

## EXPLORATORY SPATIAL DATA ANALYSIS (ESDA) ON COVID-19 CASES IN MALAYSIA

Syerrina Zakaria\*, Nur Edayu Zaini, Siti Madhiah Abdul Malik, Wan Saliha Wan Alwi

Faculty of Ocean Engineering Technology and Informatics,  
University Malaysia Terengganu, 21030 Kuala Terengganu,  
Terengganu, Malaysia

### Article history

Received

30 May 2021

Received in revised form

1 September 2021

Accepted

15 September 2021

Published Online

20 October 2021

\*Corresponding author  
syerrina@umt.edu.my

### Graphical abstract



### Abstract

The Malaysian government implemented The Movement Control Order (MCO) on 18 March 2020 to control the spread of the COVID-19 outbreak. However, the third wave that started in September 2020 during the Recovery Movement Control Order (RMCO) phase saw a continuous increase in the number of cases. In this study, the exploratory spatial data analysis (ESDA) was used to analyse the existence of COVID-19 spatial clusters. Moran's index was used to map the spatial autocorrelation (cluster) to showcase the spreading patterns of the COVID-19 pandemic in Malaysia. The study results indicated significant changes in the COVID-19 hotspots over time. At the beginning of 2020, the state of Selangor and Sarawak were the first locality to become a significant COVID-19 hotspot. Furthermore, this research showed all affected areas during the study period. Overall, a non-random distribution of COVID-19 occurrences was detected, thus suggesting a positive spatial autocorrelation. Many parties are affected by the COVID-19 pandemic, especially those involved in healthcare provision, financial assistance allocation, and law enforcement. Other sectors such as the economy, education, and religion are also affected. Therefore, the findings from this study will provide useful information to all the related governmental and private agencies, as well as policymakers and researchers.

**Keywords:** MCO, RMCO, exploratory spatial data analysis, Moran's index, spatial autocorrelation, COVID-19

### Abstrak

Kerajaan Malaysia telah melaksanakan Perintah Kawalan Pergerakan (PKP) pada 18 Mac 2020 bagi membendung penularan wabak Covid19. Walaubagaimanapun, gelombang ketiga yang bermula pada September 2020 semasa fasa Perintah Kawalan Pergerakan Pemulihan telah menunjukkan peningkatan bilangan kes secara drastik. Didalam kajian ini, penerokaan analisis data ruang (ESDA) digunakan bagi mengenalpasti kewujudan kluster COVID-19. Index Moran digunakan dalam bentuk pemetaan autokolerasi kelompok (kluster) untuk menunjukkan corak penyebaran wabak COVID-19 di Malaysia. Keputusan yang diperoleh menunjukkan terdapat perubahan signifikan pada kesan pengelompokan atau hotspot bagi kes COVID-19 dari masa ke masa. Pada awal tahun 2020, Selangor dan Sarawak adalah lokasi pertama yang menjadi kawasan panas COVID-19 yang signifikan. Selanjutnya, penyelidikan ini menunjukkan semua kawasan yang terjejas dalam tempoh kajian. Secara keseluruhan, kejadian penularan COVID-19 bukan rawak dikesan, sehingga menunjukkan autokorelasi ruang yang positif. Banyak pihak yang

terkesan dengan wabak COVID-19, terutama yang terlibat dalam penyediaan penjagaan kesihatan, peruntukan bantuan kewangan, dan penegakan undang-undang. Sektor lain seperti ekonomi, pendidikan, dan agama juga terjejas. Oleh itu, penemuan dari kajian ini akan memberikan maklumat yang berguna kepada semua agensi kerajaan dan swasta yang berkaitan, serta pembuat dasar dan penyelidik.

*Kata kunci:* PKP, PKPP, Index Moran, penerokaan analisis data ruwang (ESDA), Autokolerasi ruwang, COVID-19

© 2021 Penerbit UTM Press. All rights reserved

## 1.0 INTRODUCTION

In early 2020, countries around the world, including Malaysia was devastated by the COVID-19 pandemic. The outbreak was reported as an unknown disease in Wuhan, the capital of Hubei Province, China [1][2]. Later, the causative agent of the disease was established as a novel strain of the virus, known as SARS-CoV-2. It was later labelled as COVID-19 [3]. People with low immune systems, especially the elderly, infants, and those with comorbidities are prone to the infection. COVID-19 spreads through respiratory droplets that come into direct or indirect contact with the eyes, mouth, and nose [4]. A person infected by COVID-19 can experience high body temperature (fever), sore throat, cough, headache, diarrhoea, flu, and fatigue [5], [6], [7].

To date, a total of 73,345,971 confirmed cases, 1,631,169 deaths, and 51,483,717 recovered cases of COVID-19 have been reported worldwide since 15 December 2019. There is no vaccine or established cure so far. Thus, the number of active cases remains high at 20,231,085, approximately 28% of the total confirmed cases [8]. Meanwhile, in Malaysia, the first COVID-19 cases were reported on 25 January 2020 when two Chinese tourists from Wuhan entered Malaysia via Singapore. As of 15 December 2020, the number of confirmed cases has increased to 84,846 in Malaysia [8].

By mid-to-late March 2020, the numbers of confirmed cases in China have decreased following the implementation of various interventions such as restriction of movement order, compulsory quarantine of positive patients, cancellation of mass gathering, and other preventive measures [9] [1]. Out of all the preventive actions taken to control the spread of COVID-19, social distancing was proven to be the most effective in reducing the transmission rate [10].

The Movement Control Order (MCO) was announced on 16 March 2020 in Malaysia [11]. Since then, Malaysians needed to practice physical distancing of 1 m and they were restricted from traveling across states. It was also compulsory to wear a facial mask in public places and only one person from each family could leave homes to buy essential supplies.

This study aimed to provide a better understanding of the spatial correlation of COVID-19 occurrences by evaluating the spatial factors and measuring the

similarities in the conditions between neighbouring areas. Over the past few decades, many spatial studies based on computer programming have been published. There is a wealth of research evidence that focused extensively on mapping, especially in spatial practice [12], [13]. The mapping using geographical and socio-economic features can simplify the researcher's task to investigate the effects of the spread of cases.

One of the main epidemiological lessons in a pandemic is spatial clustering. Data on the relationship between exposure and the cases can provide valuable information. Yet, the traditional mapping practice is still not linked to robust statistical analysis. For example, spatial scan statistics, Getis's G index, and Tango' C index can only detect spatial clustering [14]. An alternative method of tackling these problems is the use of local indicator of spatial association (LISA) such as Moran's I and Geary's  $c_i$ . LISA has the advantage of being more applicable in settings whereby its components are related to the measurement of global statistics. Hence, the application of LISA enables the diagnosis of local instability to ascertain any significant global associations [15]. Previous research has shown that this index is very useful in measuring the multidimensional poverty index [16]. The study reported an evaluation of the relationship between the poverty concentration and prosperity in Gauteng between 2001 and 2011.

Since the pandemic, Moran's I has been widely applied to tackle the COVID-19 problem [17]. A study on the spread of the illness was conducted to determine the association between the hotspot areas and spatial analysis in South Africa using Moran's I. Meanwhile, another research calculated Moran's I statistics to measure the spatial autocorrelations of COVID-19 in the city of São Paulo, Brazil [18]. It studied the relationship between COVID-19 and the number of deaths, number of available hospital beds, human development index by cities, and population density. The LISA output specified Sao Paulo city as a significant hotspot for both confirmed and death cases, while other cities were considered cold spots [18].

In a more recent spatial epidemiological study of COVID-19 by [1], [19], Moran's I was applied to test the correlation between COVID-19 infection and different neighbourhoods in China. The spatial spread of the COVID-19 pandemic was observed. The results showed

that most models, except for medical care-based connection models, indicated a significant spatial association of COVID-19 infections from as early as 22 January 2020 [1]. The results showed that the spatial Moran index managed to deduce that COVID-19 had a significant spatial correlation in China. The prompt action by the Chinese government to implement lockdown at the Hubei province was effective in preventing the spread as shown by the reduction of the primary regeneration number ( $R_0$ ) to below 1. The basic regeneration number ( $R_0$ ) is a crucial component of transmissibility to explain the average number of secondary cases generated by previous infectious cases [19].

Generally, researchers from South Africa, Brazil, and China have mastered the practice of investigating the spatial autocorrelation of COVID-19 using exploratory spatial data analysis (ESDA). Compared to the Exploratory Data Analysis (EDA), ESDA is more advantageous as it can detect the spatial pattern. Specifically, the global Moran's I is used to calculate the global spatial autocorrelation whereas the local Moran's index measures the spatial autocorrelation of individual locations and provides information about the impact on the neighbouring regions or areas.

At present, this pandemic is still spreading across the globe with no signs of abating. Most of the prior work in Malaysia is limited to a subset of COVID-19 cases using statistical methods. While some other studies have been done using spatial analysis, they were limited to district mapping on certain specific issues. ESDA has also been conducted in Malaysia to analyse the spatial cluster of crime cases in Peninsular Malaysia [14]. Another study [20] provided a general view of COVID-19 based on trend distribution and temporal pattern analysis. However, no spatial statistical analysis was done. Until now, no studies have been published on the spatial autocorrelation of the COVID-19 condition in Malaysia.

With the increasing number of daily cases in Malaysia, comprehensive mapping with statistical verification is vital to provide detailed information about the pandemic. This study was conducted to analyse the spatial cluster of COVID-19 cases in Malaysia by using ESDA. The use of ESDA in this study could contribute important information towards the medical and spatial analyses of COVID-19 outbreaks in Malaysia. This study would provide a clear idea of the effects of COVID-19 based on the spatial pattern of the hotspots. Further studies are needed in the future to better delineate the impact of this pandemic on the economic, social, environmental, and health sectors of Malaysia.

## 2.0 METHODOLOGY

### 2.1 The Data

The data of COVID-19 cases from 17 April 2020 to 31st October 2020 were retrieved from the Malaysian Ministry of Health website (<http://COVID-19.moh.gov.my/>). It included data from all 144 districts

in Malaysia. The number of COVID-19 was based on the number of daily new cases registered at the hospitals and health clinics. Meanwhile, the daily global data was collected from 9 January 2020 to 14 October 2020 from the website of (<https://ourworldindata.org/coronavirus>).

Next, the number of COVID-19 cases was transformed into incidence rate per 100,000 people. The objective of transformation was to portray the actual frequency of COVID-19 cases over a specific period, as a proportion of the people at risk for the disease (population). This rate is a good reflection of the state of the pandemic in a community [21]. The incidence rate can be calculated as:

$$\begin{aligned} \text{Incidence Rate (IR)} &= \frac{\text{The number of cases}}{\text{Total population}} \times 100,000 \end{aligned} \quad (1)$$

### 2.2 The Moran's Index

In spatial analysis, the autocorrelation can be calculated as globally or locally [22]. Besides, spatial autocorrelation can provide information about each district to determine if the value of a certain indicator is similar or different from its sub-regions. Moran's I local statistics was used to calculate the localised spatial autocorrelation in the study site.

The Moran's I statistic is used to calculate the global spatial autocorrelation or clustering, it is defined as:

$$I = \frac{1}{S^2} \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^n \sum_{j=1}^n w_{ij}} \quad (2)$$

Where

$$s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (3)$$

$x_i$  represents the case number in the site  $i$  and  $x_j$  represents the case number in site  $j$  ( $j \neq i$ ).  $\bar{x}$  is an average value of  $x_i$  with  $n$  sample size.  $w_{ij}$  is represent a weight that calculates connectivity in site  $i$  with neighbour site  $j$ . This study applied rook contiguity weight (sharing a common border) as written below:

$$w_{ij} = \begin{cases} 1 & \text{if location } i \text{ and } j \text{ sharing boundary} \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

The range of potential values of  $I$  falls between -1 and 1. The positive value shows the existence of a clustered pattern where it shares the same values with its neighbour. In contrast, negative values show a spread-out pattern while the value 0 indicates no spatial autocorrelation, i.e. a random pattern. The global Moran's I value was use to check the clustering of spatial pattern in the study area globally, but it did not show the location of the cluster. Therefore, the Moran's I is used to calculate the local spatial correlation [15], as estimated by:

$$I_i = \frac{(x_i - \bar{x})}{s^2} \sum_{j=1}^n w_{ij}(x_j - \bar{x}) \quad (5)$$

Based on the estimation, a high negative value of  $I_i$  means that the site under study is a spatial outlier. In contrast, a high positive value of  $I_i$  indicates the sites are spatially clustered. There are four categories of spatial autocorrelation, namely high-high cluster, high-low cluster, low-high cluster, and low-low cluster. A positive value of  $I_i$  consists of high-high clusters and low-low clusters. Meanwhile, a negative value of  $I_i$  consists of high-low clusters and low-high clusters. A high-high cluster indicates that neighbouring sites with high values are surrounding a site with high values. Similarly, low-high can also be an outlier cluster that shows a site with a low value surrounding neighbouring sites with low values. The remaining districts that are not categorised show no significant clustering of cases [15].

In this study, the statistical significance of the Local Moran's I index was used to test the null hypothesis, i.e. there was no spatial clustering in the study area. Statistical inference was based on the 999 permutations testing of the statistical significance using the Monte Carlo simulation. The Moran's I and pseudo p-value were computed to evaluate the statistical significance. The null hypothesis would be rejected if the pseudo p-value was small [15].

GeoDa was used to construct the Moran's I index maps. The Geographic Information System, ArcGis 9.3 software (ESRI, U.S) was applied to determine the Moran's I.

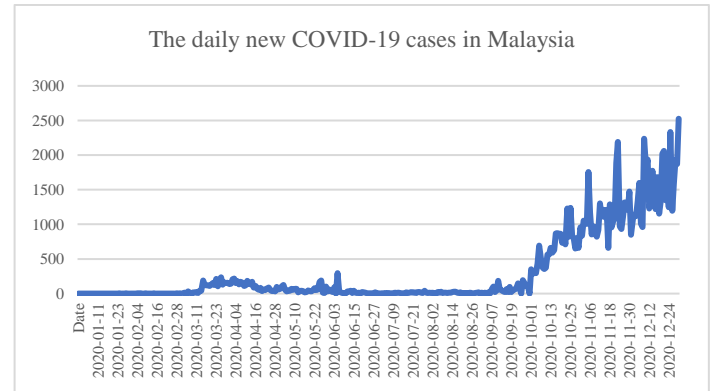
### 3.0 RESULTS AND DISCUSSION

#### 3.1 The Analysis of COVID-19 Daily Cases

As a primary analysis, Figure 1 illustrates the daily number of COVID-19 cases in Malaysia from 25 January 2020 to 31 December 2020. As shown in the figure, the distribution of the confirmed cases skewed to the right of the study period. The spike of cases began from 15 March 2020 until the end of March. After that, the cases showed an upward trend until mid-April before slowly decreasing to a plateau at the end of June 2020. However, the cases began to rise again on 27 September 2020. The highest number of cases during the study period was 2525 on 31st December 2020. Following that, the cases continued to increase significantly with daily cases higher than 2500.

There was clear evidence to show that the fluctuating number of cases during the study period was related to the preventive action taken by the government, especially the Movement Control Order (MCO) [11]. Since the outbreak at a religious mass gathering in Sri Petaling, Kuala Lumpur in early March, the number of cases increased rapidly [23]. Before MCO was implemented, the total number of cases reported was 553 by 16 March 2020. MCO was implemented as a proactive action to control the

spread of COVID-19. The first phase of MCO was enforced from 18 March 2020 until 2 April 2020. During this phase, the total number of cases stood at 2355 and the daily cases kept on increasing gradually.



**Figure 1** The daily number of active COVID-19 cases in Malaysia

Consequently, MOH extended the MCO for another 14 days until 28 April 2020. During the second phase of MCO, a total of 1909 cases were recorded. This was a decline by 18.9% compared to the first phase of MCO, thus showing that MCO was a success in controlling the outbreak. The MCO was extended for a second time until 3 May 2020 and the total number of infected patients during this third phase of MCO was 1003. This corresponded to a reduction of 47.5%. In view of the success in bringing down the cases, the fourth phase of MCO continued to be implemented until 9 June 2020. Again, this led to a 64.5% of reduction in the cases, bringing down the total cases to only 356.

After that, the Malaysian government loosened up the MCO by introducing the Recovery Movement Control Order (RMCO) [24]. During this fifth phase, certain social events could be held with minimal attendees, businesses were allowed to operate, and schools were reopened. RMCO started on 10 June 2020 until 1 July 2020. Unfortunately, the total number of cases spiked up during RMCO to 2153 cases, equivalent to an increment of 504.8%. Some of the infections were believed to originate from Malaysians who returned home from overseas but did not undergo self-quarantine.

As the third wave of the pandemic swept on in Malaysia, the number of cases increased during the sixth phase to 1005. The increment was partly driven by the fact that people were allowed to travel interstate during RMCO, except those in the Enhanced Movement Control Order (EMCO) areas [25]. On 28 August, it was announced that RMCO would be extended until 31 December 2020 except for certain areas. The Kota Setar and Tawau districts were placed under EMCO from 11 September to 23 September 2020. All the residents in the two districts were prohibited from leaving their homes. Similarly, no outsiders were allowed to enter the EMCO areas.

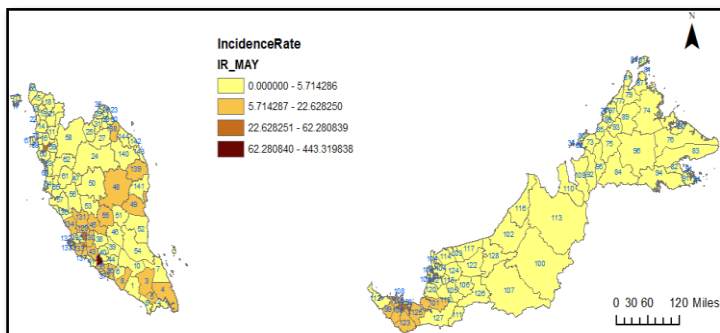


Apart from that, starting from 7 October 2020, the Conditional Movement Control Order (CMCO) was implemented in Kota Kinabalu, Penampang, Putatan in Sabah, followed by the Klang Valley on 14 October 2020. The CMCO lasted until 27 October 2020. As of 31 December 2020, Selangor recorded the highest cumulative confirmed cases of 1205 cases (82.6%), followed by Sabah with 299 cases (11.8%) and Malacca with 239 cases (9.5%). It is important to note that the analysis was performed based on the COVID-19 incidence rate to better understand the spatial pattern [26].

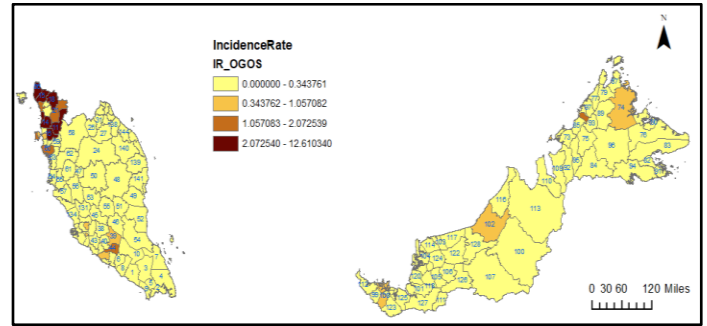
### 3.2 The Analysis of COVID-19 Monthly Incidence Rate

The mapping of COVID-19 cases was finalised by using the ArcGIS software to obtain an accurate picture of the pandemic situation. Figures 2, 3, and 4 show the COVID-19 incidence rates for May 2020, August 2020, and October 2020 in all the Malaysian states. The incidence rate mapping for other months is available in Appendix B. Next, spatial autocorrelation analysis was also performed. Specifically, the global Moran's I and local indicator of spatial association (LISA) were computed using GeoDa. The list of each district can be referred to in Appendix A.

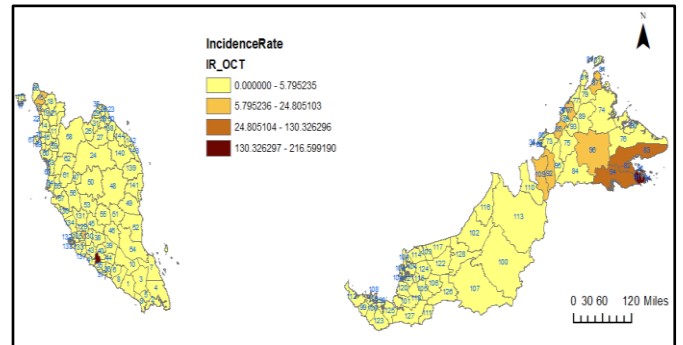
In May 2020, the highest incidence rate was recorded in Kuala Lumpur and Rembau, as shown in Figure 2. Meanwhile, in August, the COVID-19 cases were spreading north to Perlis and Kedah, involving Baling, Kuala Muda, Kubang Pasu, Kulim, Padang Terap and Yan shown in Figure 3. Furthermore, the highest incidence rate of COVID-19 was mostly reported in Sabah during October, as shown in Figure 4. For example, the districts with high incidence rates included Kota Kinabalu, Lahad Datu, Papar, Putatan, Tuaran, Labuan, Klang, Kuala Lumpur, Betong and Hilir Perak.



**Figure 2** The incidence rate COVID-19 in Malaysia for May 2020



**Figure 3** The incidence rate of COVID-19 in Malaysia for August 2020



**Figure 4** The incidence rate COVID-19 in Malaysia for October 2020

### 3.3 Spatial Analysis of COVID-19

#### 3.3.1 The Global Moran's I Analysis

Subsequent analysis was a globally autocorrelation test. Notably, the result of global Moran's I stated that the global Moran's Indices were all statistically significant with a p-value < 0.05 [15]. Table 2 displays the common trend of global spatial autocorrelation from April 2020 until December 2020. Using Geoda Software, the output gathered appeared to show positive spatial autocorrelation with significant p-values for several months, except for May 2020 and June 2020. A positive spatial autocorrelation also highlights a spatial clustering pattern with a similar value in the study area. For May 2020, there was no positive spatial autocorrelation because the p-value was not significant. The Moran's I value for June 2020 shows a negative value and with a non-significant p-value (0.306). Thus, it shows that the greater the distances between the districts, they more they tended to be more dissimilar as compared to being next to each other. In other words, there existed a spatial outlier. The Global Moran's I output in Table 2 presented the differences in COVID19 cases across space. Later, local indicator spatial autocorrelation (LISA) is required to measure the local spatial cluster for the district individually.

**Table 2** The Global Moran's I of COVID19

Month	Moran's I	p-value
Apr	0.189	(0.010)
May	0.014	(0.100)
Jun	-0.006	(0.306)
Jul	0.164	(0.017)
Aug	0.401	(0.001)
Sep	0.641	(0.001)
Oct	0.381	(0.002)
Nov	0.380	(0.002)
Dec	0.360	(0.001)

### 3.3.2 The Local Indicator Spatial Autocorrelation (LISA) Analysis

If spatial heterogeneity existed, then the global spatial autocorrelation could not imitate the local spatial correlation within the geographic unit. To overcome this, LISA analysis was used to identify the heterogeneity of spatial association across different geographic units [27]. Based on Figures 5, 6, and 7, the areas with darker red colour represented the hotspots for three consecutive months while the areas with a darker blue colour were cool spots. The hotspot area is defined as an area with a positive and statistically significant p-value in the same month throughout the study. The hotspot area for COVID-19 incidence rate means that the area with a high number of cases also tends to have neighbouring places with high cases (high-high cluster) or vice versa (low-low cluster).

Additionally, the nine months of study period were divided into three interval periods, namely April to June, July to September, and October to December to investigate the existence of cluster similarity (hotspot) for three consecutive months throughout the observed period. The map of hotspots was constructed by overlapping all the hotspots of COVID-19 incidence rate in April until June 2020 (Figure 5), July until September 2020 (Figure 6), and October until December 2020 (Figure 7). For the districts of Labuan and Langkawi that did not share borders with any other districts, they were excluded for the analysis.

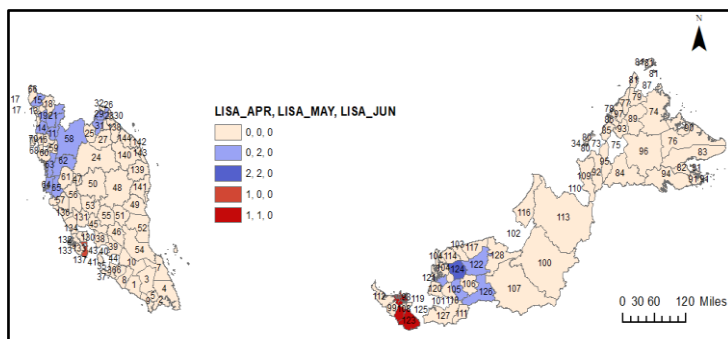
**Figure 5** The cluster map of the total number of cases for April and May 2020

Figure 5 shows a total of five types of time-based hotspots based on the incidence rate of COVID-19 in

Malaysia from April 2020 to June 2020. However, June 2020 was labelled as "0" as its result was insignificant and there was no similarity. The hotspots of these three months were overlapped to be categorised into the five types as shown in Figure 5. A district was categorised as "1, 1, 0" if it fell under the hotspot group for two consecutive months (April 2020 and May 2020). If the output was "1, 0, 0", it indicated that these areas were designated as hotspots in April 2020. As for the "2, 0, 0" group, the district was a cool spot for April 2020 while "2, 2, 0" referred to the district being a cool spot for two consecutive months (April 2020 and May 2020). The outcome of "0, 0, 0" showed that the sites were not hotspots in the three consecutive months.

There were 17 districts in the four categories of hotspots, whereas another 104 districts were not categorised as a hotspot of COVID-19 at all during the three consecutive months. Figure 5 shows the two districts under the "1, 1, 0" and one district under the "1, 0, 0" hotspot groups respectively. For the "1, 1, 0" group, Serian and Kuching in the state of Sarawak remained as hotspots for two consecutive months (April 2020 and May 2020). Sepang in Selangor was under the "1, 0, 0" group that became hotspot for April 2020. Thus, it can be concluded that the LISA analysis provided significant local results for some district even though the global Moran's I result for May 2020 was insignificant (Table 2).

For the cool spots, the Low-Low cluster (blue) refers to a spatial association group whose regions showed a low incidence rate, i.e. below the average, surrounded by regions that also had low values. In Figure 5, Sibul registered under the "2, 2" group, which meant that Sibul became a cool spot for April and May 2020. A total of 17 districts registered under the "0,2" group, indicating that the district became a cool spot for May 2020. The districts were Baling, Kubang Pasu, Pedang, Sik, Kota Bharu, Pasir Mas, Tanah Merah, Hulu Perak, Kuala Kangsar, Larut, Matang and Selama, Manjung, Perak Tengah, Beaufort, Keningau, Sarikei, Selang and Song. Due to the MCO, the number of hotspots was under control even during the Eid Celebration from 24th to 26 April 2020. The citizens were encouraged to have a simple celebration in their own homes without visiting others or traveling back to the hometowns.

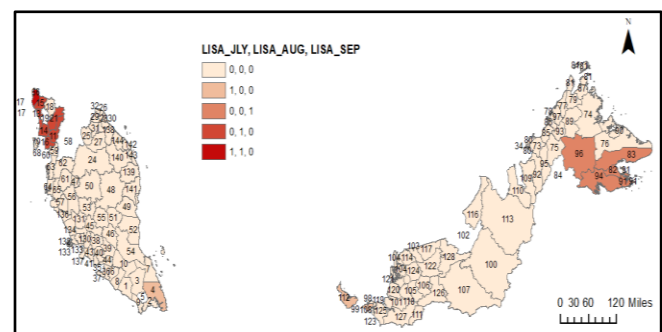
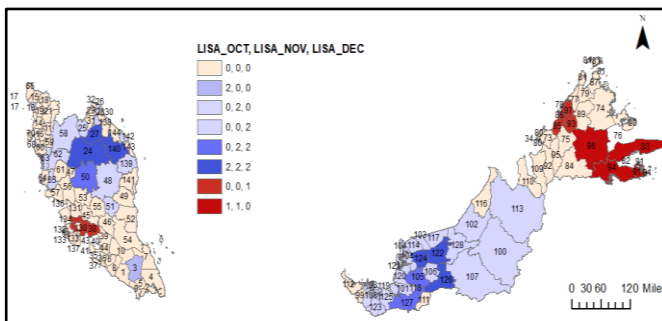
**Figure 6** The cluster map of COVID-19 cases for July 2020, August 2020 and September 2020

Figure 6 presents the cluster map of COVID-19 cases based on Moran's I in July, August, and September 2020. Only one district registered as a hotspot in the "1,1,0" category, i.e. Perlis. Perlis was a hotspot in July and August 2020. No district registered as a repetitive hotspot for July, August, and September 2020. For July 2020 hotspot, three districts registered under the "1, 0, 0" category, including Kota Tinggi, Lundu, and Kota Samarahan.

However, about seven areas in northern Malaysia were reported as hotspots in the "0, 1, 0" category for August 2020, including Kulim, Yan, Baling, Kota Setar, Kubang Pasu, and Kuala Muda. These hotspots resulted from the Sivagangga cluster that started from Kedah before spreading to neighbouring Penang and Perlis. The index case of this cluster was a restaurant owner. He returned from India and he did not follow the home quarantine rules [28] (The Straight Times, 2020). In addition, seven districts registered under the "0, 0, 1" category, namely Lahad Datu, Kunak, Semporna, Tawau, and Tongod that became hotspots in September 2020. During this period, a total of 108 districts was in the "0, 0, 0" category and never registered as a hotspot.

These findings might have been the main reason behind the government's decision to place 35,643 citizens in the six villages of Semporna, Sabah under EMCO on 14 October 2020. During the 14 days of EMCO, the Sabah government took initiatives to distribute food and daily essentials for these villagers since many sectors were prohibited from operating. Furthermore, the National Security Council sent in extra army and police personnel to tighten the security at the east coast of Sabah. Various restrictions were imposed on the movement of boats, ships, and vessels along the coastlines [29].



**Figure 7** The cluster map of COVID-19 cases for October 2020, November 2020 and December 2020

Figure 7 shows that four districts in Sabah, namely Lahat Datu, Semporna, Tongod, and Tawau were registered in the "1, 1, 0" category as hotspots for October and November 2020. Another ten areas were also listed as hotspots under the "0, 0, 1" category during December 2020. These districts included Jelebu in Negeri Sembilan, Kota Kinabalu, Penampang, Papar, Tambunan, Tuaran, and Putatan in Sabah, Gombak, Hulu Langat, as well as Petaling in Selangor. Thus, the

majority of the hotspot districts during this period were from Sabah, Malaysia.

For the cool spots, six districts registered under the "2, 2, 2" category from October 2020 until December 2020. These included Gua Musang and Kuala Krai in Kelantan, Hulu Terengganu in Terengganu, Selangau, Sibul, and Song in Sabah. Besides, Lipis in Pahang, Julau, Pakan, and Sri Aman in Sarawak were registered under the "0, 2, 2" category that became cool spots during November 2020 and December 2020. Only the Kluang district registered as a cool spot for October 2020. Meanwhile, Machang, Kuala Kangsar, Perak Tengah, and Kota Bharu registered as cool spots for November 2020. For December 2020, 27 districts registered as cool spots under the "0, 0, 2" category. During this period, a total of 79 districts did not register as a hotspot at all.

In conclusion, in early 2020, the spread of the pandemic in Malaysia centred in the Serian and Kuching in East Malaysia, the capital city, as well as Putrajaya and Sepang in Selangor. Subsequently, the incidence rate was highest in the northern part of Malaysia from July 2020 until September 2020, including Kedah and Perlis. Later, at the end of 2020, the incidence rate of COVID-19 increased throughout many of the districts in Sabah. This spatial analysis established that the neighbourhood features and surrounding circumstances significantly affected the number of cases in the areas. Thus, it is vital to recognise and investigate all the possible issues to obtain a better picture of COVID-19 status in each region. While this study was limited to spatial autocorrelation mapping of COVID-19, future studies can also be expanded to include other statistical validation analysis.

## 4.0 CONCLUSION

In short, this research conveyed the empirical outcome of the spatial variations in the COVID-19 incidence rate of all districts in Malaysia. The findings from this spatial analysis helped to identify appropriate prevention and control measures at each level. Moreover, hotspot regions in the country were recognised and described accordingly to provide a better description of the pandemic situation in Malaysia. This research also provided detailed information on the districts that were spatially affected by the hotspot areas and cool spots based on the Cluster Moran's I analysis. In any outbreaks, it is essential to investigate the neighbourhood areas to understand the pattern of spread with immediate actions can be taken to control the outbreak, especially in high-risk areas. For COVID-19, its socioeconomic impact on all districts and neighbouring regions must be investigated. Such information can be beneficial for policymakers, agencies, and security forces in the country to undertake proper resource allocation and disease management in high-risk regions to cope with the pandemic effect.

Spatial analysis can be a valuable instrument to forecast the upcoming trend of COVID-19. It is often

difficult to generate an accurate prediction on the number of cases because it is contingent on socioeconomic, demographic, and environmental issues. So far, no artificial intelligence program is available on the forecasting of COVID-19 conditions. However, analysis of retrospective data can be valuable in predicting future pandemics. The analysis can be focused on certain populations with the inclusion of good environmental and socioeconomic background data. Thus, systematic spatial data collection is essential. Lastly, this was one of the first studies to explore the spatial pattern of COVID-19 incidence rate. In future, an exploration of the relationship between COVID-19 incidence rate and other independent variables using statistical modelling can be considered.

### Acknowledgement

The data was obtained from the Malaysian Ministry of Health website (<http://COVID-19.moh.gov.my/>) and the international database at (<https://ourworldindata.org/coronavirus>).

### References

- [1] Kang, D., Choi, H., Kim, J.-H., & Choi, J. 2020. Spatial Epidemic Dynamics of the COVID-19 Outbreak in China. *International Journal of Infectious Diseases*. 94(1): 96-102. <https://doi.org/10.1016/j.ijid.2020.03.076>.
- [2] Disease outbreak news. 2020. The World Health Organization (WHO). <https://www.who.int/csr/don/05-january-2020-pneumonia-of-unknown-cause-china/en/>.
- [3] World Health Organization. 2020. Laboratory Testing for 2019 Novel Coronavirus (2019-nCoV) in Suspected Human Cases. The World Health Organization (WHO). Retrieved from <https://www.who.int>.
- [4] Lu, C., Wei, Liu, X. fen, & Jia, Z. Fang. 2020. 2019-nCoV Transmission through the Ocular Surface Must Not Be Ignored. *The Lancet*. 395(10224): e39. <https://doi.org/10.1016/S0140-6736>.
- [5] Wang, D., Hu, B., Hu, C., Zhu, F., Liu, X., Zhang, J., Wang, B., Xiang, H., Cheng, Z., Xiong, Y., Zhao, Y., Li, Y., Wang, X., & Peng, Z. 2020. Clinical Characteristics of 138 Hospitalised Patients With 2019 Novel Coronavirus-infected Pneumonia in Wuhan, China. *JAMA*. 323(11): 1061-1069. <https://doi:10.1001/jama.2020.1585>.
- [6] Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., Zhang, L., Fan, G., Xu, J., Gu, X., Cheng, Z., Yu, T., Xia, J., Wei, Y., Wu, W., Xie, X., Yin, W., Li, H., Liu, M., ... Cao, B. 2020. Clinical Features of Patients Infected with 2019 Novel Coronavirus in Wuhan, China. *The Lancet*. 395(10223): 497-506. [https://doi.org/10.1016/S0140-6736\(20\)30183-5](https://doi.org/10.1016/S0140-6736(20)30183-5).
- [7] Guan, W. J., Ni, Z. Y., Hu, Y., Liang, W. H., Ou, C. Q., He, J. X., ... & Du, B. 2020. Clinical Characteristics of Coronavirus Disease 2019 in China. *New England Journal of Medicine*. 382(18): 1708-1720. [https://doi.org/10.1016/S0140-6736\(20\)30183-5](https://doi.org/10.1016/S0140-6736(20)30183-5).
- [8] Coronavirus Disease (COVID19) Dashboard. 2020. <https://COVID19.who.int/>.
- [9] Chang, S. L., Harding, N., Zachreson, C., Cliff, O. M., & Prokopenko, M. 2020. Modelling Transmission and Control of the COVID-19 Pandemic in Australia. arXiv preprint arXiv:2003.10218. <http://doi.org/10.1038/s41467-020-19393-6>.
- [10] Götz, T., & Heidrich, P. 2020. Early-stage COVID-19 Disease Dynamics in Germany: Models and Parameter Identification. *Journal of Mathematics in Industry*. 10(1). <https://doi.org/10.1186/s13362-020-00088-y>.
- [11] Shah, A. U. M., Safri, S. N. A., Thevadas, R., Noordin, N. K., Rahman, A. A., Sekawi, Z., Ideris, A., & Sultan, M. T. H. 2020. COVID-19 Outbreak in Malaysia: Actions Taken by the Malaysian Government. *International Journal of Infectious Diseases*. 97: 108-116. <https://doi.org/10.1016/j.ijid.2020.05.093>.
- [12] Chaikaew, N., Tripathi, N.K. & Souris, M. 2009. Exploring Spatial Patterns and Hotspots of Diarrhea in Chiang Mai, Thailand. *Int J Health Geogr*. 8: 36. <https://doi.org/10.1186/1476-072X-8-36>.
- [13] Berke, O. 2001. Choropleth Mapping of Regional Count Data of Echinococcus Multilocularis among Red Foxes in Lower Saxony, Germany. *Preventive Veterinary Medicine*. 52: 119-131. [https://doi.org/10.1016/S0167-5877\(01\)00246-X](https://doi.org/10.1016/S0167-5877(01)00246-X).
- [14] Zakaria, S., & Abd. Rahman, N. 2014. Analysing the Violent Crime Patterns In Peninsular Malaysia : Exploratory Spatial Data Analysis (ESDA) Approach. *Jurnal Teknologi*. 72(1): 131-136. <https://doi.org/10.11113/jt.v72.1816>.
- [15] Anselin, L. 1995. Local Indicators of Spatial Association – LISA. *Geographical Analysis*. 27(2): 93-115. <https://doi.org/10.1111/j.1538-4632.1995.tb00338.x>.
- [16] Katumba, S. 2018. Spatial Statistical Analyses to Assess the Spatial Extent and Concentration of Multidimensional Poverty in Gauteng using the South African Multidimensional Poverty Index. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*. 42(4W8): 85-92. <https://doi.org/10.5194/isprs-archives-XLII-4-W8-85-2018>.
- [17] Arashi, M., Bekker, A., Salehi, M., Millard, S., Erasmus, B., Cronje, T., & Golpaygani, M. 2020. Spatial Analysis and Prediction of COVID-19 Spread in South Africa after Lockdown. ArXiv. <http://arxiv.org/abs/2005.09596>.
- [18] Alcantara, E., Mantovani, J., Rotta, L., Park, E., Rodrigues, T., Carvalho, F., & Souza Filho, C. 2020. Investigating Spatio Temporal Patterns of the COVID-19 in Sao Paulo State, Brazil. 1-15. <https://doi.org/10.1101/2020.05.28.20115626>.
- [19] Huang, R., Liu, M., & Ding, Y. 2020. Spatial-temporal Distribution of COVID-19 in China and its Prediction: A Data-driven Modeling Analysis. *The Journal of Infection in Developing Countries*. 14(03): 246-253. <https://doi.org/10.3855/jidc.12585>.
- [20] Yahya, M. S. S., Safian, E. E. M., & Burhan, B. 2020. The Trend Distribution and Temporal Pattern Analysis of COVID-19 Pandemic using GIS framework in Malaysia. *AJIR Preprints*. July, 1–14. <https://doi.org/10.21467/preprints.174>.
- [21] Tosepu, R., Gunawan, J., Effendy, D. S., Lestari, H., Bahar, H., & Asfian. 2020. P. Correlation between Weather and Covid-19 Pandemic in Jakarta, Indonesia. *Science of the Total Environment*. 725: 138436.
- [22] Moran, P. A. P. 1950. Notes on Continuous Stochastic Phenomena. *Biometrika*. 37: 17-23. <https://doi.org/10.1093/biomet/37.1-2.17>.
- [23] Koh Boon Yau, E., Pang Tze Ping, N., Shoosmith, W. D., James, S., Nor Hadi, N. M., & Loo, J. L. 2020. The Behaviour Changes in Response to COVID-19 Pandemic within Malaysia. *The Malaysian Journal of Medical Sciences: MJMS*. 27(2): 45-50. <https://doi.org/10.21315/mjms2020.27.2.5>.
- [24] Amir Yusof. 2020. Malaysia's Recovery Movement Control Order Extended to 31 December 2020, Tourists Still Not Allowed in: PM Muhyiddin. *CAN Insider*. 28 August 2020. <https://www.channelnewsasia.com/news/asia/malaysia-recovery-movement-control-order-extended-dec-31-mco-13047724>.
- [25] Bernama. Interstate Travel among Activities Allowed from Wednesday – Muhyiddin. *The Sun Daily*. 7 June 2020. <https://web.archive.org/web/20200608041411/https://www.thesundaily.my/home/interstate-travel-among-activities-allowed-from-wednesday-muhyiddin-HN2539119>.
- [26] Dg of Health. Kenyataan Akhbar KPK 31 Disember 2020 – Situasi Semasa Jangkitan Penyakit Coronavirus 2019 (COVID-19) di Malaysia. 31 December 2021. <https://kpkkesihatan.com/2020/12/31/kenyataan-akhbar>



- kpk-31-disember-2020-situasi-semasa-jangkitan-penyakit-coronavirus-2019-covid-19-di-malaysia/.
- [27] Anselin, L. 1988. *Spatial Econometric: Methods and Models*. Dordrecht.
- [28] Anselin, L. 1999., Rao's Score Test in Spatial Econometrics. *Journal of Statistical Planning and Inference*. 97: 113-139.
- [29] The Straight Times. 8 August 2020. Sivagangga COVID-19 Cluster Spreads Out of Kedah in Malaysia. The Straits Times. <https://www.straitstimes.com/asia/se-asia/sivagangga-COVID-19-cluster-spreads-out-of-kedah-in-malaysia>.
- [30] Nor Ain Mohamed Radhi and Dhesegaan Bala Krishnan. 9 October 2020. 6 Areas in Semporna, One in Kunak under EMCO from 14 October [NSTTV]. News Straight Times. <https://www.nst.com.my/news/nation/2020/10/630952/6-areas-semporna-one-kunak-under-emco-oct-14>.

## Appendix A

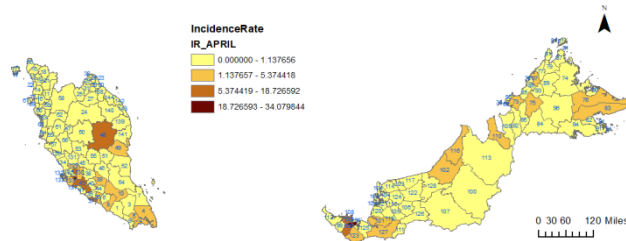
Number	District
1	Batu Pahat
2	Johor Bharu
3	Kluang
4	Kota Tinggi
5	Kulai
6	Tangkak
7	Mersing
8	Muar
9	Pontian
10	Segamat
11	Baling
12	Bandar Baharu
13	Kota Setar
14	Kuala Muda
15	Kubang Pasu
16	Kulim
17	Langkawi
18	Padang Terap
19	Pendang
20	Pokok Sena
21	Sik
22	Yan
23	Bachok
24	Gua Musang
25	Jeli
26	Kota Bharu
27	Kuala Krai
28	Machang
29	Pasir Mas
30	Pasir Putih
31	Tanah Merah
32	Tumpat
33	Kuala Lumpur
34	Labuan
35	Alor Gajah
36	Jasin
37	Melaka Tengah
38	Jejebu
39	Jempol
40	Kuala Pilah
41	Port Dickson

42	Rembau
43	Seremban
44	Tampin
45	Bentong
46	Bera
47	Cameron Highlands
48	Jerantut
49	Kuantan
50	Lipis
51	Maran
52	Pekan
53	Raub
54	Rompin
55	Temerloh
56	Batang Padang
57	Hilir Perak
58	Hulu Perak
59	Kampar
60	Kerian
61	Kinta
62	Kuala Kangsar
63	Larut, Matang, Selama
64	Manjung
65	Perak tengah
66	Perlis
67	Barat daya
68	Seberang Perai Selatan
69	Seberang Perai Tengah
70	Seberang perai Utara
71	Timur Laut
72	Putrajaya
73	Beaufort
74	Beluran
75	Keningau
76	Kinabatangan
77	Kota Belud
78	Kota Kinabalu
79	Kota Marudu
80	Kuala Penyu
81	Kudat
82	Kunak
83	Lahad datu

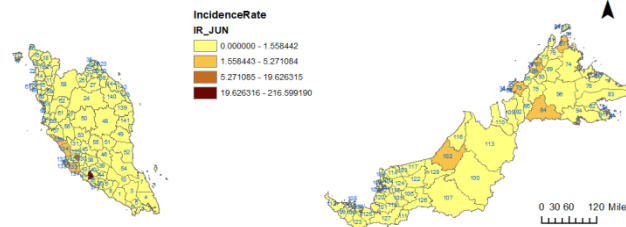
84	Nabawan
85	Papar
86	Penampang
87	Pitas
88	Putatan
89	Ranau
90	Sandakan
91	Semporna
92	Sipitang
93	Tambunan
94	Tawau
95	Tenom
96	Tongod
97	Tuaran
98	Asajaya
99	Bau
100	Belaga
101	Betong
102	Bintulu
103	Dalat
104	Daro
105	Julau
106	Kanowit
107	Kapit
108	Kuching
109	Lawas
110	Limbang
111	Lubok Antu
112	Lundu
113	Marudi
114	Matu

115	Meradong
116	Miri
117	Mukah
118	Pakan
119	Kota Samarahan
120	Saratok
121	Sarikei
122	Selangau
123	Serian
124	Sibu
125	Simunjan
126	Song
127	Sri Aman
128	Tatau
129	Gombak
130	Hulu Langat
131	Hulu Selangor
132	Klang
133	Kuala Langat
134	Kuala Selangor
135	Petaling
136	Sabak Bernam
137	Sepang
138	Besut
139	Dungun
140	Hulu Terengganu
141	Kemaman
142	Kuala Terengganu
143	Marang
144	Setiu

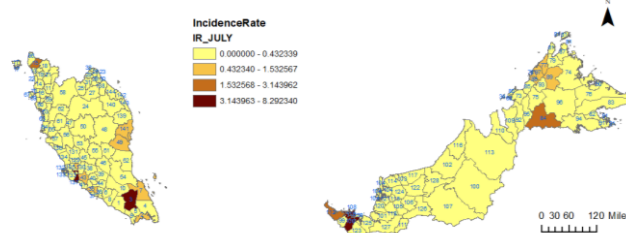
### Appendix B



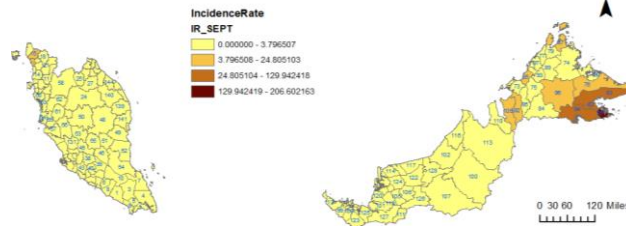
The incidence rate COVID-19 in Malaysia for April 2020



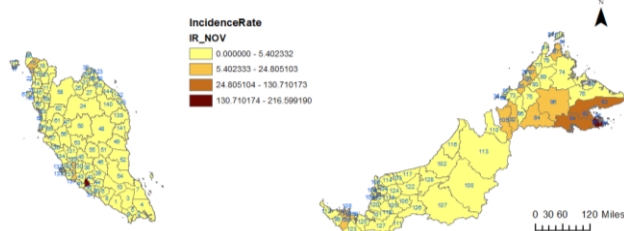
The incidence rate COVID-19 in Malaysia for June 2020



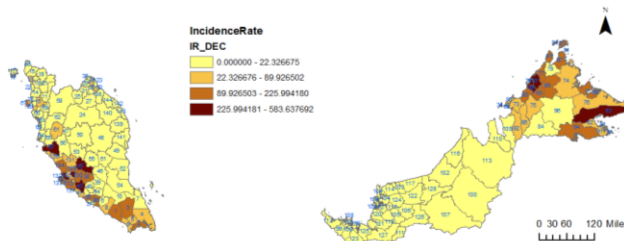
The incidence rate COVID-19 in Malaysia for July 2020



The incidence rate COVID-19 in Malaysia for September 2020



The incidence rate COVID-19 in Malaysia for November 2020



The incidence rate COVID-19 in Malaysia for December 2020