RESEARCH PAPER

USING MORAN'S I FOR DETECTION AND MONITORING OF THE COVID-19 SPREADING STAGE IN THAILAND DURING THE THIRD WAVE OF THE PANDEMIC

Parichat Wetchayont[1*], Katawut Waiyasusri[2]

[1]Department of Geography, Faculty of Social Sciences, Srinakharinwirot University, 114 Sukhumvit 23, Khlong Toei Nuea, Wattana District, Bangkok 10110 Thailand

[2]Geography and Geo-Informatics Program, Faculty of Humanities and Social Sciences, Suan Sunandha Rajabhat University, 1 U-Thong Nok Road, Dusit, Bangkok 10300 Thailand

***Corresponding author:** parichatw@g.swu.ac.th

Received: August 31th, 2021 / Accepted: November 9th, 2021 / Published: December 31st, 2021

https://doi.org/10.24057/2071-9388-2021-090

ABSTRACT. Spatial distribution and spreading patterns of COVID-19 in Thailand were investigated in this study for the 1 April – 23 July 2021 period by analyzing COVID-19 incidence's spatial autocorrelation and clustering patterns in connection to population density, adult population, mean income, hospital beds, doctors and nurses. Clustering analysis indicated that Bangkok is a significant hotspot for incidence rates, whereas other cities across the region have been less affected. Bivariate Moran's I showed a low relationship between COVID-19 incidences and the number of adults (Moran's I = 0.1023-0.1985), whereas a strong positive relationship was found between COVID-19 incidences and population density (Moran's I = 0.2776-0.6022). Moreover, the difference Moran's I value in each parameter demonstrated the transmission level of infectious COVID-19, particularly in the Early (first phase) and Spreading stages (second and third phases). Spatial association in the early stage of the COVID-19 outbreak in Thailand was measured in this study, which is described as a spatio-temporal pattern. The results showed that all of the models indicate a significant positive spatial association of COVID-19 infections from around 10 April 2021. To avoid an exponential spread over Thailand, it was important to detect the spatial spread in the early stages. Finally, these findings could be used to create monitoring tools and policy prevention planning in future.

KEYWORDS: COVID-19, Spatio-temporal, Detection, Moran's I, Socioeconomic, Health care

CITATION: Parichat Wetchayont, Katawut Waiyasusri (2021). Using Moran's I For Detection And Monitoring Of The Covid-19 Spreading Stage In Thailand During The Third Wave Of The Pandemic. Geography, Environment, Sustainability, Vol.14, No 4, p. 155-167 https://doi.org/10.24057/2071-9388-2021-090

ACKNOWLEDGEMENTS: We appreciate everyone who worked hard in the healthcare system to combat the COVID-19 epidemic. The authors are sincerely grateful for supporting from the Faculty of Social Sciences, Srinakharinwirot University and the Digital Government Development Agency (DDC) of Thailand for providing daily COVID-19 confirmed case data which was used in the study. We gratefully acknowledge the National Statistical Office of Thailand for providing the demographic data. This study was performed under the Earth and Atmospheric Research Laboratory (EAR).

Conflict of interests: The authors reported no potential conflict of interest.

INTRODUCTION

On 13 January 2020, the Thailand government reported they had first detected a tested positive of coronavirus disease (COVID-19) from a Wuhan resident who had travelled to Bangkok. A COVID-19 outbreak then rapidly spread from Bangkok into all provinces of Thailand. To prevent the coronavirus from spreading, Thailand's government announced a nationwide lockdown and curfew from March 26 to May 31, 2020. These actions immediately changed the daily life activities of Thai people, had impacts on the environment, pollution, the economy, jobs and other things, and has since been the subject of various debates (Wetchayont 2021, Wetchayont et al. 2021). Figure 1 illustrates a time series plot of daily confirmed new cases (DDC 2021). Considering that COVID-19 was spreading like an ocean wave, the first wave of COVID-19 in Thailand, which began in March 2020 and extended across 68 provinces, had initially started at boxing events and nightclubs in Bangkok. By May 25 2020, the number of total confirmed cases had reached 3,042 cases with 57 deaths (Rajatanavin et al. 2021). The second wave was spurred by some Thai workers who entered Thailand illegally from Myanmar and were not apprehended by state quarantine. They transported the infection and disseminated it across Thailand's Northern provinces. In addition, a huge number of infected migrant laborers traveled directly and illegally from Myanmar to work in industries and seafood markets in Bangkok's neighboring province of Samut Sakhon. Between December 18, 2020 and February 27, 2021, new cases rise from 3,042 cases in the first wave to 21,584 cases in the second wave. However, daily incidences had reduced to less than 100 by the last week of February 2021, indicating that the problem was progressively being brought under control (Rajatanavin et al. 2021). The most recent wave of the coronavirus outbreak in Thailand is now number three. As of 23 July 2021, 467,707 cases of COVID-19 were officially

confirmed in Thailand, including 3,811 deaths, with a record high of 14,575 new COVID-19 cases and 114 fatalities over one 24 hour period. For every day since 10 April 2021, the total number of COVID-19 cases in Thailand increased rapidly due to the Songkran holiday because of a new variant virus. The Songkran holiday, which runs from April 10 to 15, saw a large number of workers return to their hometowns from Bangkok and other big cities where infection rates were high, causing COVID-19 to spread nationwide. Over 2020 Thailand had a commendably low number of cases and successfully slowed the rate of infections by using strict measures. Nevertheless, on this current wave, the infection rate is rising exponentially and has not stopped up till now. This has caused a hospital bed shortage and left medical officials exploring home isolation for asymptomatic cases. Regular hospitals have been advised to treat only those patients with moderate symptoms or severe conditions. Early in the first wave of this COVID-19 pandemic, sex differences in the incidence of the disease were observed. The incidence rate in males was about 15% higher than in females in younger age groups, as shown in Figure 2. The same results were also found in the European region. A study by Green et al. (2021) provided evidence of sex and age differences in the case-fatality rates (CFR) of infection from COVID-19 in seven countries throughout Europe. They found that higher CFRs in males within younger age groups could be related to hormonal factors. Meanwhile, in the second wave, females tended to have higher incidence rates by about 30%, but then the incidence rates came down to become equal. There could be several factors responsible for the increased number of cases in the third wave. One of the reasons is that the mutant virus is more effective at transmitting and has a shorter incubation period (Li et al. 2021).

Previous studies have evaluated the geographical heterogeneity of chronic respiratory diseases (CRDs) in Thailand by using spatial statistics and found that high cluster areas of CRDs associate with population and industrial booms which generally contribute to epidemics (Laohasiriwong et al. 2018). Other studies have looked at the spatial spread of severe acute respiratory syndrome (SARS) in Beijing and mainland China (Meng et al. 2005; Fang et al. 2009). The findings show that there were different spatial connections between provinces in terms of possible pathways for the spread. Recently, there have been some studies that explored spatiotemporal clustering patterns of COVID-19 outbreaks at the provincial level by using global and local Moran's I indices in mainland China, which were also attributed to changes in social and demographic factors (Wang et al. 2021; Kang et al. 2020; Li et al. 2020; Zhang et al. 2020). COVID-19 incidence rates showed spatial connections between the district level and socioeconomic determinants, according to studies conducted in Brazil and Iran, which employed the global Moran's index and the Local Index of Spatial Association (LISA) as tools (Raymundo et al. 2021; Ramírez-Aldana et al. 2020). Moreover, a study among European nations, discovering a substantial positive association between income/population and COVID-19 cases/deaths, implied that these two variables could be important control variables for assessing COVID-19-related human casualties (Sannigrahi et al. 2020).

There was an assessed effort over temporal and spatiotemporal reproduction numbers to track transmission of COVID-19 dynamics in Thailand during the first wave (Rotejanaprasert et al. 2020). Results showed that the outbreak could contain reproduction numbers when it was under the control of strict measures. However, the real situation was difficult to determine in the early transmission stage due to limited incidence numbers. Additionally, Triukose et al. (2021) reported on the causes and lessons learned from the first wave of COVID-19 spread in Thailand. Their study used effective reproduction numbers to characterize the spread of COVID-19 across the country in five stages. Their results revealed that COVID-19 was mostly limited to Bangkok in the Early stage, but then started to cross over Bangkok's borders during the Spreading stage. After that, it spread nationwide, primarily due to a massive movement to people's hometowns after the Bangkok lockdown on 26 March 2020 in the Intervention I stage. In the Intervention II stage, the infectious rate downturned and became stable. Finally, the average of daily confirmed cases declined to zero and held the country in the Easing stage.

So far, the number of studies on spatial association of COVID-19's spread in Thailand are limited. Understanding the early stages of the COVID-19 outbreak's spatial spread is critical, as it may aid in the prediction of local epidemics and the development of public health policies based mainly on outbreak data (Kang et al. 2020). Even if those studies show how well temporal and spatial reproduction numbers track the transmission of COVID-19 dynamics in Thailand, in order to utilize these tools for tracking a spreading wave, or for preventive and crisis management policy planning, it requires extensive study to assess the effect of other variables on the spread of COVID-19. Thus, the goal of this study is to detect early transmission by using spatial distribution of COVID-19 incidence in Thailand's provinces and its relationship with sociodemographic factors. The expected results of this study could gain better understanding of the social environment and the epidemic's spread for utilizing tools needed to inform disease control activities in the future.

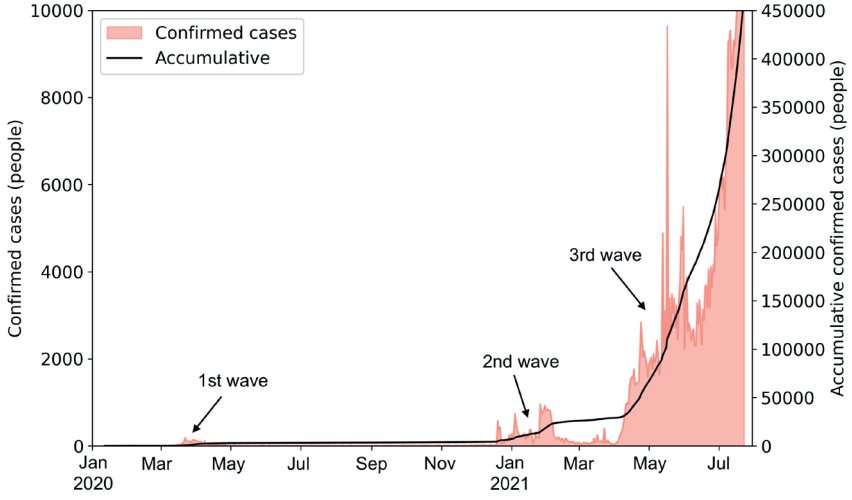

Fig. 1. Trend chart of new confirmed cases in Thailand during the period of January 1, 2020 to July 23, 2021. The first (March-May 2020), second (January-February 2021) and third waves (April-July 2021) are indicated by the daily new number of confirmed cases

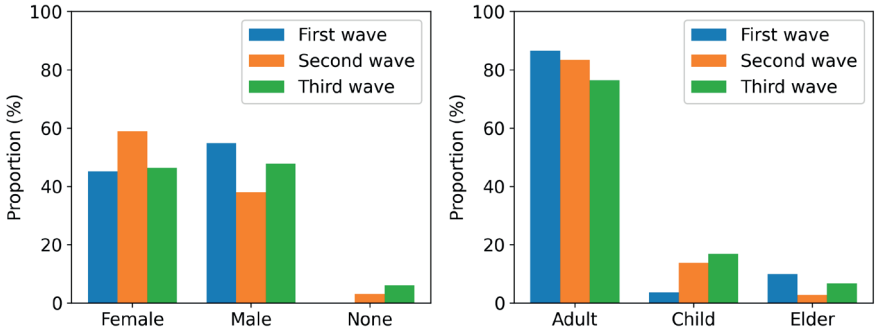

Fig. 2. Distribution by gender and age groups of patients admitted for COVID-19 during the first (March-May 2020), second (January-February 2021) and third waves (April-July 2021)

MATERIALS AND METHODS

Data collections

The number of confirmed COVID-19 cases, as registered by the Department of Disease Control of Thailand and updated daily, were obtained from their Thai website at the Digital Government Development Agency (https://data.go.th/dataset/covid-19-daily). There are 76 provinces and one special administrative area (Bangkok) in Thailand. This study used 4 months' data from 1 April to 19 July 2021, which was during the early stages of the third wave of COVID-19 in Thailand. Other datasets were obtained from the National Statistical Office of Thailand (NSO) website, including population, population density, number of doctors, and hospital beds by province (NSO 2021). All population-related and medical resource datasets were collected in 2020, making this the most up-to-date demographic information available (Table 1).

Figure 3 depicts a province-by-province map of monthly cumulative cases. The total number of instances is the sum of newly confirmed cases from 1 April–23 July 2021. The largest number of cases was in Bangkok, which is the capital city. Figure 4(a) presents the total population and population density (population/km^2) for each province in 2020. Metropolitan Bangkok had the highest population density, number of adult people and mean household income. As shown in Figure 4(d) – (f), Bangkok also has the highest number of hospital beds, doctors and nurses, whereas Chiang Mai and Nakhon Ratchasima Provinces ranks third and fifth for the number of hospital beds and doctors, respectively.

Data analysis

The number of COVID-19 cases as confirmed from the laboratory in the 76 provinces and one special administrative area (Bangkok) in Thailand, including the total population in each province, were used to compute the incidence rates in this section. The cumulative number of all reported cases between 1 January and 23 July 2021, as announced by the Department of Disease Control (DDC) of Thailand on July 23, 2021 (DDC 2021), were calculated to get the incidence of COVID-19 in each province according to the following formula:

$$COVID-19\ incidence = confirmed\ case / total\ population \quad (1)$$

The most commonly used measure of global spatial autocorrelation is Moran's I, which represents the overall distribution of deviations from randomness (Tu and Xia 2008). To provide information on spatial clusters and outliers, as well as forms of spatial correlation, we presented both global and local Moran's I. By using global Moran's Index (global Moran's I) based on Queen's contiguity spatial-lag of order 1: immediate neighbors (Moran 1950), the COVID-19 incidence rates in Thailand for each month from April to July 2021 were used as a variable. Global Moran's I is a spatial autocorrelation test that determines whether or not the spatial patterns of COVID-19 incidence are clustered by considering that space and location presents an influence on a single variable. Global Moran's I and P test values were used to determine the number of incidence rates with different types of neighbors using the spatial weight matrix, based on geographical adjacency at a 5% level of significance. Spatial autocorrelation was measured using the global Moran's I statistic, which was derived as follows:

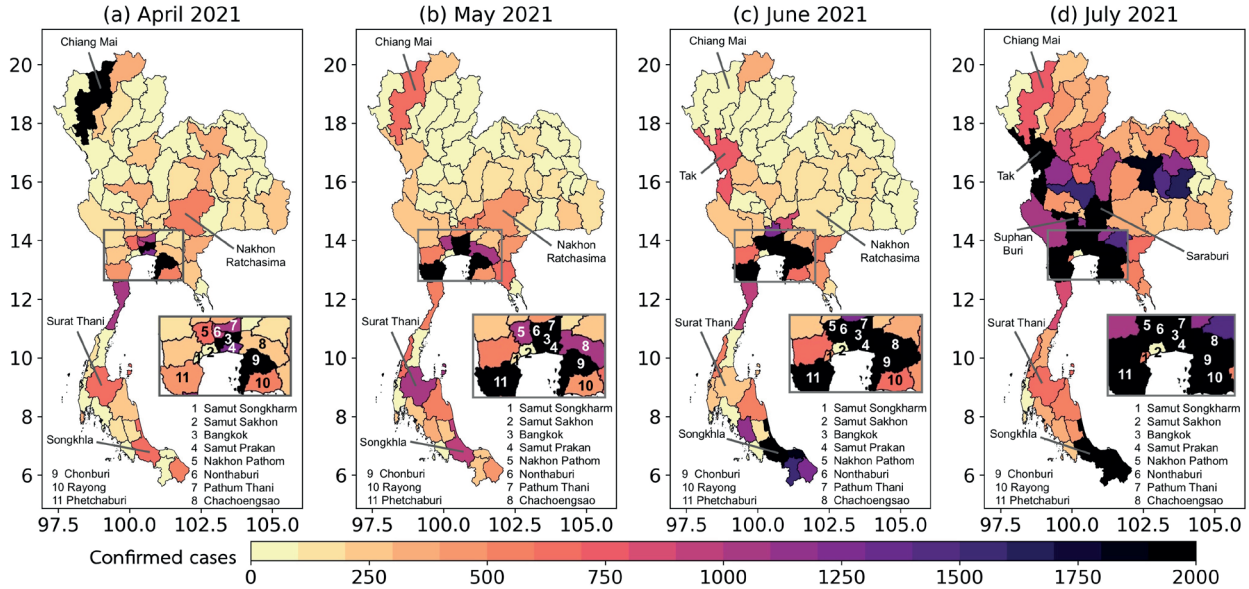

Fig. 3. Spatial distribution of monthly accumulative COVID-19 confirmed cases in Thailand by province during the period of April 1, 2021 to July 23, 2021

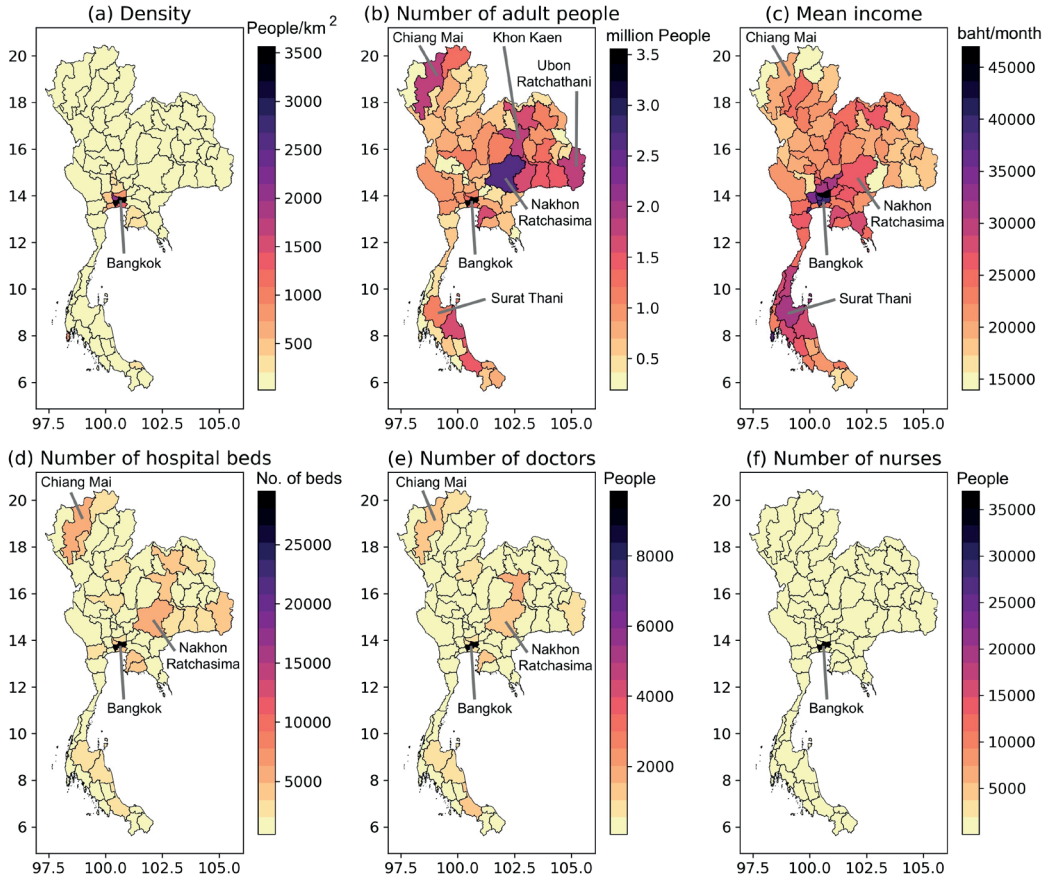

Fig. 4. Spatial distribution of (a) population density (people/km2), (b) number of adult people (million), (c) mean household income (baht/month), (d) number of hospital beds, (e) doctors (people) and (f) nurses (people) in 2020

Table 1. Description of parameters, data sources, and periods used in this study

Type	Description of variable	Data source	Period
Pandemic	COVID-19 confirmed cases	https://data.go.th/dataset/covid-19-daily	January-July 2021
Demographic	Total population of province	http://statbbi.nso.go.th/staticreport/page/sector/th/index.aspx	2020
	Population density		
	Number of adults in province		
Healthcare	Number of Hospital beds in province	http://statbbi.nso.go.th/staticreport/page/sector/th/index.aspx	2020
	Number of Doctors in province		
	Number of Nurses in province		
Socioeconomic	Mean monthly household income in province	http://statbbi.nso.go.th/staticreport/page/sector/th/index.aspx	2020

$$I = \frac{N}{\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}} \times \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} z_i w_{ij} z_j}{\sum_{i=1}^{N} z_i^2} \quad (2)$$

where N is the total number of analysis locations, zi is the variable value for each location i, and wij is the spatial weight (or connectivity) for location i and j. Notice that locational information for this formula is found in the weights. For non-neighboring tracts, the weight is zero, so these add nothing to the correlation. Global Moran's I ranges from -1 to +1, where +1 indicates strong positive spatial autocorrelation (clustered pattern), 0 indicates no spatial autocorrelation (random pattern), and -1 indicates strong negative spatial autocorrelation (checkered pattern).

The local Moran's I, or local indicators of spatial association (LISA), (Anselin 1995) was used to identify the location of local clusters for each individual location in the region as either a hot or cold spot, which is the likelihood of a unit's attribute value being associated with values in neighboring regions. The LISA spreading maps (LISA map) and LISA significance maps were then created with the following equation:

$$I_i = z_i \sum_j w_{ij} z_j \quad (3)$$

where zi and zj are the deviation of the variable with respect to the mean values for location i and j, and j is a location in the neighboring region of location i.

The LISA significance map and cluster map are used to visualize clusters of COVID-19 incidence more thoroughly. The maps highlight significant provinces with the LISA statistics using four alternative methods of clustering. Each one is based on a province's local spatial relationship with its neighbors (Anselin and Bao 1997; Anselin 2019): High-High (indicates a province with a high value surrounded by provinces of high values), Low-High (a province with a low value surrounded

by provinces of high values), Low-Low (a province with a low value surrounded by provinces of low values), and High-Low (a province with a high value surrounded by provinces of low values). The spatial dependence of the provinces are represented on a map and color-coded based on the type of interaction. Positive local spatial correlations are detected as spatial clusters: High-High and Low-Low areas (red and light green color, respectively), whereas negative local spatial correlations are defined as spatial outliers in the High-Low and Low-High sectors (orange and dark green color, respectively).

Moreover, bivariate Moran's I statistic is a spatial version of the correlation coefficient. It captures the changing relationship between two variables at multiple locations and embeds this information in a geographical context (Lee 2001). We also calculated the bivariate spatial autocorrelation between province-level demographics, healthcare, socioeconomic parameters and COVID-19 incidence rates. The spatial correlations of COVID-19 incidence based on the two variables are as follows (Anselin 2019):

$$I_{x,y} = \frac{N}{\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}} \times \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}\left(x_i - \bar{x}\right)\left(y_j - \bar{y}\right)}{\sqrt{\sum_{i}^{N}\left(x_i - \bar{x}\right)^2 \sum_{j}^{N}\left(y_j - \bar{y}\right)^2}} \quad (4)$$

Six distinct types of neighborhood were selected, as in a prior study (Kang et al. 2020; Meng et al. 2005). Three groups of parameters (Table 1) that may describe the spatial relationship with incidence rates of the COVID-19 pandemic were examined using bivariate Moran's I statistic: demographic (population density and age), socioeconomic (mean income) and healthcare (hospital beds, number of doctors and nurses). To employ the bivariate Moran's I method, the incidence rates were analyzed within those parameters, providing six models in this stage. Because COVID-19 travels from person to person, the coronavirus spreads from places with the largest to the smallest population size (Kang et al. 2020; Meng et al. 2005). Model 1 took into account population density. As seen in Fig.2, the key inflection age group is adults, thus, numbers of adult people were important factors to consider for Model 2. In Model 3, we considered the mean household income in each province by assuming that COVID-19 spread refers to socioeconomic factors, from the largest to the smallest provinces (Mollalo et al. 2020). In terms of healthcare resources, Models 4, 5 and 6 took into account the number of hospital beds, as well as doctors and nurses, respectively. Its assumes that the spread of COVID-19 goes from where the largest to the smallest healthcare resources exist (Kang et al. 2020; Mollalo et al. 2020; Meng et al. 2005). Spatial autocorrelation analyses in this study were implemented using a PySAL package, Version 1.14.0 (Rey and Anselin 2007), based on a Python Version 3.8 package.

RESULTS

Spatial Distribution of monthly COVID-19 incidence

Monthly COVID-19 incidence rates in the third wave are shown in Figure 5(a-d) for April, May, June and July, respectively. The incidence rates in Thailand clearly increased from month to month. The infection rate first increased in Bangkok, Chonburi and Prachuap Khiri Khan Provinces, then came first-order neighboring provinces like Nonthaburi, Pathum Thani, Samutprakarn, and Phetchaburi (Figure 5b), and eventually second-order neighboring provinces such as Nakhon Pathom, Chachoengsao and Samut Songkhram (Figure 5c). Finally, the epidemic spread to neighboring regions in the third-order, including Nakhon Nayok, Saraburi, Ayutthaya, Suphan Buri, Ang Thong, Sing Buri, Lopburi, Tak, Ratchaburi, Ranong, Songkhla, Pattani, Yala and Narathiwat (Figure 5d). This confirms that COVID-19 is regionally dispersed and that investigating spatial dependency is necessary.

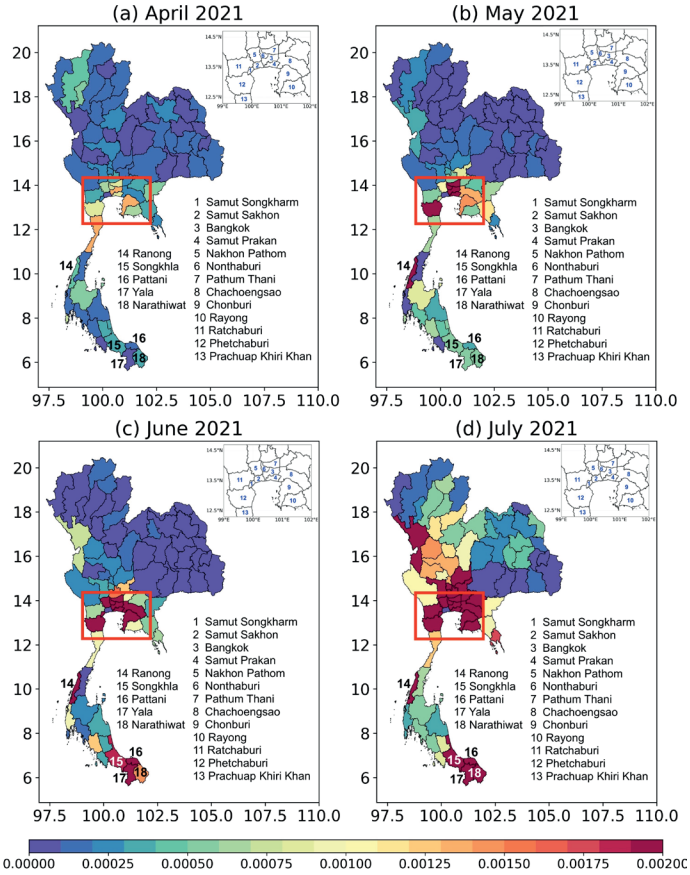

Fig. 5. Spatial distribution map of Thai COVID-19 incidence rates at the provincial level during the period of (a) April 2021, (b) May 2021, June 2021 and July 2021

Univariate spatial correlation

The goal of this research was to examine if COVID-19 incidences in Thailand had a spatial correlation. To measure spatial dependency, Moran's I statistic was utilized to determine COVID-19 incidence in various types of neighborhoods. The global Moran's I value for COVID-19 incidence was 0.3290, 0.3739, 0.5789 and 0.7605 for April, May, June and July, respectively ($P < 0.05$), which indicates that COVID-19 incidence at the province-level showed a very significant spatial dependence. Additionally, Figure 6 presents the LISA map (upper panel) and its p-value (lower panel) for each month during the third wave of the pandemic at Thailand's province level. The LISA map highlights the main clusters of COVID-19 incidence, where the High-High clustering zones (red) indicate regions with high numbers of incidence relative to the average, surrounded by regions that also present high numbers. The Low-Low clustering zones (light green) denote spatial association groups with low numbers of incidence (below the average) surrounded by low-valued regions. In April, most provinces in Central Thailand (including Bangkok) and surrounding provinces belonged to a High-High clustering zone. In contrast, Samut Sakorn and Samut Songkarm Provinces were close to these High-High clustering zones but contained significant Low-Low clustering zones. Since those provinces were the main COVID-19 clusters in the second wave, most of its citizens received a COVID-19 vaccine afterward. Therefore, these regions have had low incidence rates in the third wave. In May, the High-High clustering zone from East of Bangkok decreased, and a Low-Low clustering zone appeared in North and Northeast Thailand, indicating that COVID-19 had spread across the country (Fig.9a). It became apparent during May to July (Fig 9b-d) that a Low-Low clustering zone is continuously increasing its distribution in the North and Northeast regions, and a High-High clustering zone has also expanded.

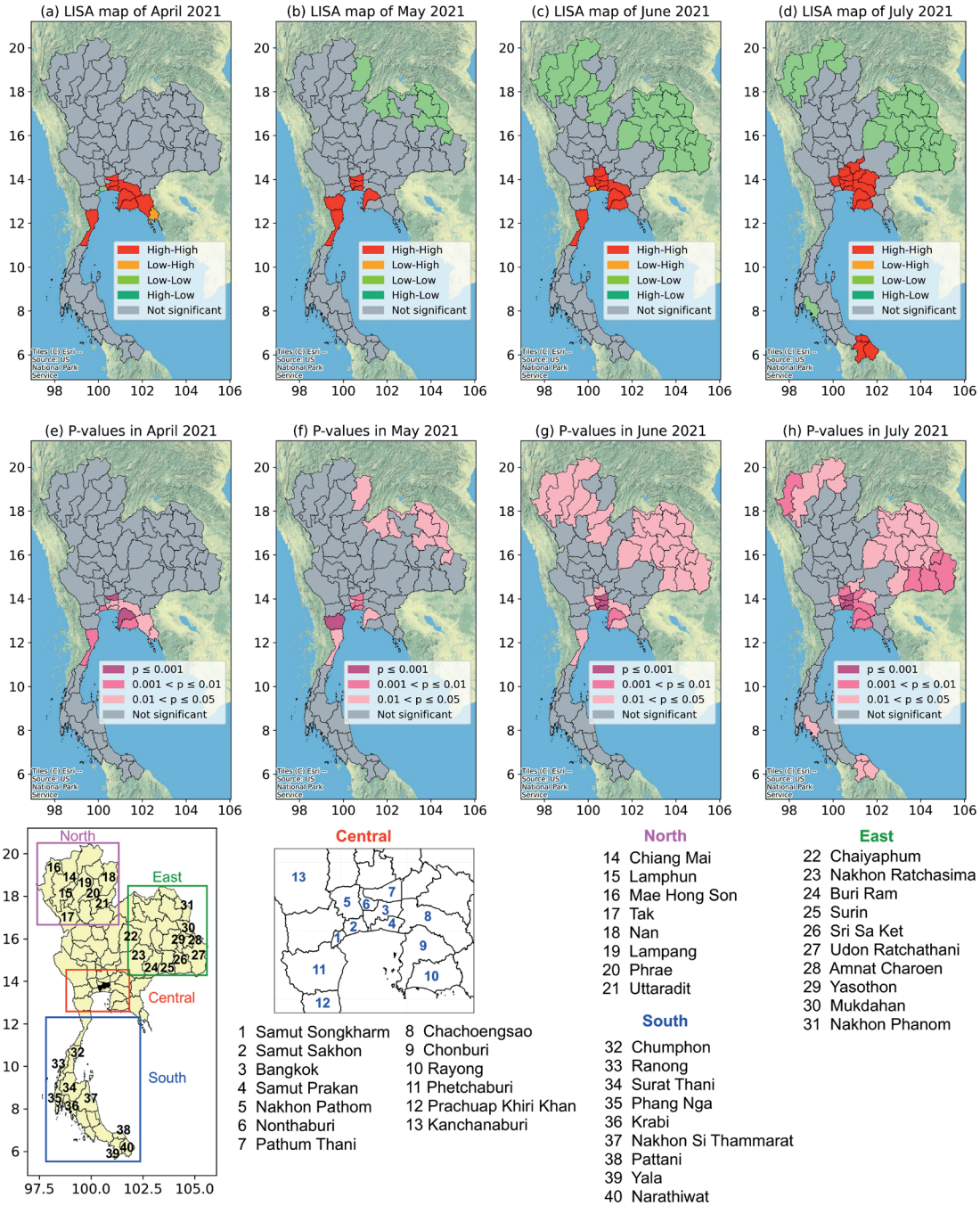

Fig. 6. Global Moran's I and univariate LISA cluster map of monthly COVID-19 incidence in provinces of Thailand, (a) April 2021, (b) May 2021, (c) June 2021, and (d) July 2021. Significance of the LISA map for monthly COVID-19 incidence in (e) April 2021, (f) May 2021, (g) June 2021, and (h) July 2021

Bivariate spatial correlation

Figure 7 shows the summary of results obtained from Pearson correlation analysis. Factors of population density and mean income showed the highest significant positive correlation (0.55) with COVID-19 incidence rates, followed by healthcare resources, which also represented a significant positive correlation (0.33). Meanwhile, the adult age group showed a weak positive (0.2) association with COVID-19 incidence rates. Additionally, population density showed a strong significant positive correlation (0.8) with the number of hospital beds, as well as doctor and nurse numbers. Accordingly, the distribution of medical resources in Thailand leans towards regions with greater population size and economic development.

Bivariate Moran's I analysis of demographics, health care and socioeconomics with COVID-19 incidence revealed a positive spatial autocorrelation, with Moran's I values shown in Table 2. With a significance threshold of 0.05, these values were statistically significant, except for Model 2: number of adult people in April and May. Based on these results, all the Models exhibit strong spatial association for COVID-19 incidence, and Moran's I value increases as the number of COVID-19 incidences increase. Figure 8 shows changes in the global Moran's I and P-values over time. The incidence rates of confirmed COVID-19 cases in all models showed significant global spatial correlation (Moran's $I > 0.0$, $p < 0.05$), except for the first few days in April 2021 and in early May 2021 of Model 2 (Figure 8b). Features of Moran's I index trend changes are evident in three stages: they first increase in a low slope from April to May 2021; then, an increasing slope appears from May to June; after that, the trend becomes steadier in July (Fig. 8). These results are inconsistent with a change in the incidence trend that rapidly increased from May to July. This indicates that, even though the incidence rates were increasing, the number of clusters did not change or were slightly less than previously (Fig. 5 and 6). Cluster features still influence global spatial correlation, however, reflecting that COVID-19 incidence tends to develop into a spread. Global bivariate Moran's I revealed that Model 1 had the highest spatial dependence between COVID-19 cases and population density in July 2021 (Moran's I = 0.6022, p = 0.001), as shown in Table 2 and Fig. 8a. The primary High-High clustering zones were located in big cities of Thailand such as Bangkok, Rayong, Samut Prakan, Patumthani and Nonthaburi (Fig. 9a-d). The Low-Low clustering zones were mainly located in Northern Thailand. This implies that high COVID-19 incidence rates occurred in regions with high population density such as big cities, then distributed to rural areas with low population density. Because COVID-19 spreads primarily from people to people by touching, breathing and talking, high population density should be considered to have a high possibility of infection. Model 2 indicates that the number of adult people determines the highest proportion among COVID-19 infections in people. Global bivariate Moran's I dramatically increased each month with weak positive spatial autocorrelation (Moran's I = 0.1023-0.1985), as shown in Table 2 and Fig. 8b. The LISA map for April and June presented three groups of clustering zones: High-High, Low-Low and Low-High. A High-High clustering zone was located in Bangkok, Samut Prakan and Chonburi, similar to Model 1, indicating a high incidence rate with high numbers of adult people in the surrounding area. A Low-Low clustering zone was found in Suphanburi, Nakhonsawan and Chiang Mai, and a Low-High clustering zone appeared in Samut Sakorn, Chaiyaphoom and Buriram, which indicates low incidence rates with a low and high number of adult people in the surrounding area, respectively. Then in July, the Low-Low clustering zone in Chiang Mai disappeared, while the Low-Low clustering zone in Suphanburi became a High-Low clustering zone. These results suggest that Chiang Mai successfully reduced the incidence rate, whereas Suphanburi increased its incidence rate in other age groups. For the mean household income variable, Model 3 had strong positive spatial autocorrelation (Moran's I = 0.2980-0.3523), as shown in Table 2 and Fig. 8c. Most of the LISA map reveals two primary clustering zones - one High-High and one Low-Low (Fig. 9 i-l). It can be observed how dynamically changes between the High-High and Low-High clustering zones occurred, depending on how the incidence rates changed. The results indicate that a region with high income tends to have a high possibility of infection due to being densely populated. Model 4 (hospital beds), Model 5 (doctor numbers) and Model 6 (nurse numbers) indicated moderate positive spatial autocorrelation (Moran's I = 0.1986-0.4853), as shown in Table 2 and Fig. 8d-f, with a uniform pattern of High-High and Low-Low clustering zones (Fig.10). The results show that hospital beds were not more essential than doctors or nurses in COVID-19's third wave, and they were significant throughout the whole period.

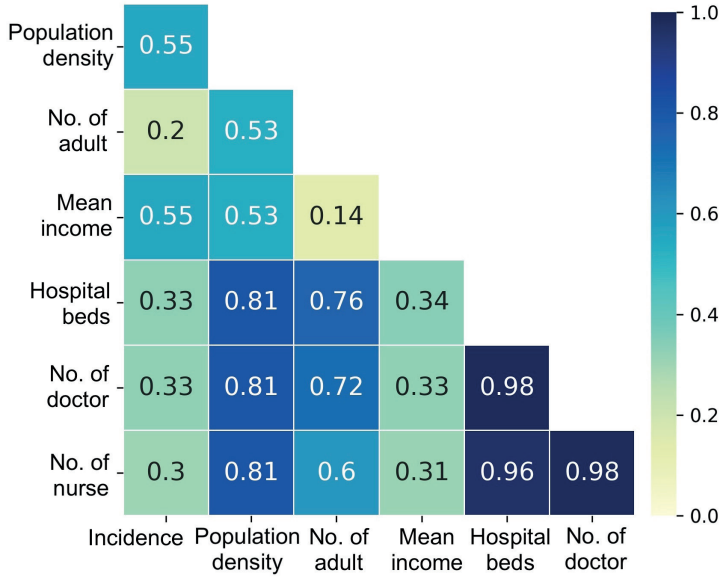

Fig. 7. Summary of Pearson correlation analysis between COVID-19 incidence rates and factors related to demographic, socio-economic, and healthcare resources.

Table 2. Moran's I statistical and P-values for Models 1 to 6

	April		May		June		July	
	Moran's I	P-value	Moran's I	P-value	Moran's I	P-value	Moran's I	P-value
Model 1	0.2776	0.0020	0.3788	0.0010	0.5765	0.0010	0.6022	0.0010
Model 2	0.1023	0.0570	0.1081	0.0840	0.1985	0.0340	0.1885	0.0400
Model 3	0.2980	0.0010	0.3908	0.0010	0.4663	0.0010	0.4853	0.0010
Model 4	0.1986	0.0210	0.2237	0.0080	0.3424	0.0020	0.3457	0.0040
Model 5	0.1881	0.0140	0.2223	0.0030	0.3440	0.0010	0.3523	0.0050
Model 6	0.1718	0.0210	0.2144	0.0050	0.3295	0.0020	0.3340	0.0070

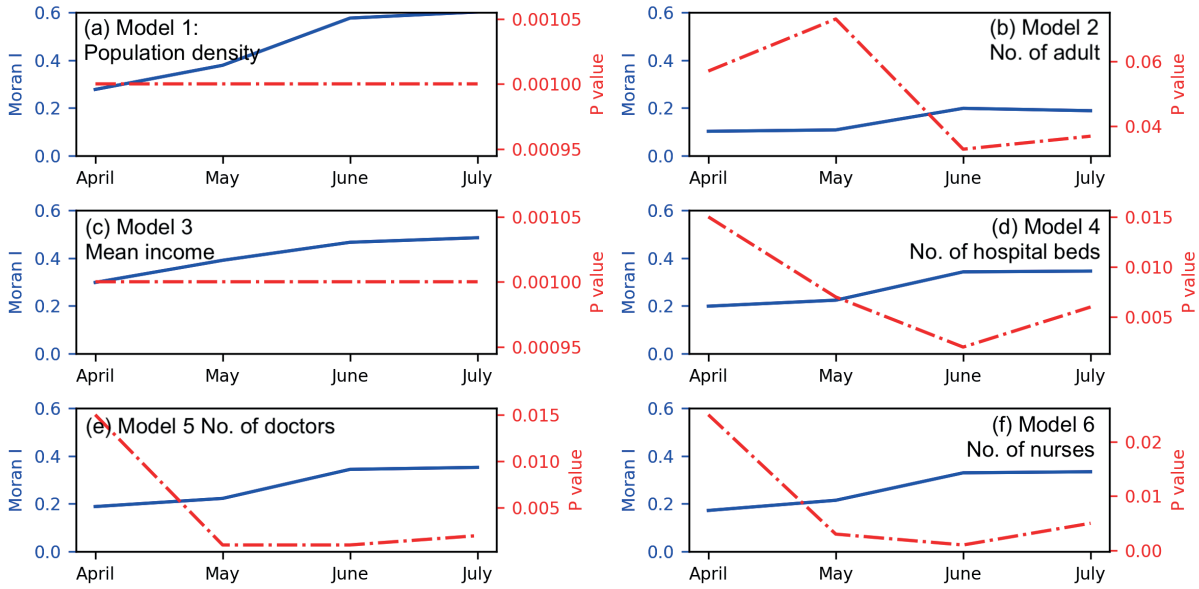

Fig. 8. Monthly changes in local Moran's I statistics and p-values from April 1, 2021 to July 23, 2021 in Thailand for six models: (a) Model 1: population density, (b) Model 2: number of adults, (c) Model 3: mean income, (d) Model 4: number of hospital beds, (e) Model 5: number of doctors, and (f) Model 6: number of nurses

For each day, Moran's I statistics and associated p-values are shown in Figure 10. Except for the first few days and Model 2, the p-values in Figure 11 are all very close to zero. All the models show systematically spatial clustering correlation with the COVID-19 incidence rates since 10 April, being especially strongest in Model 1 (Fig.10a). In Figure 11, the features of Moran's I index trend changes are seen in three phases: they first increase and then fall from April 10; after that, they increase again and then fall from May 20; after that, they increase again from June 7. Interestingly, after they went down, they then increased with higher Moran's I values, indicating that the COVID-19 spread was continuously increasing. Moreover, the different Moran's I value in each stage demonstrated the transmission level of infectious COVID-19, particularly in the Early (first phase) and Spreading stages (second and third phases).

In addition, daily changes in the bivariate global Moran's I value expressed the key factors affecting the spread of the COVID-19 epidemic. The results reveal that COVID-19 incidence has a significant spatial dependence on all six variables, indicating that effecting factors on the outbreak and transmission of an epidemic occur primarily through the source of infection, the channel of transmission, and the receptive population. As seen in Model 1, the results point to the fact that COVID-19 expanded significantly, from crowded areas to adjacent areas. During the periods with no significant spatial dependence in Model 2, we demonstrated that COVID-19 was spreading to other age groups. The third wave of the COVID-19 pandemic in Thailand started with a spike in daily confirmed cases on 10 April associated with the Songkran holiday. Workers went back to their hometowns and spread COVID-19 to their family members, resulting in a wider spread and transmitting it to other age groups. It has been demonstrated that the factors that affect the outbreak and transmission of an epidemic occur mainly through influences on the sources of infection, the routes of transmission, and the susceptible population. COVID-19's spread may be linked to ecological, social, and economic factors, as well as preventative measures and controlling activities such as limiting the infection sources and shutting off transmission routes.

DISCUSSION

This study provides information on the spatial and temporal patterns of the third wave of the COVID-19 pandemic in Thailand. In the early stage of the COVID-19 outbreak, new cases occurred intensively in Bangkok and Rayong, Chonburi and Prachuap Khiri Khan Provinces. The pandemic extended to bordering provinces over time, with the first-order, second-order and third-order neighboring provinces showing a notably high number of confirmed cases after May 2021. Results on the spatial distribution patterns of COVID-19 incidence rates in Thailand show a

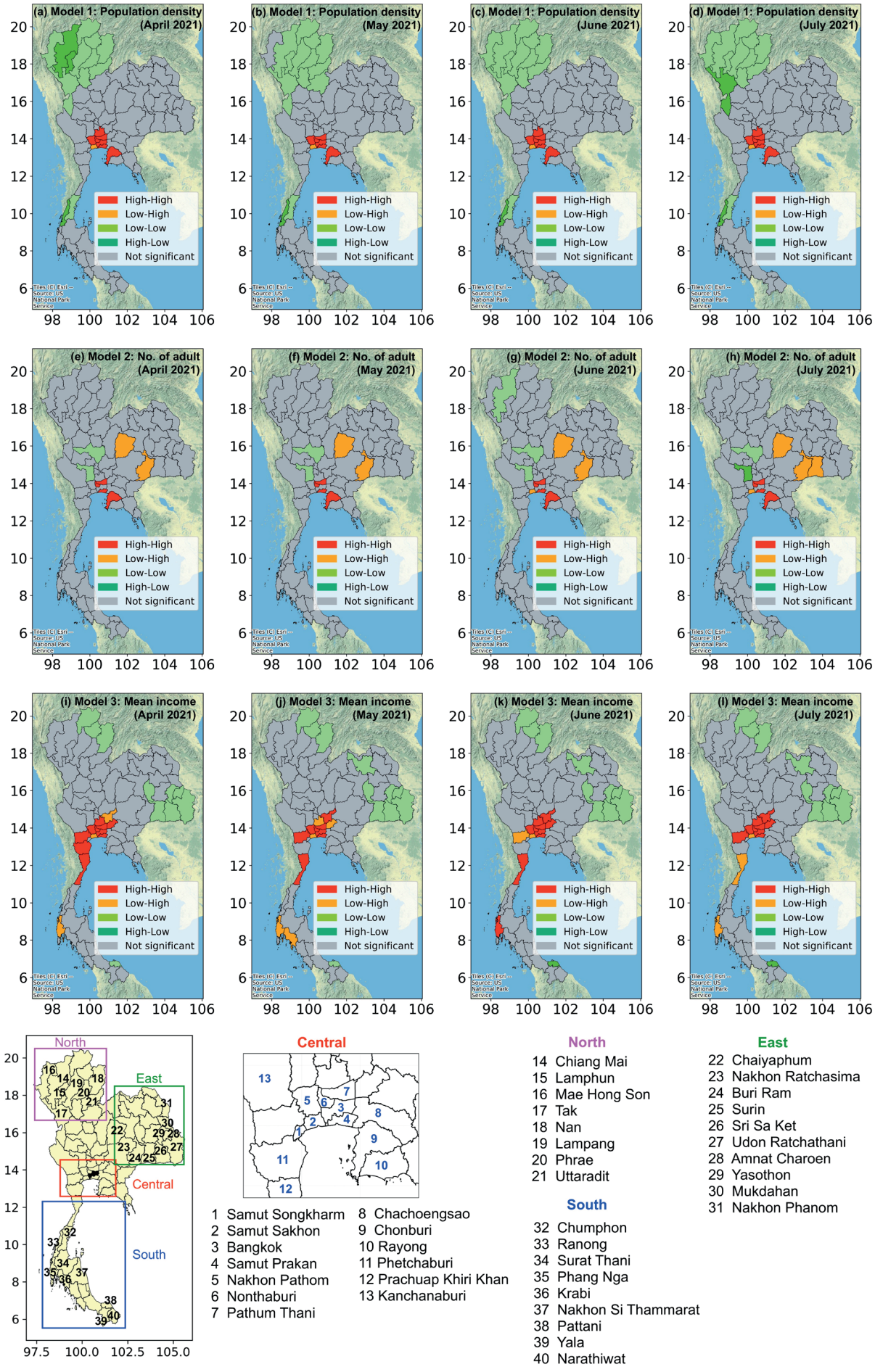

Fig. 9. Spatial distribution of LISA cluster maps between COVID-19 incidence and related parameters: Model 1: population density (a-d), Model 2: number of adults (e-h), and Model 3: mean household income (i-l) from April to July 2021

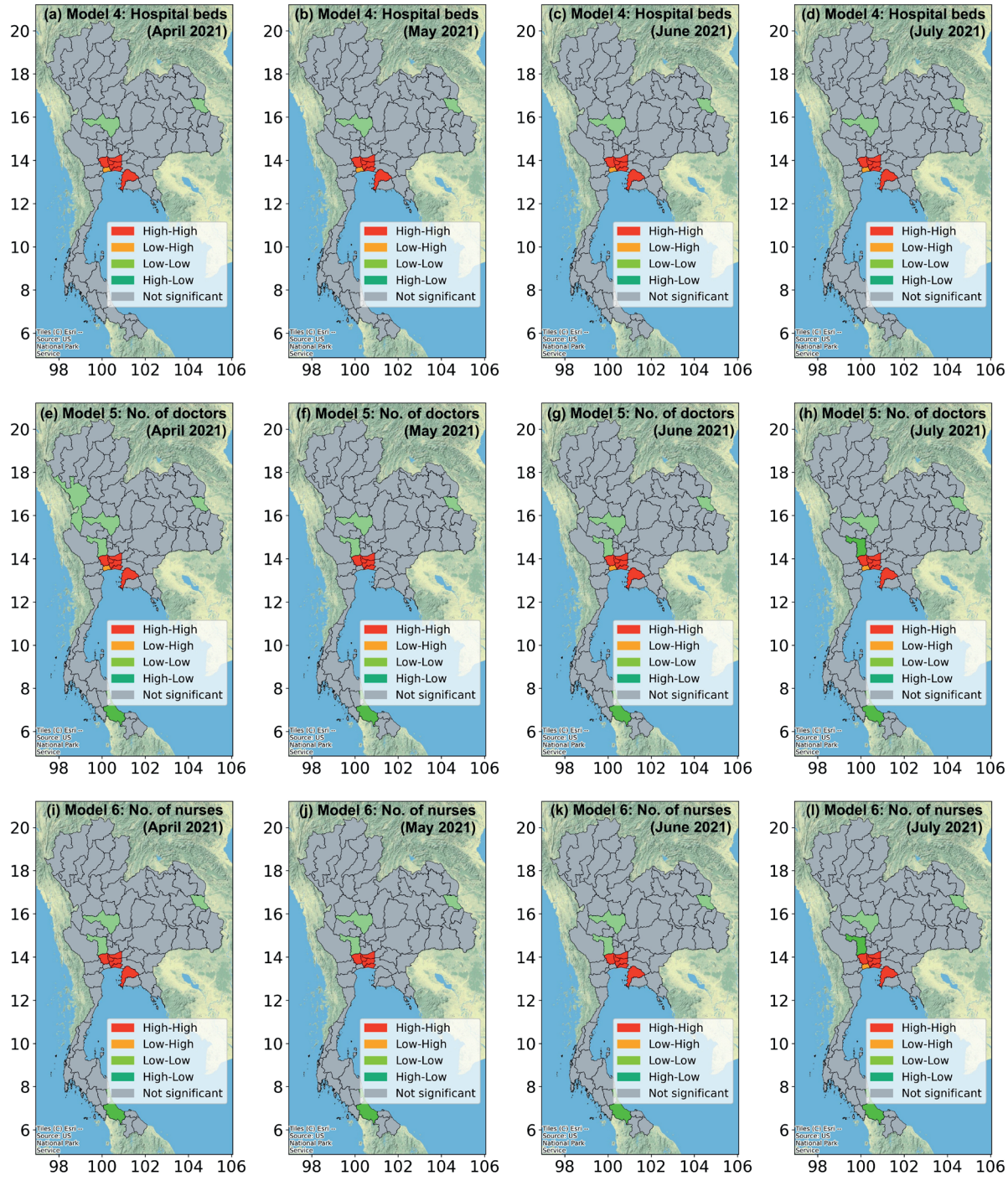

Fig. 10. Spatial distribution of LISA cluster maps between COVID-19 incidence and related parameters: Model 4: hospital beds (a-d), Model 5: number of doctors (e-h), and Model 6: number of nurses (i-l) from April to July 2021

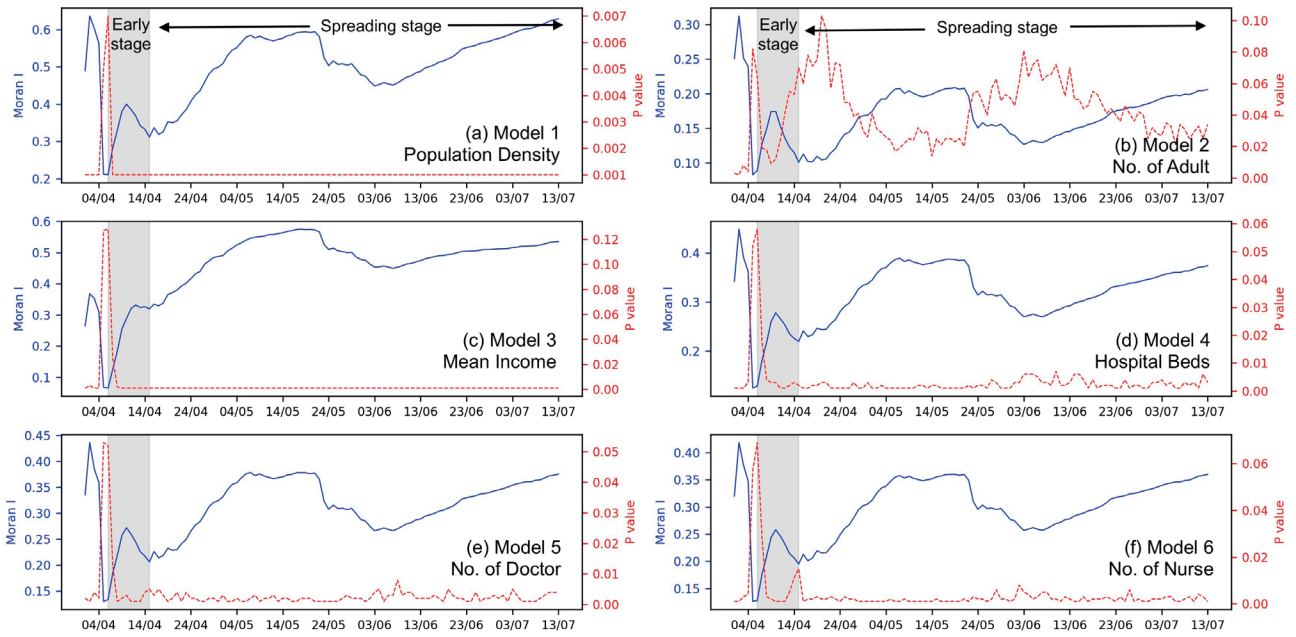

Fig. 11. Daily changes in the global Moran's I statistic and p-values from April 1, 2021 to July 23, 2021 in Thailand for six models: (a) Model 1: population density, (b) Model 2: number of adults, (c) Model 3: mean income, (d) Model 4: number of hospital beds, (e) Model 5: number of doctors, and (f) Model: 6 number of nurses

clear systematic pattern among provinces: high incidence rates surrounded by provinces with similar levels of incidence, and provinces with low incidence surrounded by provinces with low values. The highest infected clusters have occurred in Bangkok and its neighboring provinces, while lower infection rates have been in the countryside. The same pattern was observed for the relationship between incidence rates and population density, adult numbers, mean household income, hospital beds, and doctor and nurse numbers.

Meng et al. (2005) modeled the influence of demographic, socioeconomic and healthcare resources on the (SARS) respiratory syndrome in Beijing and found that population density and medical care resources did influence the spatial spread of SARS in Beijing. However, the major variables that caused the epidemic to spread were different at different times. Kang et al. (2020) provides information on the spatial and temporal patterns of the COVID-19 pandemic in mainland China. They show that, in the early phases of the COVID-19 pandemic, the disease distributed rapidly from region to region in mainland China. Moreover, they also evaluated the influence of demographic, socioeconomic and healthcare resource factors on COVID-19 cases. The results presented significant spatial autocorrelation between COVID-19 cases and population density and the number of doctors, but no significant spatial autocorrelation existed with hospital beds. Population size is a key factor in COVID-19 spread.

Their findings are in line with our results, except that we found a significant spatial autocorrelation between COVID-19 incidence and hospital beds, as it showed strong correlation between population density and healthcare resources. Having a large number of beds, doctors and nurses in a location means that the area can handle a large number of critically ill patients, this could lead to the infection spreading. Moreover, the daily change in global Moran's I values in this study could indicate the Early and Spreading stages during the COVID-19 pandemic's third wave. Similarly, Triukose et al. (2021) used effective reproduction numbers to identify the COVID-19 transmission level in Bangkok. On the other hand, the LISA statistic showed the dynamics of cluster distribution to be consistent with the Spreading stages, which were found by the daily change in global Moran's I values. Therefore, global and local Moran's I can be used as a monitoring and policy-decision-making tool to inform disease control activities in the future. (Laohasiriwong et al. 2018; Triukose et al. 2021). There are several limitations to this paper. The factors that influence the epidemic's propagation are numerous. This paper developed an indicator system based on the available data that influences the epidemic's multiple factors (Meng et al. 2005; Kang et al. 2020). Other non-quantitative factors may have been overlooked, adding to the inadequacy of the study's evaluation. Without access to precise information on those diagnosed with COVID-19, such as medical history, weight, smoking status, and so on, the study's findings may have been weakened. Nevertheless, attempts to investigate probable external environmental impacts are critical for protecting healthcare workers and containing the COVID-19 outbreak (Wang et al. 2021). Future research may be useful in delving deeper into the epidemiological factors and social environment of COVID-19.

CONCLUSIONS

A new geographic database for studying the COVID-19 epidemic was created based on the COVID-19 outbreak in 76 provinces and one special administrative region (Bangkok) in Thailand. A PySAL package based on python language was used to detect early transmission by using the spatial distribution of COVID-19 incidence in Thailand's provinces and its relationship with sociodemographic factors in order to better understand the outbreak and transmission of the disease in Thailand from April to July 2021. Thus, the following conclusions were found. Global spatial autocorrelation was employed to confirm that COVID-19 incidence in Thailand has a spatial association, and the correlation characteristics increased at a slow rate in April, and then rapidly increased in June and July 2021. However, considering local spatial autocorrelation, the correlation characteristics tended to remain steady over time and mostly consisted of High/Low aggregation zones. In the provinces surrounding Bangkok, hotspots

have stabilized over time (Nonthaburi, Pathum Thani, Samutprakarn Chonburi, and Prachuap Khiri Khan Provinces). Among social factors, population density exhibited the highest positive correlation with incidence. Our results revealed that the diverse spatiotemporal clustering patterns we discovered may represent variances in social and demographic characteristics, public health emergency readiness and response capacities, as well as differences in transmission patterns and mechanisms of these coronaviruses. Finally, daily changes in global Moran's I successfully indicated the Early and Spreading stages during the third wave of the COVID-19 pandemic. These findings could be used to create monitoring tools and aid in policy prevention planning. Future researches may be useful in delving deeper into the epidemiological factors and social environment of the COVID-19.

REFERENCES

Anselin L. (1995). Local Indicators of Spatial Association LISA, Geographical Analysis, 27, DOI: 10.1111/j.1538-4632.1995.tb00338.x.

Anselin L. (2004). GeoDa 0.95i Release Notes. Urbana-Champaign, IL: Spatial Analysis Laboratory (SAL), Department of Agricultural and Consumer Economics, University of Illinois.

Anselin L. (2019). A local indicator of multivariate spatial association: extending Geary's C. Geogr Anal, 51, 133-50, DOI: 10.1111/gean.12164.

Anselin L., and Bao S. (1997). Exploratory Spatial Data Analysis Linking Space Stat and ArcView, Chapter 3 in Recent Developments in Spatial Analysis, Spatial Statistics, Behavioral Modeling and Computational Intelligence, M.M. Fischer and A. Getis (Eds.). Berlin: Springer-Verlag, 35-59.

DDC (2021). The Digital Government Development Agency, https://data.go.th/dataset/covid-19-daily, (Accessed 23 July 2021).

Fang L.Q., Vlas S.J., Feng D., Liang S., Xu Y.F., and Zhou J.P. (2009). Geographical spread of SARS in mainland China, Trop Med Int Health, 14, DOI: 10.1111/j.1365-3156.2008.02189.x.

Green M.S., Nitzan D., Schwartz N., Niv Y., Peer V. (2021). Sex differences in the case-fatality rates for COVID-19 — A comparison of the age-related differences and consistency over seven countries, PLoS ONE, 16, 4, DOI: 10.1371/journal.pone.0250523.

Kang D., Choi H., Kim J.H., and Choi J. (2020). Spatial epidemic dynamics of the COVID-19 outbreak in China, Int J Infect Dis, 94, DOI: 10.1016/j.ijid.2020.03.076.

Laohasiriwong W., Puttanapong N., and Luenam A. (2018). A comparison of spatial heterogeneity with local cluster detection methods for chronic respiratory diseases in Thailand, F1000Research, 6,1819, DOI: 10.12688/f1000research.12128.2.

Lee S.I. (2001). Developing a bivariate spatial association measure: An integration of Pearson's r and Moran's I, J. Geograph Syst, 3, DOI: 10.1007/s101090100064.

Li B., Deng A., Li K., Hu Y., Li Z., Xiong Q., Liu ,Z., Guo Q., Zou L., Zhang H., Zhang M., Ouyang F., Su J., Su W., Xu J., Lin H., Sun J., Peng J., Jiang H., Zhou H., Zhen H., Xiao J., Liu T., Che R., Zeng H., Zheng Z., Huang Y., Yu J., Yi L., Wu J., Chen J., Zhong H., Deng X., Kang M., Pybus O., Hall M., Lythgoe K., Li Y., Yuan J., He J., and Lu J. (2021). Viral infection and transmission in a large well-traced outbreak caused by the Delta SARS-CoV-2 variant, medRxiv, Published online: July 12, 2021:2021.07.07.21260122, DOI:10.1101/2021.07.07.21260122.

Li H., Calder C.A., and Cressie N. (2007). Beyond Moran's I: testing for spatial dependence based on the spatial autoregressive model, Geogr Anal, 39, 4, DOI: 10.1111/j.1538-4632.2007.00708.x.

Li H., Li H., Ding Z., Hu Z., Chen F., Wang K., Peng Z., and Shen H. (2020). Spatial statistical analysis of Coronavirus Disease 2019 (Covid-19) in China, Geospatial Health, 15, 1, DOI: 10.4081/gh.2020.867.

Meng B., Wang J., Liu J., Wu J., Zhong E. (2005). Understanding the spatial diffusion process of severe acute respiratory syndrome in Beijing. Public Health, 119, 12, DOI: 10.1016/j.puhe.2005.02.003.

Mollalo A., Vahedi B., and Rivera K.M. (2020). GIS-based spatial modeling of COVID-19 incidence rate in the continental United States, Science of the Total Environment, 728, DOI: 10.1016/j.scitotenv.2020.138884.

Moran P.A. (1950). A test for the serial independence of residuals. Biometrika, 37, DOI: 10.1093/biomet/37.1-2.178.

NSO (2021). The National Statistical Office of Thailand, http://statbbi.nso.go.th/staticreport-/page/sector/th/index.aspx, (Accessed 23 July 2021).

Pei Z., Han G., Ma X., Su H., and Gong W. (2020). Response of major air pollutants to COVID-19 lockdowns in China, Sci Total Environ, 743, DOI: 10.1016/j.scitotenv.2020.140879.

Rajatanavin N., Tuangratananon T., Suphanchaimat R., and Tangcharoensathien V. (2021). Responding to the COVID-19 second wave in Thailand by diversifying and adapting lessons from the first wave BMJ Global Health, 6, 7, DOI: 10.1136/bmjgh-2021-006178.

Ramírez-Aldana R., Gomez-Verjan J.C., and Bello-Chavolla O.Y., (2020). Spatial analysis of COVID-19 spread in Iran: Insights into geographical and structural transmission determinants at a province level, PLoS Negl Trop Dis, 18, 14(11), DOI: 10.1371/journal.pntd.0008875. PMID: 33206644; PMCID: PMC7710062.

Raymundo C.E., Oliveira M.C., Eleuterio T. A., André S. R., da Silva M. G., Queiroz ERDS, and Medronho R.A. (2021). Spatial analysis of COVID-19 incidence and the sociodemographic context in Brazil, PLoS One, 6, 3, DOI: 10.1371/journal.pone.0247794.

Rey S.J., and Anselin L. (2010). PySAL: A Python Library of Spatial Analytical Methods. In: Fischer M., Getis A. (eds) Handbook of Applied Spatial Analysis. Springer, Berlin, Heidelberg, DOI: 10.1007/978-3-642-03647-7_11.

Rotejanaprasert C., Lawpoolsri S., Panngum W., and Maude R.J. (2020). Preliminary estimation of temporal and spatiotemporal dynamic measures of COVID-19 transmission in Thailand, PLoS ONE, 15, 9, DOI: 10.1371/journal. pone.0239645.

Sannigrahi S., Pilla F., Basu B., Basu A.S., and Molter A. (2020). Examining the association between socio-demographic composition and COVID-19 fatalities in the European region using spatial regression approach, Sustain Cities Soc, 62, DOI: 10.1016/j.scs.2020.102418.

Triukose S., Nitinawarat S., Satian P., Somboonsavatdee A., Chotikarn P., Thammasanya T., Wanlapakorn N., Sudhinaraset N., Boonyamalik P., Kakhong B., and Poovorawan Y. (2021). Effects of public health interventions on the epidemiological spread during the first wave of the COVID-19 outbreak in Thailand, PLoS ONE, 16, 2, DOI: 10.1371/journal. pone.0246274.

Tu J., and Xia Z.G. (2008). Examining spatially varying relationships between land use and water quality using geographically weighted regression I: Model design and evaluation, Sci Total Environ., 407, 358-378, DOI: 10.1016/j.scitotenv.2008.09.031.

Tu, J. and Xia Z.G., (2008). Examining spatially varying relationships between land use and water quality using geographically weighted regression I: Model design and evaluation, Sci Total Environ., 407, 358-378, DOI: 10.1016/j.scitotenv.2008.09.031.

Wang Q., Dong W., Yang K., Ren Z., Huang D., Zhang P., and Wang J. (2021). Temporal and spatial analysis of COVID-19 transmission in China and its influencing factors, Int J Infect Dis., 105, DOI: 10.1016/j.ijid.2021.03.014.

Wetchayont P. (2021). Investigation on the Impacts of COVID-19 Lockdown and Influencing Factors on Air Quality in Greater Bangkok, Thailand, Advances in Meteorology, 2021, 6697707, DOI: 10.1155/2021/6697707.

Wetchayont P., Hayasaka T., and Khatri P. (2021). Air Quality Improvement during COVID-19 Lockdown in Bangkok Metropolitan, Thailand: Effect of the Long-range Transport of Air Pollutants, Aerosol Air Qual. Res., 21, 200662, DOI: 10.4209/aaqr.200662.

Zhang X., Rao H., and Wu Y. (2020). Comparison of spatiotemporal characteristics of the COVID-19 and SARS outbreaks in mainland China, BMC Infect Dis, 20, 805, DOI: 10.1186/s12879-020-05537-y.