GIS APPLICATION FOR MODELING COVID-19 RISK IN THE MAKKAH REGION, SAUDI ARABIA, BASED ON POPULATION AND POPULATION DENSITY

Ibrahim Ali Alkhaldy

Department of Administrative and Human Research, The Custodian of the Two Holy Mosques Institute for Hajj and Umrah Research, Umm Al-Qura University. Makkah, Saudi Arabia
Email: iakhaldy@uqu.edu.sa

Abstract
COVID-19 in Saudi Arabia has emerged across different regions of the country. To date, the Makkah region has seen the highest number of COVID-19 cases, containing 39.5% of all cases in Saudi Arabia. To understand the risk of COVID-19, the population and population density of the region were investigated. Examining the number of COVID-19 cases could provide useful insights for officials when considering the adoption of different interventions. To obtain this data, a local spatial analysis using Geographically Weighted Regression (GWR) was employed to predict the risk of COVID-19 for the Makkah Region. A strong positive association (R² =0.96) was found between COVID-19 and population and population density in Makkah region. The cities with the highest population and population density were found to be at a higher risk of COVID-19.

Keywords: COVID-19, Explanatory variables, Geographically Weighted Regression, Makkah region.
1. Introduction

The Middle East has been plagued by various coronaviruses, such as severe acute respiratory syndrome coronavirus (SARS-CoV) and Middle East respiratory syndrome coronavirus (MERS-CoV) (Surveillances, 2020). Its most recent virus is the novel coronavirus named COVID-19, which first appeared in Wuhan, Hubei Province, China. It has since spread around the globe, reaching Saudi Arabia on March 2, 2020 (MOH, 2020; Surveillances, 2020). All cases that have been confirmed in Saudi Arabia are believed to have been imported from Iran (aside from one from Iraq) (Arab-Mazar, Sah, Rabaan, Dhama, & Rodriguez-Morales, 2020). All other cases appear to be the result of local transmission from interactions with those with initially confirmed cases (Arab-Mazar et al., 2020). Initially, the virus caused significant concern for health officials in late December 2019. An alert from the Wuhan Municipal Health Commission resulted in a response team being sent by the Chinese Center for Disease Control and Prevention, who in turn notified the World Health Organization (WHO) (HuiDS, Madani, Ntoumi, Koch, & Dar, 2020; WHO, 2020b; WMHC, 2020; Zhu et al., 2020). The WHO declared COVID-19 to be a global health emergency on January 30, 2020 (Velavan & Meyer, 2020). Epidemiological investigations have associated the virus with Wuhan’s Huanan Seafood Wholesale Market, which has since been closed down and disinfected (HuiDS et al., 2020; WHO, 2020b; Zhu et al., 2020).

The spread of COVID-19 was a major concern for the Chinese Government; as a result, the movement of people in and out of Wuhan was restricted (Surveillances, 2020). The Chinese Government developed and implemented control strategies in response to the epidemic, in order to prevent its spread (Surveillances, 2020). China carried the largest burden of this virus in February of 2020; however, many other countries surpassed China’s Infection rates shortly thereafter (Surveillances, 2020).

Historically, coronaviruses have been acknowledged since 1965 (Anderson et al., 2004). In 1966, Tyrell and Bynoe first described the coronaviruses, which were cultivated from patients with common colds (Tyrrell & Bynoe, 1966). At that time, only four subfamilies existed, namely alpha, beta, gamma, and delta (Velavan & Meyer, 2020). As of 2020, seven species have been shown to affect humans (Ahmadi, Fadai, Shirani, & Rahmani, 2020). Mammals and birds are considered to be a possible source of these diseases, also known as zoonotic diseases, which have been connected to SARS, MERS, and now COVID-2019, causing widespread epidemics.
COVID-19 can spread respiratory diseases directly and indirectly with a wide range of symptoms, including cold symptoms, respiratory/fever, shortness of breath, muscle pain, and even death (Liu et al., 2020). It is mostly believed to be transmitted through contact, droplets, aspirates, and feces, with an incubation period between 5 to 14 days. However, the exact route of transmission still requires identification (Li et al., 2020; WHO, 2020c).

Social distancing has been greatly encouraged by authorities globally to reduce the incidence of COVID-19 with the goal of limiting interactions between members of a community (Wilder-Smith & Freedman, 2020). It is also considered useful for regions where asymptomatic infectious individuals may not yet be identified (Wilder-Smith & Freedman, 2020). It has also been suggested that social distancing protects those who are at greater risk of fatal complications following infection (de Kadt J, 2020). If social distancing is insufficient, the next step may be implementing community containment, which has proven to be more challenging if it involves law enforcement (Wilder-Smith & Freedman, 2020).

Some Geographic Information System (GIS)-based studies have been published since the onset of COVID-19 (Abolfazl Mollalo, Vahedi, & Rivera, 2020). Many kinds of GIS applications and dashboards, such as the WHO dashboard, are capable of illustrating the spread of COVID-19 (Boulos & Geraghty, 2020). Mapping with GIS has been used to identify COVID-19 in important health care sites that are related to more vulnerable people, including those in hospices, the elderly, and disabled patients (Lakhani, 2020). These results suggest possible improvements in the quality of care during the pandemic (Gibson & Rush, 2020). GIS technology has also been used to outline boundaries between dwellings to detect the probability of COVID-19 in areas such as Cape Town, South Africa (Gibson & Rush, 2020). The findings suggested that the spread of COVID-19 can be lessened by enforcing social distancing procedures, as reinforced by buffer analysis and cluster identifications (Gibson & Rush, 2020).

Spatial models are important tools for examining the geographic relationships between many explanatory variables and disease outbreaks (A Mollalo, Alimohammadi, Shirzadi, & Malek, 2015; A Mollalo & Khodabandehloo, 2016). Mollalo et al. (2020) examined some regressive and autoregressive spatial models to see how effectively they could describe the differences of
COVID-19 in the United States, built on several factors with environmental, socioeconomic, behavioral, and demographic variables, among others (Abolfazl Mollalo et al., 2020). Mollalo’s study was the first to utilize local geography to model the spread of COVID-19 across the United States. It may help policymakers identify and implement specific interventions (Abolfazl Mollalo et al., 2020).

In the past two decades, the WHO and other health organizations have regularly used mapping and spatial analysis to control disease outbreaks (REZAEI, NOURI, PARK, & KIM, 2020). There have been some successful applications for observing and controlling some diseases, such as SARS or Ebola virus disease, through the application of GIS systems (REZAEI et al., 2020). Consequently, based on the research and experience of managing disease outbreaks, obtaining information about the locations of patients is crucial to the process of controlling and decision making (REZAEI et al., 2020). By employing such approaches from the onset of COVID-19, the local authorities and health experts of some countries have started to track and identify infected people through GIS systems (WHO, 2020a). Countries that have done this during the early stages of the outbreak have been more successful in combatting the disease than other countries (WHO, 2020a). One of the best examples in controlling the COVID-19 outbreak through a GIS application is South Korea (REZAEI et al., 2020). This country designed a unique GIS tool and mapping system for tracking, identifying, and observing infected people, and identifying the places that they visited prior to their infection being identified (MOFA, 2020). While other countries implemented a quarantine policy, South Korea had already started tracking cases through a system in which they indicated the places where patients had visited and those people who had been in direct contact with the infected person (REZAEI et al., 2020).

Geographically weighted regression (GWR) can overcome the restrictions of Ordinary Least Square (OLS) and has therefore been broadly applied to measure local-specific connections (LeSage, 2004; Su, Li, Xiao, & Zhang, 2014). GWR is capable of producing spatially varying coefficients, which promote further understanding of the phenomenon being investigated (Brunsdon, Fotheringham, & Charlton, 1998). The benefits of GWR for solving problems in many spheres of applied geography have been extensively discussed in previous literature (Ansong, Ansong, Ampomah, & Adjabeng, 2015; Cardozo, García-Palomares, & Gutiérrez, 2012; Cockx & Canters, 2015; Hu, Yang, Li, Zhang, & Xu, 2016; Ivajnšič, Kaligarič, & Žiberna,
The overall aim of this paper was to apply GWR to examining the relationships between COVID-19, population, and population density. This is the first study that has examined the risk of COVID-19 in Saudi Arabia and its transmission throughout the country’s regions.

2. Materials and methods

2.1 Study Area

The Makkah region is located in the central area of the western region of Saudi Arabia, near the coast of the Red Sea (Fig. 1). The area spans approximately 141 thousand km², comprising approximately 6.3% of the entire area of the Kingdom. According to the preliminary results of the General Census of the Population for the year 1431 (2010), the total population reached ~6.92 million people; this represents ~25.48% of the whole population of the Kingdom (~27.1 million people) (GASTAT, 2010). Out of the 17 cities (Governorates) in Makkah region, the largest city is Jeddah City, which has a population of 3,456,259 people; it also has the heaviest population density in the region. This is followed by Makkah City, which has a population of 1,675,368 people and has the second highest population density (GASTAT, 2010).
2.2 Data Analysis

The information used in this study was obtained from the Ministry of Health in Saudi Arabia’s daily data from all the regions and the 17 Cities in the Makkah region. Cases, recoveries, and mortalities data were obtained from the period of 2 March 2020 until May 24, 2020 (MOH, 2020). The population data from Saudi Arabia and Makkah region were obtained from the General Authority for Statistics (GASTAT, 2010).

For this study, two spatial methods were used to explore the relationships of COVID-19 with both population and population density in the Makkah region. Two of the most common methods used for studying spatial relationships are OLS and GRW. OLS is global regression model, but it tends to be unreliable when two or more variables show multicollinearity or high correlations (Brunsdon et al., 1998). Moreover, OLS is primarily used to display a dependent variable based
on its relationship with a few independent variables (Brunsdon et al., 1998). This global estimation technique does not capture local variation in the relationship (A. Fotheringham, Brunsdon, & Charlton, 2002; A. S. Fotheringham, Charlton, & Brunsdon, 2001). Therefore, GRW was more appropriate to use on a local scale, and it can handle spatial autocorrelation better than OLS (Brunsdon, Fotheringham, & Charlton, 2002).

To determine which method was more appropriate, the Akaike Information Criterion (AIC) (Akaike, 1974) was used along with the adjusted $R^2$. Better model performances are represented by lower AIC values and higher adjusted $R^2$ values (A. Fotheringham et al., 2002).

After comparing, the AIC was calculated from both OLS and GWR. The GWR had a better fit to the observed data, as the AIC (288) was lower and the $R^2$ (0.96) was higher, compared to OLS. Moreover, the GWR was more appropriately focused on a model that represented local regression.

2.3 GWR

GWR was employed to indicate the spatial association since it is able to determine the spatial dependency in a dataset which is based on a comprehensible traditional regression framework (A. S. Fotheringham, Brunsdon, & Charlton, 2003). To implement GWR for this study, the dependent and explanatory variables were established. The total number of COVID-19 incidences per city was established as the dependent variable, and each city’s population and population density were set as the explanatory variables. The kernel was identified as a fixed predetermined distance to decipher the regression analysis (Clement, Orange, Williams, Mulley, & Epprecht, 2009; Tu & Xia, 2008). The AIC specified the bandwidth to determine the range of the kernel (density) (A. Fotheringham et al., 2002). The number of cities or bandwidth that was utilized for each local estimation may have been the most important parameter for the GWR, as the model’s degree of smoothing was controlled (A. Fotheringham et al., 2002). GWR estimated predicted values that were used to create a model of people at risk of COVID-19. The software ArcGIS was used for analysis.
3. Results

3.1 Descriptive analysis

COVID-19 has affected many parts of Saudi Arabia. The most affected area is the Makkah region, with approximately 28,654 cases, representing 39.5% of all cases in the country. As well as having the highest number of cases of all of the regions, the region also has the most recoveries and mortalities (Table 1). In terms of the rate of COVID-19 cases, the Makkah region contains around 41 cases per 10,000 people. The second highest region is the Al-Riyad region, which has 23.2% of all cases (16,854). Other regions with a significant number of cases, ranging from 18.8% to 13.2%, include the Eastern Region and Al-Madinah regions, respectively. All other regions have low numbers of cases from 1.3% to 0.1%.

Saudi Arabia has 72,560 cases with 43,520 recoveries and 390 mortalities in total. As shown in Table 1 and Fig. 2, COVID-19 cases can be found all over the regions of Saudi Arabia, and the region with the highest rate of cases was found in Al-Madinah (54 cases per 10,000 people), followed by the Makkah region (41 cases per 10,000 people). Hail, Al-Jawf, and Najran all had much lower rates, with only 0.1 cases per 10,000 people for each of these regions.

<table>
<thead>
<tr>
<th>Region</th>
<th>Cases</th>
<th>Recoveries</th>
<th>Mortalities</th>
<th>Rate - cases per 10,000 people</th>
<th>Percentage of total cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Makkah</td>
<td>28,654</td>
<td>15,961</td>
<td>290</td>
<td>41</td>
<td>39.5%</td>
</tr>
<tr>
<td>Al-Riyad</td>
<td>16,854</td>
<td>9,604</td>
<td>24</td>
<td>25</td>
<td>23.2%</td>
</tr>
<tr>
<td>Eastern Region</td>
<td>13,652</td>
<td>8,848</td>
<td>22</td>
<td>33</td>
<td>18.8%</td>
</tr>
<tr>
<td>Al-Madinah</td>
<td>9,564</td>
<td>6,812</td>
<td>44</td>
<td>54</td>
<td>13.2%</td>
</tr>
<tr>
<td>Jazan</td>
<td>941</td>
<td>801</td>
<td>3</td>
<td>7</td>
<td>1.3%</td>
</tr>
<tr>
<td>Tabouk</td>
<td>777</td>
<td>423</td>
<td>1</td>
<td>10</td>
<td>1.1%</td>
</tr>
<tr>
<td>Al-Qassim</td>
<td>761</td>
<td>320</td>
<td>4</td>
<td>6</td>
<td>1.0%</td>
</tr>
<tr>
<td>Asir</td>
<td>571</td>
<td>404</td>
<td>1</td>
<td>3</td>
<td>0.8%</td>
</tr>
<tr>
<td>North Borders</td>
<td>212</td>
<td>55</td>
<td>1</td>
<td>7</td>
<td>0.3%</td>
</tr>
<tr>
<td>Hail</td>
<td>210</td>
<td>88</td>
<td>0</td>
<td>4</td>
<td>0.3%</td>
</tr>
<tr>
<td>Al-Bahah</td>
<td>148</td>
<td>83</td>
<td>0</td>
<td>4</td>
<td>0.2%</td>
</tr>
<tr>
<td>Najran</td>
<td>132</td>
<td>46</td>
<td>0</td>
<td>3</td>
<td>0.2%</td>
</tr>
<tr>
<td>Al-Jawf</td>
<td>84</td>
<td>75</td>
<td>0</td>
<td>2</td>
<td>0.1%</td>
</tr>
<tr>
<td>Total</td>
<td>72,560</td>
<td>43,520</td>
<td>390</td>
<td></td>
<td>100%</td>
</tr>
</tbody>
</table>
Table 2 shows the number of cases, recoveries, and mortalities in the Makkah region. Makkah City has the highest number of cases (14,132, with 8,300 recoveries and 174 mortalities). It has 49.3% of all cases in the cities (84 cases per 10,000 people).

Jeddah has 12,140 cases, which is the second highest number in the cities, having 42.4% of all cases (35 cases per 10,000 people). Other cities, such as Taif and Al-Qunfudah, have a significantly lower numbers of cases (1,701 and 175 cases, respectively) corresponding to only 5.9% to 0.6% of cases). The other cities also have very low numbers of cases, with Rabigh and Ranyah having 35 and 0 cases, respectively. Moreover, these cities have rate cases of 0 cases per 10,000 people. Figure 3 shows the location and density of COVID-19 cases inside the Makkah region. This map reveals that the areas of highest case density are located in the western area in Jeddah City and Makkah City.
Table 2 COVID-19 in Makkah region from 2/3/2020 to 24/5/2020

<table>
<thead>
<tr>
<th>Cities (Governorates)</th>
<th>Cases</th>
<th>Recoveries</th>
<th>Mortalities</th>
<th>Rate - cases per 10,000 people</th>
<th>Percentage of total cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Makkah</td>
<td>14,132</td>
<td>8,300</td>
<td>174</td>
<td>84</td>
<td>49.3%</td>
</tr>
<tr>
<td>Jeddah</td>
<td>12,140</td>
<td>6,741</td>
<td>114</td>
<td>35</td>
<td>42.4%</td>
</tr>
<tr>
<td>Taif</td>
<td>1,701</td>
<td>589</td>
<td>2</td>
<td>17</td>
<td>5.9%</td>
</tr>
<tr>
<td>Al-Qunfudhah</td>
<td>175</td>
<td>137</td>
<td>0</td>
<td>6</td>
<td>0.6%</td>
</tr>
<tr>
<td>Bahra</td>
<td>169</td>
<td>0</td>
<td>0</td>
<td>17</td>
<td>0.6%</td>
</tr>
<tr>
<td>Khulais</td>
<td>107</td>
<td>40</td>
<td>0</td>
<td>19</td>
<td>0.4%</td>
</tr>
<tr>
<td>Maysan</td>
<td>73</td>
<td>11</td>
<td>0</td>
<td>10</td>
<td>0.3%</td>
</tr>
<tr>
<td>Alordyat</td>
<td>46</td>
<td>30</td>
<td>0</td>
<td>6</td>
<td>0.2%</td>
</tr>
<tr>
<td>Rabigh</td>
<td>35</td>
<td>25</td>
<td>0</td>
<td>4</td>
<td>0.1%</td>
</tr>
<tr>
<td>Al-Muwayh</td>
<td>35</td>
<td>6</td>
<td>0</td>
<td>9</td>
<td>0.1%</td>
</tr>
<tr>
<td>Al-Khurma</td>
<td>14</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>0.0%</td>
</tr>
<tr>
<td>Al-Lith</td>
<td>13</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>0.0%</td>
</tr>
<tr>
<td>Adham</td>
<td>11</td>
<td>7</td>
<td>0</td>
<td>2</td>
<td>0.0%</td>
</tr>
<tr>
<td>Turbah</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0.0%</td>
</tr>
<tr>
<td>Al-Kamil</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0%</td>
</tr>
<tr>
<td>Al-Jumum</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0%</td>
</tr>
<tr>
<td>Ranyah</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>28,654</strong></td>
<td><strong>15,894</strong></td>
<td><strong>290</strong></td>
<td></td>
<td><strong>100%</strong></td>
</tr>
</tbody>
</table>
Fig. 3. COVID-19 cases in Makkah region per 10,000 people

3.2 GWR

The parameters used for the GWR were total population and population density for the 17 cities in Makkah region. These were used to build a prediction model of the likelihood of infection by COVID-19. The result obtained with GWR showed there was a strong positive association ($R^2 = 0.96$) between COVID-19 and population and population density in the Makkah region. Figure 3 shows the result of the GWR prediction model of COVID-19 risk for the cities in the Makkah region. People in the cities of Makkah and Jeddah have been at higher risk of COVID-19, according to the model. People in the cities of Ranyah and Taif were identified as being at lower
risk. People who were predicted to have been at mid risk of COVID-19 did not appear to differ much from those in low risk cities, as most of the COVID-19 cases have been in the cities of Makkah and Jeddah (Table 2).

![GWR prediction model of COVID-19 risk, based on population and population density](image)

**Fig. 3.** GWR prediction model of COVID-19 risk, based on population and population density

### 4. Discussion and conclusions

The COVID-19 outbreak in Saudi Arabia has mostly affected those people who live in the Makkah region. The number of COVID-19 cases that have occurred here has significantly impacted this region. This can be demonstrated through the use of GWR, which shows that this
pandemic has a spatial association. The focus of this study was to establish a causal relationship between COVID-19 and population and population density.

The most significant contribution of this study has been the usage of GWR, which enabled the spatial variability of the results to be examined. The identification of population and COVID-19 factors was therefore improved. GWR analysis facilitated the examination of variable behavior at the local level. A spatial component was also included, with its specification implying that the parameters estimated for this regression differed based on their geographical location. The results obtained from the GWR models indicate that some explanatory variables related to COVID-19, population, and population density did not have constant parameters in the study area.

Here, the implementation of spatial information analysis using GWR improved the understanding of the associations of the relationships between COVID-19, population, and population density in the Makkah region. The GWR showed that COVID-19 was strongly significant ($R^2=0.96$) to the population and population density. Cities that had larger populations and higher population densities were found to be at higher risk of COVID-19, and were capable of spreading the virus throughout the Makkah region and beyond. Moreover, the two highest risk cities (Makkah and Jeddah) had 91.7% of the total cases as well as the highest populations and population densities of the Makkah region. Furthermore, approximately 8.3% of the cases of the region were at mid and low risk of COVID-19. It can be safely inferred that there was a direct correlation between the transmission of COVID-19 and the population and population density of the cities in the Makkah region. Future studies should explore these trends, and should identify any improvements that can be implemented. More complex models should be employed to scrutinize the global and local relationships, while the associations should be compared on other geographical scales.

This study attempted to determine to what extent the factors included are related to COVID-19, population, and population density. The results were able to demonstrate that the application of statistical methods and GIS serves as a strong instrument for geographical analysis. GWR allows for the production of a wide variety of statistical maps from the results produced: AIC, local-$R^2$, and parameter estimates, amongst others. These methods can function as a methodological basis for future COVID-19 studies in this region or in other areas of the country.
References


