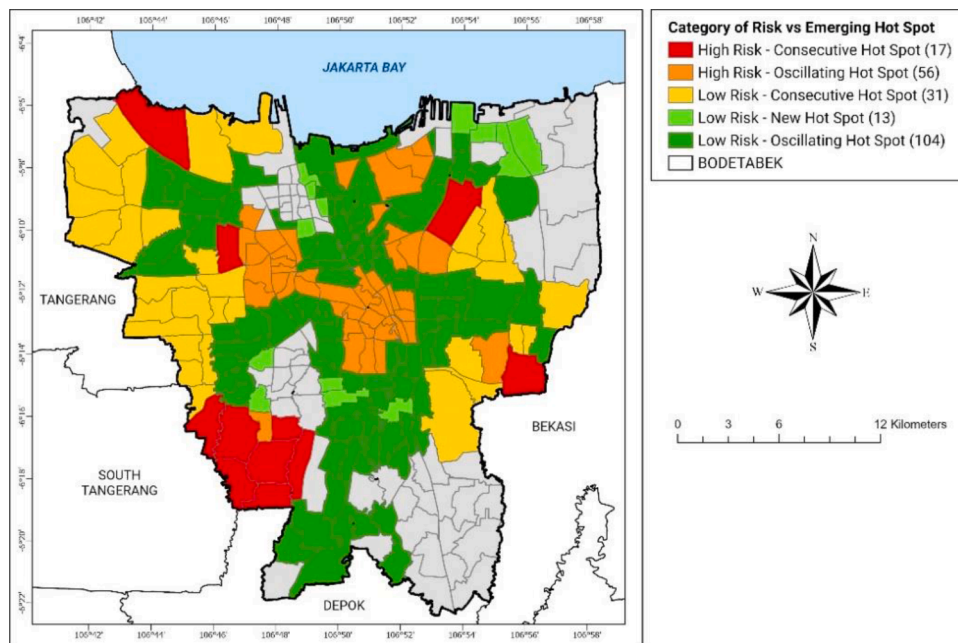
(b)

Fig. 6. Priority areas of COVID-19 surveillance based on the dynamic of hot spot clusters of cases and the dynamic of transmission risk (a) from 25th March to 12th April 2020 and (b) from 25th March to 28th April 2020 (note: number of villages is written in the bracket).

Table 5
The ANOVA result of the GWR model.

Source	Sum of Square	Degree of Freedom	Mean Square	F-value	P-value
Global Residuals	3857.195	254.000			
GWR Improvement	459.060	18.413	24.931		
GWR Residuals	3398.136	235.587	14.424	1.728407	0.034325

from home, thus, they could prevent the transmission of infected cases. Contrarily, large buildings in the South, which are surrounded by settlement areas, increased the transmission of infected cases.

The next variable, which also represents people distribution, is population density. Fig. 9 shows that the population density has a significant effect only in the Southwestern of Jakarta where land-use type
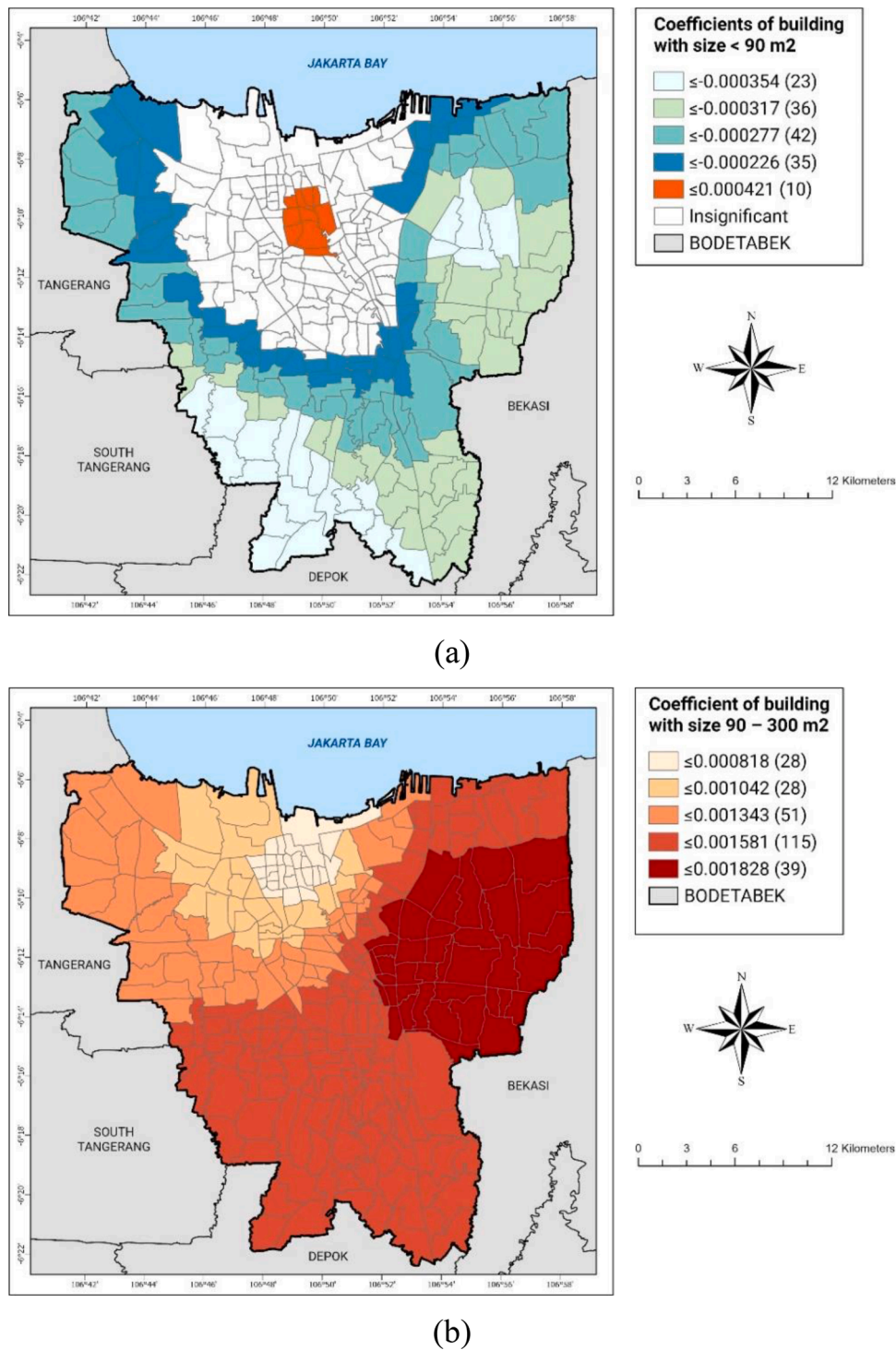
(a)

(b)

Fig. 7. Map of significant coefficients of (a) building size < 90 m^2 and (b) building size $90 - 300$ m^2 based on t-test at $\alpha = 0.1$ (note: number of villages is written in the bracket).

in this area was dominated by settlements. The increasing population density in this location may increase the number of infected cases.

The last two variables represent transportation factors that indicating people interaction and movement. Transportation variables which were considered are the number of road intersection $\geq$ three lines and the distance to the train station. In the modelling process, we have tested other transportation factors such as the distance to arterial roads, collector roads, and toll roads, but none could improve the reliability of the model. It was likely caused by the LSSR policy that strongly

restricted vehicles on the main roads, but less controlling mobility via small roads as well as commuting trips by train.

Fig. 10(a) shows that the road intersection could push infected cases almost in the entire region of Jakarta. This situation indicated that mobility and interaction in residential areas have more influence in increasing infected cases. Probably, this occurred as most of the distant mobility have been restricted during the implementation of the LSSR policy.

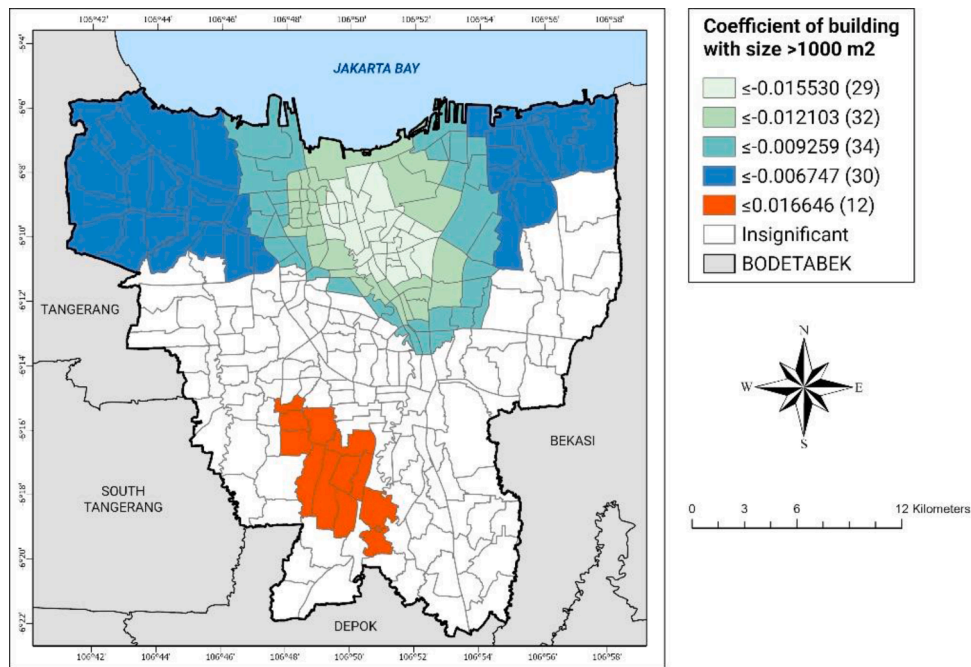Then Fig. 10(b) exhibits that villages close to the train stations were

Fig. 8. Map of significant coefficients of building size >1000 m^2 based on t-test at α=0.1 (note: number of villages is written in the bracket).
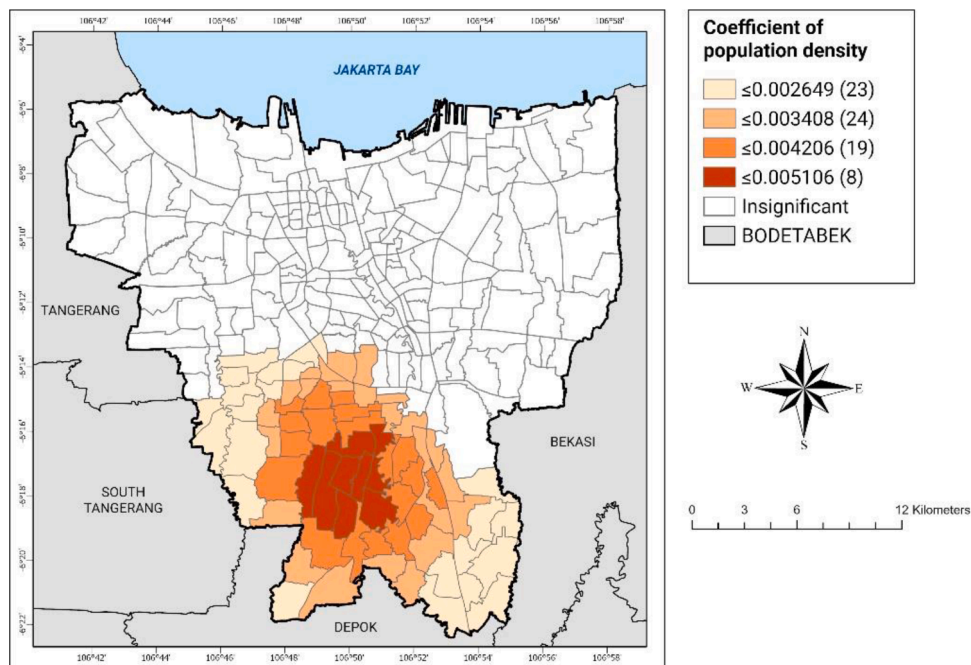

Fig. 9. Map of significant coefficients of population density based on t-test at α=0.1 (note: number of villages is written in the bracket).

likely having a higher number of infected cases. The significant co-efficients were found in the South to East of Jakarta which means commuting trips from Bogor, Depok, and Bekasi (i.e. hinterland areas of Jakarta) have become a key factor that increased spatial transmission. Closing the commuter line has become an issue since the implementation of the LSSR policy, but unfortunately, it was never done.

4. Discussion and conclusion

The results of the emerging hot spot analysis highlight two important things. First, every smaller unit of areas (village) has its own pandemic curve, which is different across the city region. The curves in some areas

have been slowing down, but others might just have started to increase. Therefore, a regular assessment should be done spatially and temporally to develop more effective and efficient surveillance.

Second, the spatial transmission continues, although the intensity of cases has become lower due to the implementation of the LSSR policy. This implies that we should concern about the spreading of infected cases, even though the daily new cases in the aggregate pandemic curve showed a decline. Neglecting this situation will lead to a misleading policy. For instance, when the LSSR policy was a bit relaxed considering the decline, it turns out, leading to higher number of new cases as the disease has been spread to wider areas. It occurred in Jakarta when the government and the people became less aware after the decline, thus the

11
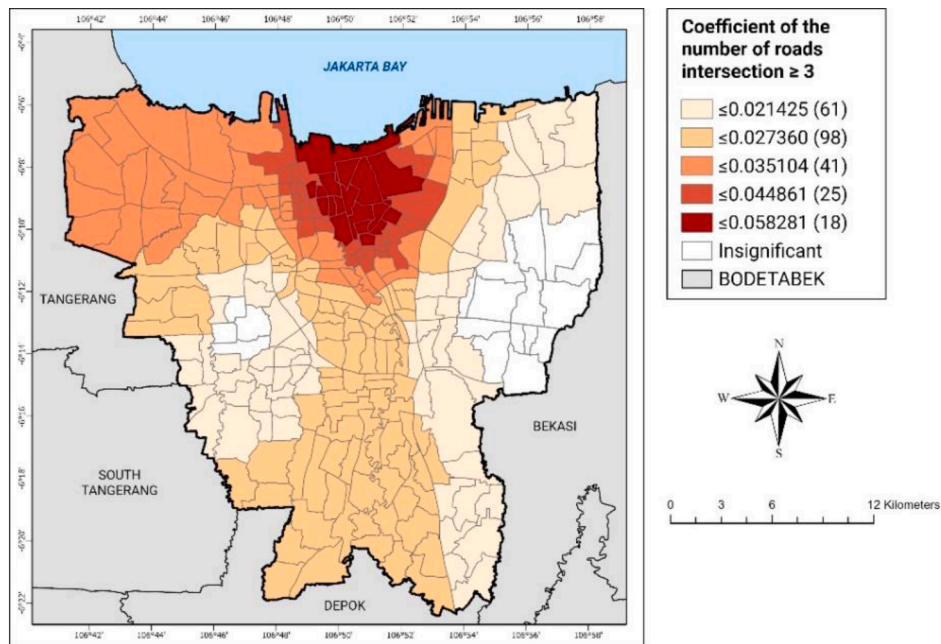
(a)
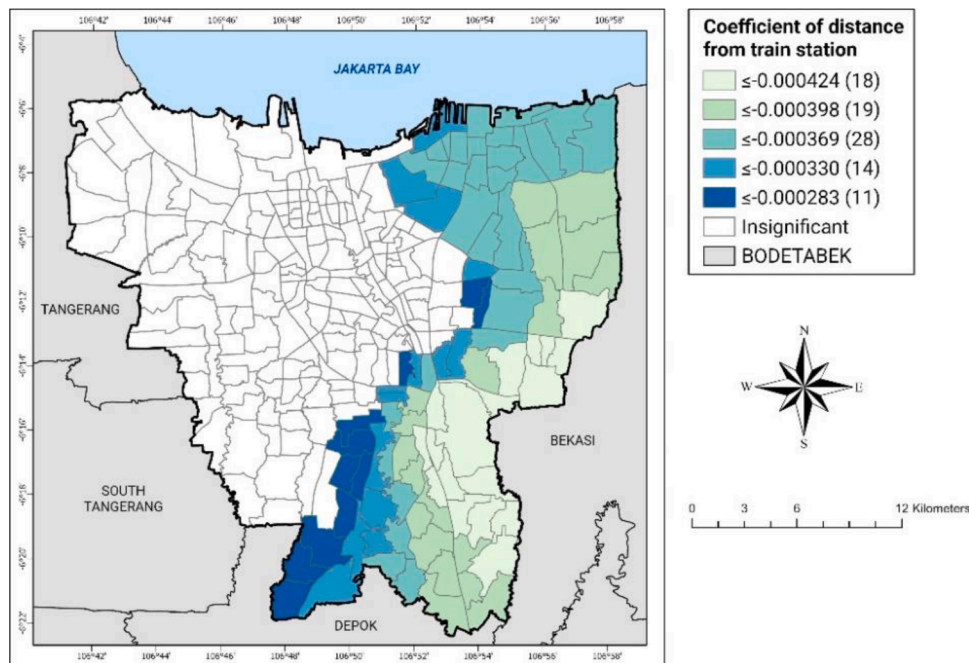
(b)

Fig. 10. Map of significant coefficients of (a) the number of road intersection ≥ 3 lines and (b) distance from the train station based on t-test at $\alpha = 0.1$ (note: number of villages is written in the bracket).

number of new cases has increased again (see Fig. 3).

Still, a regular assessment on spatiotemporal dynamic of confirmed cases is insufficient since the spatiotemporal dynamic of transmission risk should be considered as well. The results of the study highlight that the implementation of the LSSR policy was able to reduce the intensity of transmission risk. However, the number of risk clusters increased, thus, this confirms the results of the emerging hot spot analysis that the spatial transmission still continued to wider area.

Both the dynamic of confirmed cases clusters and transmission risk clusters should be mutually considered as they were complemented each other. Our results show that albeit some villages had become intensifying hot-spot (see Fig. 6a), they had a lower transmission risk due to the lesser number of populations at risk. Contrarily, some villages with lower intensity hot spot (compared to intensifying hot-spot), had a higher risk since they have more population at risk. The combination of these two approaches would help the government in preventing and controlling COVID-19 transmission as well as allocating resources such as medical supplies, medical staffs, etc.

To overcome this continuing spatial transmission, it is also important to identify driving factors that affect the transmission. We used a GWR model as the most appropriate approach for our data in Jakarta. For other areas or context, different models or approaches might be employed. However, as the problem is related to the spatial transmission, then it is important to use spatial modelling approaches that considering spatial correlation and spatial heterogeneity as done by Mollalo et al. (2020) or Guliyev (2020).

The GWR results show that there are three groups of variables that affect the number of infected cases consisting of building types distribution, population density, and transportation. These variables influence social mobility and interaction in Jakarta. Regarding the building variables, the GWR results show that increasing infected cases were likely caused by low-middle class settlement in the city centre; high-middle class settlement in the Southern and Eastern of Jakarta; and larger building that potentially make a crowd in the settlement areas in the Southern of Jakarta. Then, related to the population density, the densest population in the Southwestern of Jakarta might push the number of infected cases. While, the transportation variables that facilitate the mobility around settlements might increase infected cases. Commuter lines by trains also stimulate infected cases, especially for the commuters from peri-urban areas (e.g. Bogor, Depok, and Bekasi) to Jakarta. Normally the trains carry thousands of passengers every day. Although the number of passengers was limited during the implementation of the LSSR policy, many commuters still used this line as Ministry of Industry allowed 721 companies in Jakarta to normally operate (Kompas, 2020c). This is one factor that made the implementation of the LSSR policy in Jakarta was less successful due to unsynchronized policies between the Provincial Government and the Central Government (i.e. Ministry of Industry and Ministry of Transportation).

The results of this study can be used as a guide for the government to apply such kind of approaches for monitoring the dynamic of infected cases and transmission risk. This will increase the government capacity in doing surveillance and suppress the spatial transmission. Furthermore, the method also useful as an instrument to evaluate the effectivity of policies to reduce the transmission in certain areas and periods. More importantly, this study offers an instrument that can be used in managing health and economic issues which is usually conflicting and hard to be handled, especially in developing countries. Certainly, health issue should be put as a priority, but with more precise data, economic activities in areas with lower cases and risk still can be run by implementing health protocols. This also offers a great help when a new normal want to be applied as more massive and accurate data or information become a prerequisite to contain the disease and to save economic activities (Petersen et al., 2020).

Still, this study has some limitations. First, in the emerging hot spot we used fixed neighborhood distance and neighborhood time step which actually could be dynamic based on variability of the case number distribution over time. Second, in the scan statistic we only consider village population in calculating risk, thus it can be improved by including other factors such as number of older people, number of people with comorbid conditions, number of hospitals, etc. Third, in the GWR model, as we mainly used physical variables (buildings, transportation networks), other variables might be added to improve the reliability of the model such as social and economic condition of villages.

This research still can be further explored by incorporating three other spatial aspects as mentioned by Zhou et al, (2020), including spatial prevention and control, spatial allocation of resources, and spatial of social sentiments. Spatial prevention and control can be designed based on the spatial pattern of confirmed cases and transmission risk. Then, the spatial allocation of resources can be defined to decline the potential risk in certain areas and periods. In this case, an area with better health infrastructures and lesser number of infected cases will have a lower risk. Finally, spatial of social sentiments can be analyzed to define the most effective policies to control the spatial transmission by adjusting to the social culture and behavior of a society that could be different between regions.

To do such kind of studies, big data infrastructures should be developed. Especially, the availability of people mobility data would be very helpful to quickly prevent and control further transmission. As Liu et al. (2020) highlighted that the use of big data analysis in London has been able to overcome the pandemic, although the government was a bit late to respond to the emerging pandemic in the early period.

References

Brundson, C.F., Fotheringham, A.S., Charlton, M., 1996. Geographically weighted regression: a method for exploring spatial non-stationarity. Geogr. Anal. 28, 167–174.

Cox, W. (2019). Demographia world urban areas 2019: population, land area and urban densities. https://www.newgeography.com/content/006272-demographia-world-urban-areas-2019-population-land-area-urban-densities. Accessed on 18th May 2020.

Dailynewsindonesia. (2020). Eijkman: positif corona dari kelas menengah atas (Eijkman: positive corona mainly come from high-middle income class). https://www.dailynewsindonesia.com/news/megapolitan/eijkman-positif-corona-dari-kelas-menengah-atas. Accessed on 18th May 2020.

Desjardins, M.R., Hohl, A., Delmelle, E.M., 2020. Rapid surveillance of COVID-19 in the United States using a prospective space-time scan statistic: detecting and evaluating emerging clusters. Appl. Geogr. 118, 102–202. https://doi.org/10.1016/j.apgeog.2020.102202 https://doi.org/https://doi.org/.

Detik. (2020). Pesan jangan berlebihan pada kurva corona DKI yang mulai mendingan (Do not overestimate the pandemic curve of Jakarta that showing a declining trend). https://news.detik.com/berita/d-4995199/pesan-jangan-berlebihan-pada-kurva-corona-dki-yang-mulai-mendingan/. Accessed on 17th May 2020.

Djalante, R., Lassa, J., Setiamarga, D., Sudjatma, A., Indrawan, M., Haryanto, B., Mahfud, C., Sinapoy, M.S., Djalante, S., Rafliana, I., Gunawan, L.A., Surtiari, G.A.K., Warsilah, H., 2020. Review and analysis of current responses to COVID-19 in Indonesia: Period of January to March 2020. Progr. Disaster Sci. 6, 100091 https://doi.org/10.1016/j.pdisas.2020.100091.

ESRI. (2020). How emerging hot spot analysis work. https://pro.arcgis.com/en/pro-app/tool-reference/space-time-pattern-mining/learnmoreemerging.htm. Accessed on 10th April 2020.

Guliyev, H., 2020. Determining the spatial effects of COVID-19 using the spatial panel data model. Spatial Stat. 38, 100443 https://doi.org/10.1016/j.spasta.2020.100443.

Hohl, A., Delmelle, E.M., Desjardins, R.M., Lan, Y., 2020. Daily surveillance of COVID-19 using the prospectivespace-timescan statistic in the United States. Spatial Spatio-Temporal Epidemiol. 34, 100354 https://doi.org/10.1016/j.sste.2020.100354.

Herdiyeni, Y., Poetri, O.N., Indrawan, D.R., Musthafa, Basri, C., Jaya, I., Mulyati, H., 2020. Model prediksi COVID-19 di Indonesia untuk perencanaan mitigasi resiko terhadap dampak epidemiologi, sosial dan ekonomi (Prediction model of COVID-19 in Indonesia for planning risk mitigation on epidemiology, social and economy impacts). IPB University, Bogor.

Homburg, S., 2020. Effectiveness of corona lockdowns: evidence for a number of countries. Discussion Paper Number 671. The Economist's Voice (forthcoming).

Kompas. (2020a). 543 Perusahaan di Jakarta Langgar PSBB, 76 Sudah Disegel (543 companies in Jakarta broke the LSSR policy, 76 have been sealed). https://nasional.kompas.com/read/2020/04/27/17224661/543-perusahaan-di-jakarta-langgar-psbb-76-sudah-disegel. Accessed on 17th May 2020.

Kompas. (2020b). Evaluasi PSBB Jakarta, masih ada kerumunan massa hingga perusahaan curi-curi operasi (Evaluation of LSSR's Jakarta, still finding crowds and companies which still operate). https://megapolitan.kompas.com/read/2020/04/22/19511081/evaluasi-psbb-jakarta-masih-ada-kerumunan-massa-hingga-perusahaan-curi. Accessed on 17th May 2020.

Kompas. (2020c). Kemenperin izinkan 721 perusahaan di Jakarta tetap beroperasi (Ministry of Industry allowed 721 companies in Jakarta keep operating). https://megapolitan.kompas.com/read/2020/04/22/13373471/kemenperin-izinkan-721-perusahaan-di-jakarta-tetap-beroperasi-selama-psbb. Accessed on 18th May 2020.

Kulldorff, M. (2018). SaTScanTM user guide for version 9.6. https://www.satscan.org/. Accessed on 10th April 2020.

Liu, M., Ning, J., Du, Y., Cao, J., Zhang, D., Wang, J., Chen, M., 2020. Modelling the evolution trajectory of COVID-19 in Wuhan, China: experience and suggestions. Public Health 183, 76–80. https://doi.org/10.1016/j.puhe.2020.05.001.

McMillen, D.P., 1996. One hundred fifty years of land values in Chicago: a nonparametric approach. J. Urban Econ. 40, 100–124.

McMillen, D.P., McDonald, J.F., 1997. A nonparametric analysis of employment density in a polycentric city. J. Reg. Sci. 37, 591–612.

Mollalo, A., Vahedi, B., Rivera, K.M., 2020. GIS-based spatial modeling of COVID-19 incidence rate in the continental United States. Sci. Total Environ. 728, 138884 https://doi.org/10.1016/j.scitotenv.2020.138884.

Nakaya, T., Charlton, M., Brunsdon, C., Lewis, P., Yao, J., Fotheringham, S. (2016). GWR 4 windows application for geographically weighted regression modelling., accessed 1.3.2020, https://sgsup.asu.edu/geodacenter-redirect.

Petersen, E., Wasserman, S., Lee, S.-S., Go, U., Holmes, A.H., Abri, S.AlA., McLellan, S., Blumberg, L., Tambyah, P., 2020. COVID-19–We urgently need to start developing an exit strategy. Int. J. Infect. Dis. 96, 233–239. https://doi.org/10.1016/j.ijid.2020.04.035.

Provincial Government of Jakarta. (2020a). Data pemantauan COVID-19 DKI Jakarta (Data monitoring on COVID-19 cases at Jakarta). https://corona.jakarta.go.id/id/data-pemantauan. Accessed on 17th May 2020.

Provincial Government of Jakarta. (2020b). Berapa kepadatan penduduk DKI Jakarta saat ini? (What is the current population density of Jakarta?) http://statistik.jakarta.go.id/berapa-kepadatan-penduduk-dki-jakarta-saat-ini/. Accessed on 17th May 2020.

Ren, X., 2020. Pandemic and lockdown: a territorial approach to COVID-19 in China, Italy and the United States. Eurasian Geogr. Econ. 1–12. https://doi.org/10.1080/15387216.2020.1762103.

RIHN (Research Institute for Humanity and Nature), 2014. Kyojuu kankyou ruikei kara Megacity no guroobaruna renkan to doutai o toraeru. Zen-kyuu toshi zenshi kenkyuu jouhou-sho sougou chikyuu kankyou gaku kenkyuusho. In: Fukami, N., Yamada, K., Uchiyama, Y. (Eds.), Capturing the dynamics and changes of global megacity from living environment point of view. The Report on Global City History. The Megacity Project. Research Institute for Humanity, RIHN, Japan.

Rustiadi, E., Pribadi, D.O., Pravitasari, A.E., Indraprahasta, G.S., Iman, L.S., 2015. Jabodetabek megacity: from city development toward urban complex management system. Urban development challenges, risks and resilience in Asian mega cities. Springer, pp. 421–445.

Sarwar, S., Waheed, R., Sarwar, S., Khan, A., 2020. COVID-19 challenges to Pakistan: is GIS analysis useful to draw solutions? Sci. Total Environ. 730, 139089 https://doi.org/10.1016/j.scitotenv.2020.139089.

WHO. (2020). Coronavirus disease (COVID-19) outbreak situation. https://www.who.int/emergencies/diseases/novel-coronavirus-2019?gclid=Cj0KCQjwnv71BRCOARIsAIkxW9EvYk32e05zL4zJsJSULcIAWPKzAFSgWd4P-6leWqXFCWLYtKUzvvoaAjYeEALw_wcB. Accessed on 16th May 2020.

Winarso, H., Firman, T., 2002. Residential land development in Jabotabek , Indonesia : triggering economic crisis ? Habitat Int. 26, 487–506.

Zhou, C., Su, F., Pei, T., Zhang, A., Du, Y., Luo, B., Cao, Z., Wang, J., Yuan, W., Zhu, Y., Song, C., Chen, J., Xu, J., Li, F., Ma, T., Jiang, L., Yan, F., Yi, J., Hu, Y., Xiao, H., 2020. COVID-19: challenges to GIS with Big Data. *Geogr. Sustain.* https://doi.org/10.1016/j.geosus.2020.03.005.