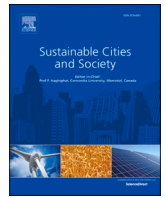




Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Sociodemographic determinants of COVID-19 incidence rates in Oman: Geospatial modelling using multiscale geographically weighted regression (MGWR)

Shawky Mansour^{a,b,*}, Abdullah Al Kindi^c, Alkhattab Al-Said^d, Adham Al-Said^d, Peter Atkinson^e

^a Geography Department, College of Arts and Social Sciences, Sultan Qaboos University, El Khodh, Aseeb, Muscat, Oman

^b Department of Geography and GIS, Faculty of Arts, Alexandria University, Al Shatby, Alexandria, Egypt

^c Department of Mass Communication, College of Arts and Social Sciences, Sultan Qaboos University, El Khodh, Aseeb, Muscat, Oman

^d Department of Economic and Finance, College of Economics and Political Sciences, Sultan Qaboos University, El Khodh, Aseeb, Muscat, Oman

^e Faculty of Science and Technology, Lancaster University, Lancaster, LA1 4YR, UK

ARTICLE INFO

Keywords:

GIS
Geospatial modeling
GWR
MGWR
COVID-19
Sociodemographic determinants
Oman

ABSTRACT

The current COVID-19 pandemic is evolving rapidly into one of the most devastating public health crises in recent history. By mid-July 2020, reported cases exceeded 13 million worldwide, with at least 575,000 deaths and 7.33 million people recovered. In Oman, over 61,200 confirmed cases have been reported with an infection rate of 1.3. Spatial modeling of disease transmission is important to guide the response to the epidemic at the subnational level. Sociodemographic and healthcare factors such as age structure, population density, long-term illness, hospital beds and nurse practitioners can be used to explain and predict the spatial transmission of COVID-19. Therefore, this research aimed to examine whether the relationships between the incidence rates and these covariates vary spatially across Oman. Global Ordinary Least Squares (OLS), spatial lag and spatial error regression models (SLM, SEM), as well as two distinct local regression models (Geographically Weighted Regression (GWR) and multiscale geographically weighted regression MGWR), were applied to explore the spatially non-stationary relationships. As the relationships between these covariates and COVID-19 incidence rates vary geographically, the local models were able to express the non-stationary relationships among variables. Furthermore, among the eleven selected regressors, elderly population aged 65 and above, population density, hospital beds, and diabetes rates were found to be statistically significant determinants of COVID-19 incidence rates. In conclusion, spatial information derived from this modeling provides valuable insights regarding the spatially varying relationship of COVID-19 infection with these possible drivers to help establish preventative measures to reduce the community incidence rate.

1. Introduction and background

Modeling the incidence rate of communicable diseases is vital during epidemics and pandemics. The ongoing coronavirus pandemic (COVID-19) was first reported in Wuhan city in China and has since spread globally causing thousands of fatalities. On 30th January 2020, the World Health Organization declared COVID-19 as a Public Health Emergency of International Concern (PHEIC) (WHO, 2020a, 2020b). The severe acute and infectious COVID-19 not only has posed threats to

local communities but has also brought disastrous consequences to the global economy (Buheji et al., 2020; Civelek & Xiarewana, 2020).

It is recognized widely that COVID-19 is not bound by territorial geography or national borders and, hence, the contemporary intensification of globalization, movement, communications, and socioeconomic activities has increased the potential speed, frequency, and geographical effects of the disease. Consequently, the transnational spread of COVID-19 has been considered as emanating from a range of globalizing forces beyond the control of individual states or governments, making the

* Corresponding author at: Department of Geography, College of Arts and Social Sciences, Sultan Qaboos University, P.O. Box: 42, Al-Khod P.C. 123, Muscat, Oman.

E-mail addresses: shmansour@squ.edu.om (S. Mansour), kindik@squ.edu.om (A. Al Kindi), h.alsaid@squ.edu.om (A. Al-Said), adham@squ.edu.om (A. Al-Said), pma@lancaster.ac.uk (P. Atkinson).

<https://doi.org/10.1016/j.scs.2020.102627>

Received 5 August 2020; Received in revised form 6 November 2020; Accepted 26 November 2020

Available online 2 December 2020

2210-6707/© 2020 Elsevier Ltd. All rights reserved.

outbreak a key challenge for collective action from a geographical perspective. In particular, the fundamental division between the national and international realms is no longer appropriate. Indeed, the outbreak of the disease in both developed and developing countries has demonstrated that the governance needs surrounding disease transmission and incidence are multifaceted and complex.

The existing literature is limited concerning the spatial variation of COVID-19 transmission and incidence at subnational boundaries. Moreover, there is growing evidence that the identification of the major driving forces of disease incidence rates and spatial diffusion is an intricate process. Recent literature has investigated the causes of disease transmission, mortality, and morbidity. Indeed, there have been many attempts to understand how, and to what extent, the transmission dynamics of COVID-19 is associated with prime environmental factors (Muhammad, Long, & Salman, 2020; Qu, Li, Hu, & Jiang, 2020; Saadat, Rawtani, & Hussain, 2020; Zhu & Xie, 2020). In India, Kumar et al. (2020) analysed the influences of anthropogenic emissions switch-off on ambient PM_{2.5} in five Indian cities during COVID-19 pandemic and found that improvements in air quality during the lockdown and restrictions procedures led to a reduction in air pollutants and gas emissions. Other studies have examined the associations between sociodemographic factors, public health interventions and the COVID-19 outbreak (Choi, Denice, Haan, & Zajacova, 2020; Qiu, Chen, & Shi, 2020; Sirkeci & Yucesahin, 2020; Wang, Zhong, & Hurd, 2020). The epidemiology of COVID-19 has been also considered in relation to demographic characteristics of the population, particularly age as an important determinant (Tian et al., 2020; WHO, 2020a, 2020b). High prevalence rates of long-term illnesses such as diabetes are significantly associated with increased COVID-19 risks in geographic zones (Fadini, Morieri, Longato, & Avogaro, 2020; Li et al., 2020; Muniyappa & Gubbi, 2020). Furthermore, individuals with diabetes are more at risk of a poorer outcome of COVID-19 infection (Guo et al., 2020; Shah & Hux, 2003). To protect workers in healthcare system from COVID-19 infection, Ge et al. (2020) assessed the exposure risks of the disease in the hospitals environments. The findings showed that infected patients in hospitals increase the nucleic acid of SARS-CoV-2 in the air and surface. Accordingly, monitoring and disinfection are necessary measures to lessen incidence rates during the pandemic. In a recent seminal work, Sannigrahi, Pilla, Basu, and Basu (2020), Sannigrahi, Pilla, Basu, Basu, and Molter (2020) examined the influences of several socio-demographic parameters on COVID-19 incidence and mortality across the European regions. Using local and global spatial models such as OLS and GWR, an explicit spatial modelling was performed. The findings of this study showed that population size, poverty, and household income were the key predictors of disease incidence in the European region.

Spatial analysis and modeling studies have attempted to simulate the effects of various explanatory variables on COVID-19 incidence rates (e.g. DiMaggio, Klein, Berry, & Frangos, 2020; Mollalo, Rivera, & Vahedi, 2020; Mollalo, Vahedi, & Rivera, 2020; Qiu et al., 2020; Sannigrahi, Pilla, Basu, Basu, 2020; Sannigrahi, Pilla, Basu, Basu, Molter, 2020; Scala et al., 2020). The advent and continued development of geospatial technology have enabled the global and local modeling of socioeconomic and environmental conditions that influence the prevalence of COVID-19. Indeed, GIS and geospatial techniques can play crucial roles in analyzing big data of the COVID-19 outbreak globally (Zhou et al., 2020). Providing valuable spatial information support for decision-making, several methods can be utilized for the analysis of disease transmission such as data aggregation, spatial tracking, simulation and prediction, spatial distribution, and clustering (Caprarello & Fletcher, 2014).

In India, climatic, geographical and topographical variables such as air temperature, precipitation, actual evapotranspiration, solar radiation, humidity, wind speed, topographic altitude and population density were utilized to model the number of COVID-19 infections (Gupta, Kumar Patel, Sivaraman, & Mangal, 2020; Gupta, Banerjee, & Das,

2020). In an early analysis of COVID-19 transmission, Velraj and Haghghat (2020) investigated the theory of respiratory droplet drying within different weathering and indoor environments in India. Areas with low humidity were found to lead to high rates of disease infection. In another study, quantile regression was used to predict the effects of globalization, population development, and demographic characteristics on COVID-19 transmission globally (Sigler et al., 2020). A seminal work by Mollalo, Rivera et al. (2020), Mollalo, Vahedi et al. (2020) estimated the COVID-19 cumulative incidence rate in the USA using a multilayer perceptron (MLP) neural network. The findings revealed a 65 % association with the reference data for the holdout samples, while the output of the logistic regression model explained the presence/absence of disease incidence hotspots across the USA. Chen, Jiao, Bai, and Lindquist (2020) used global Ordinary Least Squares (OLS) and local Geographically Weighted Regression (GWR) models to investigate and simulate the spatial factors of COVID-19 in New York city, indicating that specific factors of green space, mean travel distance, male percentage, and commuting (walking, carpooling, and public transit) were significantly associated with higher rates of COVID-19 incidence. Spatio-temporal patterns and the spatial dependency of COVID-19 in the early stages of the disease in China were investigated using Moran's *I* statistics (Kang, Choi, Kim, & Choi, 2020), showing that from 22nd January, the outbreak spread rapidly from Wuhan to other neighboring areas through the transportation network.

Although the cumulative severity patterns among developing countries are more muted, by the mid of June 2020, many low and lower-middle-income countries (LICs and LMICs) such as Brazil and India showed a high daily severity of COVID-19, with Iran in the Middle East particularly affected. Chronologically, up to mid-March 2020, Iran reported 14,991 confirmed cases, ranked as the third most affected country after China and Italy, and a significant source of the outbreak spreading in the Middle East. Arab-Mazar, Sah, Rabaan, Dhama, and Rodriguez-Morales (2020) addressed the spatial distribution of cases across the Iranian provinces, finding that the transition was from the north-central provinces such as Tehran and Qom towards surrounding areas. The spatial distribution and spreading patterns of COVID-19 in Iran have been linked to environmental and spatial drivers, such as intra-provincial movement, temperature, precipitation, humidity, wind speed, and average solar radiation (Ahmadi, Sharifi, Dorosti, Ghoushchi, & Ghanbari, 2020). Asna-ashary, Farzanegan, Feizi, and Sadati (2020) adopted a panel vector autoregressive (PVAR) approach to investigate the relationship between the disease incidence rate and air pollution across the Iranian provinces, whereas Azarafza, Azarafza, and Tanha (2020) utilized long short-term memory-based deep learning for time series forecasting at the provincial geographical scale, indicating that machine learning was an effective method in modeling the disease incidence.

A few nonspatial studies have been conducted to address various issues associated with COVID-19 effects in GCC countries. For instance, in response to the disease outbreak, the Umrah pilgrimage services were suspended in Saudi Arabia, which hugely impacted the economy and placed a burden on several public and private sectors, particularly tourism and hospitality, airlines and transportation, as well as other local businesses (Ebrahim & Memish, 2020). Rahman et al. (2020) modelled a framework of a data-driven dynamic clustering to alleviate the adverse economic effects of Covid-19 lockdown restrictions. The results indicated that the proposed algorithms improved the relevant metrics by approximately 50 % in the lockdown experiments and 60–80 % in the conceivable lessening of economic loss. In another study (Sun & Zhai, 2020) predicted the infection probability of COVID-19 through investigating social distancing and ventilation strategies as effective measures to mitigate disease infection risks and transmission. Based on fostering the capabilities of communities to tackle disease spread, a key study by Megahed and Ghoneim (2020) found that designing healthy urban infrastructures and sustainable architectures may be effective planning strategies in response to the COVID-19 pandemic to diminish

the risk of infection. In Oman, a system dynamic epidemic spread model was developed to simulate disease prevalence temporally, indicating that during stringent social distancing and testing strategies, a small perturbation can lead to higher rates of infection (Zia & Farooq, 2020).

Despite the growing body of research regarding COVID-19 epidemics in several countries within North America, Europe, Asia and Latin America (e.g. Mollalo, Rivera et al., 2020; Mollalo, Vahedi et al., 2020; Sannigrahi, Pilla, Basu, Basu, 2020; Sannigrahi, Pilla, Basu, Basu, Molter, 2020; Sigler et al., 2020; Wang et al., 2020), spatial assessment and modeling of disease transmission in the GCC is rare. Accordingly, and drawing from spatial modeling of COVID-19 incidence that has been conducted elsewhere, this research attempts to bridge this gap by developing a local spatial model of disease incidence rates in Oman. Various sociodemographic determinants, three regressive global models (OLS, SLM, and SEM) and two local models (GWR, and MGWR) were employed to predict spatial variation in the COVID-19 distribution across the Omani subnational boundaries.

The main aims of this research were:

- To determine which sociodemographic risk factors relate to COVID-19 incidence rates.
- To investigate whether the relationship between these socio-demographic characteristics and disease incidence rates vary geographically.
- To examine the impact that sociodemographic parameters have on the pandemic incidence, particularly how these affect the relationships between the sociodemographic risk factors on infection rates geographically.

To our knowledge, this is the first study in the GCC countries and Oman which adopts advanced spatial techniques to analyze spatial patterns of disease prevalence. Such analysis could be valuable for health governance and provide key insights into alleviating the risks of disease transmission.

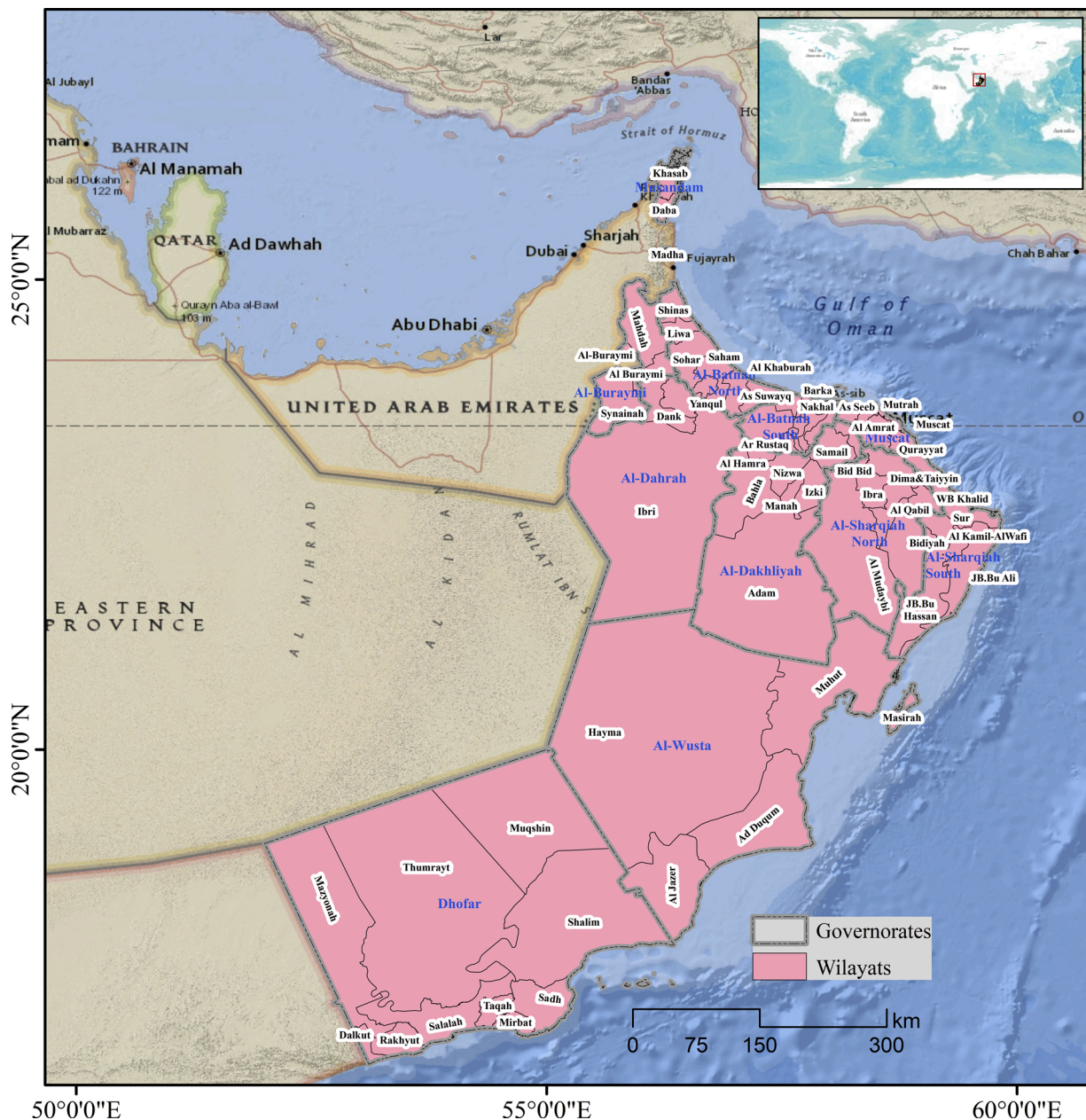


Fig. 1. Location of the study area.

2. Study area and dataset

2.1. Study area

The Sultanate of Oman is situated in the far south-eastern corner of the Arabian Peninsula, extending from latitudes 16.40–26.20 degrees north and longitudes 51.50–59.40 degrees east. Oman shares borders with the United Arab Emirates (UAE) in the north and west, the Kingdom of Saudi Arabia in the west and the Republic of Yemen in the south (Fig. 1). The country covers a total land area of approximately 309,500 km² including several islands in the Gulf of Oman and the Arabian Sea. The country covers a diverse range of topographies, including mountain ranges, arid deserts, valleys (82 %), fertile plains, oasis, and islands (MRMWR, 2008). Most of central Oman is covered with a vast gravel desert plain, while the Al Hajar mountain ranges extending from the north-west of the country toward the south-east constitutes a barrier between the northern coasts and the interior desert (Garzanti, Andò, Vezzoli, & Dell’era, 2003). The climate is hot and dry in the interior areas, and humid in coastal areas. During the summer months, when most of the surrounding regions experience increasing temperatures, the monsoon occurs in the south of the country, particularly the Dhofar governorate, resulting in cooling rains and pleasant temperatures (Edgell, 2006).

Administratively, Oman is divided into 11 governorates and 61 Wilayats (states), with the Muscat governorate being the largest in terms of population size. It consists of six Wilayats including Muscat City as the capital of Oman. Approximately, 50 % of the total population (4.7 million) is concentrated in Muscat and the Batinah coastal plain governorates northwest of the capital (National Center for Statistics and Information (NCSI, 2020). In 2017, the annual population growth rate had decreased to 3.3 %, while the population density was 14.7 persons per km² (Year book, 2018). The life expectancy at birth is 77.9 years, with a crude death rate of 2.4 per 1000 population in 2019 (UN, 2019). The highest population density was recorded in Muscat governorate with 364.8 persons per km².

2.2. Dataset

The data used in this study were derived from the Ministry of Health (MOH) in Oman, which is responsible for monitoring COVID-19 daily at the Wilayat level across the country. Disease prevalence from 24th February to June was collected and the crude incidence rate was computed at the subnational level (Fig. 1). A geodatabase was developed within the GIS environment and ArcGIS Desktop 10.6 was utilized to link the health and demographic explanatory and response variables to the administrative boundary shapefile of the Omani administrative geographic units (Table 1). On 4th July 2020, the overall number of deaths was 203 47 % of which were aged 60 years and above (MOH, 2020).

Five global (OLS, SLM, and SEM) and local (GWR and MGWR) models were implemented to understand which sociodemographic factors are related to the incidence of COVID-19 in Oman. The local modeling process is an effective approach that builds upon traditional global regression by allowing non-stationary (local) rather than stationary parameter estimates to be computed (Fig 2).

3. Methods and model specification

3.1. Global regression modeling

3.1.1. Ordinary least squares (OLS)

The OLS linear regression is an inferential technique utilized primarily to regress a dependent variable on other explanatory variables. The technique assumes a stationary and constant relationship over space. Accordingly, the implicit independence assumptions associated with the geographical data of interest may not hold (Hutcheson, 2011;

Table 1

Description of explanatory variables and data sources.

Parameters	Description	Source	Rationale to disease incidence rates
Population density	The number of people per Wilayat calculated by dividing the total number of people by total land area.	NCSI, Oman	There is a significant association between population density, overcrowding and COVID-19 transmission (Gupta, Kumar Patel et al., 2020; Gupta, Banerjee et al., 2020).
Number of hospital beds	The total number of beds in all hospital and health centers that are regularly maintained and available for patient care in each Wilayat.	NCSI, Oman	The capacity of healthcare system and hospital beds provide protection for non-infected population through isolating and treating infected people (Khanjehani, 2020; Liang, Tseng, Ho, & Wu, 2020).
Population aged 65+	Population aged 65 and above as a percentage of the total population in each Wilayat.	NCSI, Oman	There is a significant association between population density, overcrowding and COVID-19 transmission.
Diabetes rate	The prevalence of diabetes among adults calculated as the number of people with diabetes divided by the total population aged 18 and above.	MOH, Oman	Diabetes is considered as a risk factor for COVID-19 infection. High diabetes rate is likely to be associated with high COVID-19 infection (Guo et al., 2020; Shah & Hux, 2003).
South Asian immigrants	The number of South Asian immigrants (Indian, Bangladeshi, Pakistani, Phillipian, and Seri Lankan) divided by the total number of immigrants in each Wilayat.	NCSI, Oman	South Asian immigrants are the largest groups in Oman and quite often they live in isolated and overcrowding households (Mansour, 2017).
Western immigrants	Number of immigrants from western countries divided by the total number of immigrants in each Wilayat.	NCSI, Oman	Examining whether the impacts of COVID-19 vary among immigrants from different groups (Choi et al., 2020).
Arab immigrants	Number of immigrants from Arab countries divided by the total number of immigrants in each Wilayat.	NCSI, Oman	Examining whether there is an association between disease incidence rate and this group of immigrants.
Crude death rate	The total number of deaths divided by the total population in each Wilayat and multiplied by 1000.	MOH, Oman	Examining whether there is a significant correlation between crude death rate and COVID-19 incidence.
Number of physicians	The total number of registered medical physicians in each Wilayat calculated as the doctor to population ratio of 1:1000.	MOH, Oman	The number of physicians is critical parameter in isolating suspected patients, and supporting infection prevention policy (Buerhaus, Auerbach, & Staiger, 2020).
Number of nurses	The total number of registered medical nurses and midwives in each Wilayat calculated as nurses and midwives to population ratio of 1:1000.	MOH, Oman	Examining whether there is a relationship between the number of nurses per 1000 population and COVI-19 infection.

Pohlmann & Leitner, 2003). In this research, the relationship between a response variable (y) and a collection of regressors or predictors (x₁, x₂, x₃, x_n...) is presented as a line of best fit. The OLS is characterized by:

$$y_i = \beta_0 + b_1x_1 + b_2x_2 + b_nx_n + e \tag{1}$$

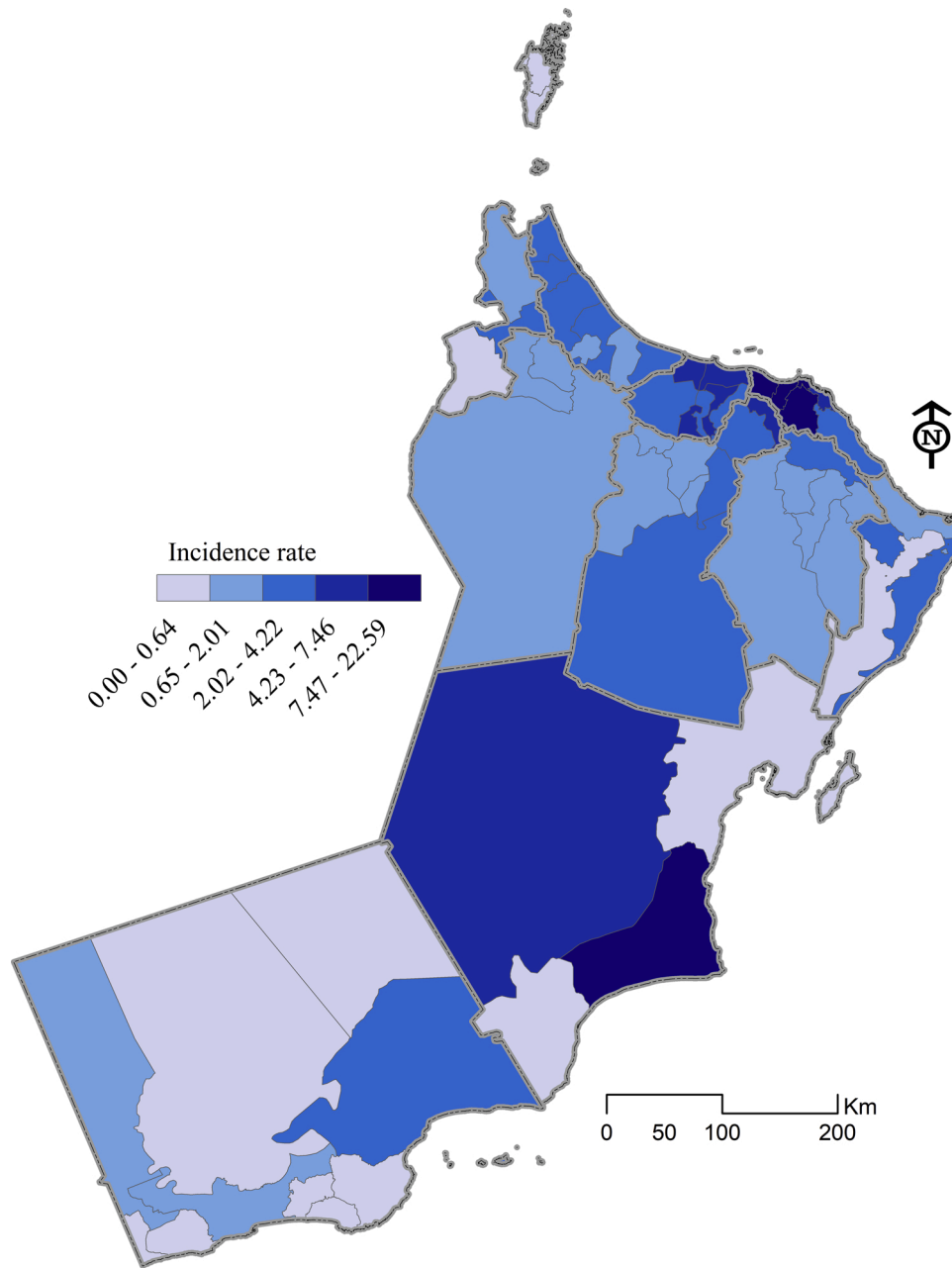


Fig. 2. Distribution of the dependent variable (COVID-19 incidence rate) across subnational boundaries.

where y_i signifies the dependent variable observation (COVID-19 incidence rate) at the i th locations (Omani Wilayats), β_0 is the estimated intercept and indicates the value of y when x equals to zero, b_1 is the parameter estimate for x_1 . x_n denotes the set of explanatory variables, b_n indicates the regression coefficients that describe changes in the dependent variable y when x changes by one unit.

To assess the magnitude of any multicollinearity in the regression model, the variance inflation factor (VIF) was adopted. For every explanatory variable, a VIF value greater than 10 indicates a state of high intercorrelation with other predictors and that the variable should be removed (Montgomery, Peck, & Vining, 2012). The VIF factor is calculated as follows:

$$VIF_i = \frac{1}{1 - R^2_i} \tag{2}$$

where R^2 represents the coefficient of determination in the regression

equation.

3.1.2. Spatial autoregressive models

Although the OLS model assumes that COVID-19 incidence rates at the subnational level are independent of each other, it does not consider spatial dependence, and many explanatory variables may be spatially correlated and omitted from the model. Therefore, the OLS was considered as a misspecified model (Ward & Gleditsch, 2018), and two spatial autoregressive models SLM and SEM were employed since both are variants of OLS and account for spatial weights and dependence (Anselin, 1988, 2003).

3.1.2.1. Spatial lag model (SLM). The model estimates the spillover effects of a specific variable over space, assuming the dependency between the response variables and a set of explanatory variables. The SLM is calculated as follows:

$$y_i = \beta_0 + x_i\beta + \rho w_i + \varepsilon_i \quad (3)$$

where ρ is the spatial lag coefficient (spatial autoregressive parameter), while w is a spatial weights matrix demonstrating the distance relationship between observations i and j . The weights matrix and the spatially lagged dependent variable both depict the spillover effects from adjacent locations.

3.1.2.2. Spatial error model (SEM). The SEM utilises a form of spatial dependence which works through the error terms instead of the dependent variable, in this case, the distribution of errors over spatial units (Anselin, 2003; Oud & Folmer, 2008). The SEM model includes a spatial autoregressive error term and is computed as follows:

$$y = \alpha + \rho W y + (\beta)x + \varepsilon \text{ with } \varepsilon = \lambda(W)\varepsilon + \mu \quad (4)$$

Where y represents an element vector of observations of the dependent variable COVID-19 incidence rate and α is the intercept, $W y$ specifies the spatially lagged dependent variable for weighting contiguity matrix W , while ρ illustrates the spatial autoregressive parameter of $W y$, which is estimated for the whole model, ε denotes the vector of error term, λ designates the coefficient of spatially lagged autoregressive errors, $(W)\varepsilon$, while μ indicates the vector of independent identically distributed errors.

3.2. Local regression modeling

3.2.1. Geographically weighted regression (GWR)

The above global regression models assume that the relationship between the response and explanatory variables is stationary (i.e., constant) across a study area which means these relationships do not vary over space (Brunsdon, Fotheringham, & Charlton, 1996; Fotheringham, Charlton, & Brunsdon, 1998). To relax this assumption and allow the parameters to vary spatially, GWR assumes that non-stationary relationships exist between the response variable and the explanatory variable(s). Therefore, the model estimates a local parameter for each location separately (Brunsdon et al., 1996; Charlton, Fotheringham, & Brunsdon, 2009; Fotheringham et al., 1998). The GWR model is calculated as follows (Fotheringham, Brunsdon, & Charlton, 2003):

$$y_i = \beta_{0i}(u_i, v_i) + \sum_{n=1}^k \beta_{ni}(u_i, v_i)x_{ni} + \varepsilon_i \quad (5)$$

where y_i is the COVID-19 rate at Wilayat i , (u_i, v_i) represents the coordinates of the centroid of Wilayat i , β_{0i} , β_{ni} indicates the local estimated intercept and effect of variable n for Wilayat i , respectively, x_{ni} refers to the values of the i th explanatory variables while ε_i denotes a random error term. Parameter estimates for each independent variable and at each Wilayat in matrix form is given by (Fotheringham & Oshan, 2016):

$$\hat{\beta}(i) = (X'W(i)X)^{-1}X'W(i)y \quad (6)$$

where $\hat{\beta}$ indicates the vector of parameter estimates ($p \times 1$), X displays the matrix of the selected independent variable ($n \times p$), $W(i)$ is the matrix of spatial weights ($n \times n$) while y implies the vector observation of COVID-19 rates ($p \times 1$). The matrix $W(i)$ is constructed from the weights of each spatial unit according to its distance from location i . Implementing the model calibration, the Gaussian and biquare weighting kernel functions are the most common techniques where the Wilayats nearby to i have larger effects on the estimation of $\beta_{ni}(u_i, v_i)$ than those located farther from i . The kernel function and bandwidth should be specified where the bandwidth is determined based on the Euclidean distance and number of nearest neighbors. Selecting different bandwidths affects the type of neighborhoods in which local weighting occurs (Mollalo, Rivera et al., 2020; Mollalo, Vahedi et al., 2020).

3.2.2. Multiscale geographically weighted regression (MGWR)

GWR provides advantages to the regression modeling process and accounts for spatial variations, considering the spatial scale to be constant over space, however, in many cases, a fixed spatial scale is not valid where phenomena involve numerous spatial processes with various spatial scales. MGWR allows the relationship between the response variable and explanatory variables to vary spatially and at different scales (Fotheringham, Yang, & Kang, 2017; Mollalo, Rivera et al., 2020; Mollalo, Vahedi et al., 2020; Yu et al., 2020), incorporating various bandwidths over the study area surface and is calculated as follows:

$$y_i = \sum_{j=0}^m \beta_{b_{wj}} x_{ij} + \varepsilon_i \quad (7)$$

where b_{wj} represents the bandwidths, which are utilized to calibrate the j th conditional relationship (Fotheringham et al., 2017). For each process, every single bandwidth exhibits a unique scale where the model relies on a backfitting algorithm that derives a set of bandwidths for the j processes. Compared to GWR, the model has several advantages, particularly it can accurately depict spatial heterogeneity, diminish collinearity, and lessen the bias in the parameter estimates (Oshan, Li, Kang, Wolf, & Fotheringham, 2019; Wolf, Oshan, & Fotheringham, 2018).

3.3. Model fitting

A range of sociodemographic and health variables are included within the modeling process to determine which factors are associated with COVID-19 incidence rate. Giving the relatively large number of explanatory variables, a stepwise forward procedure was implemented to eliminate the non-significant predictors and identify a single model with the best fit. Furthermore, the VIF was used to check the multicollinearity among the independent variables and consequently, the uncorrelated regressors were selected as the input of the regression model.

To express the potential for interaction between observations at each pair of spatial units, a spatial weight matrix was generated based on first-order Queens' contiguity which specifies whether spatial units share a boundary or not. The key function of the spatial weight matrix is to represent the structure of spatial features (polygons) and quantify the existing relationships among them (Wang et al., 2020). As the spatial autocorrelation is embedded in the local model (GWR), the spatial weight matrix is an indispensable component in the model (Brunsdon, Fotheringham, & Charlton, 2002). According to Tobler's first law of geography (Tobler, 1970) and when the model parameters are estimated, the neighboring zones (Wilayats) of the model's variables should obtain more weights than distant ones. In this study, the selected first-order Queens' contiguity reflects how administrative spatial zones (Wilayats) interact with each other.

Developing the local models, the kernel function type of adaptive biquare with its bandwidth size was specified. The bandwidth is adaptive defined as the proportion of data points involved in the calibration process of a local estimate, which eliminates the influence of spatial units outside the neighborhood. Evaluating and comparing the model fit and performance, a corrected Akaike Information Criterion (AICc) was used to select the optimal bandwidth (Guo, Ma, & Zhang, 2008; Oshan et al., 2019). All estimates of the coefficients included within the final fitted model were statistically highly significant with p-values less than 0.005. The best model fit is indicating by a larger adjusted R^2 and a smaller AIC value.

4. Findings

The outcome of the OLS global model is shown in Table 2. The coefficient estimates are highly significant with p-values less than 0.005. The VIF values of all independent variables revealed low multicollinearity (all values were less than the threshold of 5). The coefficient

Table 2
Summary statistics of global OLS model.

Variable	Coefficient	St. Error	t-Statistic	Probability	VIF
Intercept	6.4842	1.5567	4.1651	0.0001	–
Population 65+	0.1853	0.0381	4.8602	0.0000	1.4064
N. of hospital beds	–0.0642	0.0032	–2.0103	0.0042	1.9766
Population density	0.0771	0.0010	7.0855	0.0000	1.8069
Diabetes rate (per 1000)	0.0660	0.0301	2.1909	0.0026	1.2695

estimates are a mix of positive and negative values. For example, the coefficient estimates of elderly population (aged 65 and above) are positive indicating that an increase in elderly population size is associated with an increased rate of disease incidence. The coefficient estimates also indicate that the rate of the elderly population was the most influential variable, followed by population density and diabetes rate.

The global OLS model demonstrated a low adjusted R² (0.58) indicating that 41 % of the variance in COVID-19 incidence rates across the Omani Wilayats remain unexplained and caused by unknown variables (Table 4) spatial non-stationarity exists in the relationship between the response and explanatory variables. Hence, the OLS global model which assumes a constant functional structure across space is not sufficient to describe the underlying relationship. The performance of global modeling was improved by using the SAR models (SLM and SEM) to characterise the relationship between the sociodemographic drivers and the disease incidence rate. The lag coefficients were strongly positive (p < 0.000) and the adjusted R² of both SLM and SEM increased to 0.62 and 0.65, respectively (Table 3), while the AICc slightly decreased. Nonetheless, although the SLM and SEM models provided a closer fit than OLS according to the R² and AICc diagnostic criteria, the models were unable to reveal any spatial variation that might occur in the inter-variable relationships across the study area. This can be attributed to the neglected scale of spatial processes involved in modeling the disease incidence rates. Accordingly, a spatially non-stationary local modeling approach was adopted next.

To explore the local spatial variation in the relationships with the COVID-19 incidence rates, GWR and MGWR were applied to the same set of predictors used in the global models. The diagnostics of GWR indicated a relatively improved adjusted R² and AICc. Fitting GWR using optimal bandwidth of 54.0, the model R² increased to 0.69 while the AICc decreased to 122.55 which is substantially smaller than that of the global model (304.16) indicating a better fit. However, among all fitted models, MGWR represented the largest adjusted R² (0.71) and lowest AICc (120.141). Figs. 3 and 4 map the coefficients of the GWR and MGWR for the statistically significant variables. The coefficient estimates of the elderly population (65+) vary across the study area (Fig. 3). In both GWR and MGWR, the variable represents similar patterns in describing the spatial distribution of COVID-19 incidence rates at the Wilayat level. The largest positive values in some areas are associated with higher COVID-19 incidence rates, specifically within Muscat governorates. However, there were also positive coefficient estimates in some Wilayats where the percentage of the population aged 65 and

Table 3
Summary statistics of SLM and SEM models.

Variable	Coefficient		St. Error		Z-score		P-value	
	SLM	SEM	SLM	SEM	SLM	SEM	SLM	SEM
Intercept	5.9343	6.4712	0.1804	1.4864	1.3930	4.3536	0.1636	0.2310
Population 65+	0.7215	0.0076	1.5271	0.0010	3.8859	7.2550	0.0001	0.0000
N. of hospital beds	–0.1743	–0.1853	0.0010	0.0367	6.8964	–5.0473	0.0000	0.0000
Population density	–0.0621	–0.0064	0.0370	0.0030	–4.6994	–2.1012	0.0000	0.0006
Diabetes rate (per 1000)	0.0530	0.0663	0.0030	0.0294	–2.0420	2.2544	0.0004	0.0001
Rho	0.2513	–	0.0316	–	11.7538	–	0.0019	–
Lambda	–	0.0935	–	0.2541	–	0.3679	–	0.0079

Table 4
Comparison of the goodness of fit measures for the global and local models.

Criterion	OLS	SLM	SEM	GWR	MGWR
Adj.R ²	0.581	0.621	0.651	0.697	0.711
AICc	307.261	304.162	303.612	122.556	120.142

above was low, such as ASeeb, Bowsher (Muscat governorate), and Barka (South Al-Batnah). The number of hospital beds in each Wilayat was an influential factor in explaining the variation in the COVID-19 incidence rates across the northwest study area (e.g. Saham, Liwa) and within the Muscat governorate (Aseeb, Qurayat).

The coefficient estimates of hospital beds show a negative association with COVID-19 incidence rates in Wilayats of the Al-Wusta and Dhofar governorates. Other Wilayats, such as Adam and Iski (Al-Dakhaliya) and Dank (Al-Dhahra) where the GWR and MGWR models were less predictive, have a low number of hospital beds but a statistically negative association with disease incidence rate.

Fig. 4 illustrates that in both GWR and MGWR, the population density was a significant regressor in describing the spatial distribution of COVID-19 incidence rates in the study area, particularly in the Wilayats located in the north and northeast. The GWR coefficient map of population density demonstrates that positive estimates in the spatial pattern stretched from the northwest to the southeast, in particular, Al-Batnah South as well as some of Al-Sharqiyah Wilayats, suggesting a large correlation between population concentration and disease incidence in these areas. A similar spatial structure appears in the MGWR coefficient map, where a variation of coefficient estimates across the regions can be depicted. Overall, the smaller negative coefficient estimates were found in the northern and southern parts of the country, indicating a weaker relationship between population density and COVID-19 incidence rates. In contrast, the larger estimates in the north of the country are associated with a larger population.

Fig. 4 also shows the GWR and MGWR coefficient results for diabetes rate, with a range of positive GWR coefficient values located in the north of the study area, specifically in the South Al-Batnah and Muscat governorates. The coefficient estimates are smaller in the southeast some of the areas, particularly along the coastal Wilayats. This shows that the diabetes covariate explains the COVID-19 incidence rate where people with diabetes are more likely to be at increased risk of infection. The spatial patterns of estimation coefficient for diabetes follow the observed pattern of COVID-19 incidence rates, where some Wilayats within Al-Dakhliyah, South Al-Batnah, and Muscat governorates show notably elevated observed incidence rates. However, the impact of diabetes on COVID-19 incidence rates was inconsistent between the GWR and MGWR models.

Fig. 5 reveals the spatial heterogeneity in terms of subnational fitting which was reflected in the spatially varying local R² of both the GWR and MGWR models. The GWR model fits better in the northwest, specifically Ibri, Liwa, Shinas, and Musandam, whereas the local R² was consistently low in the central and south of the country. Similarly, the local R² of MGWR indicates that the model also predicts accurately and explains local relationships better (R² > 0.7) in the northern parts and

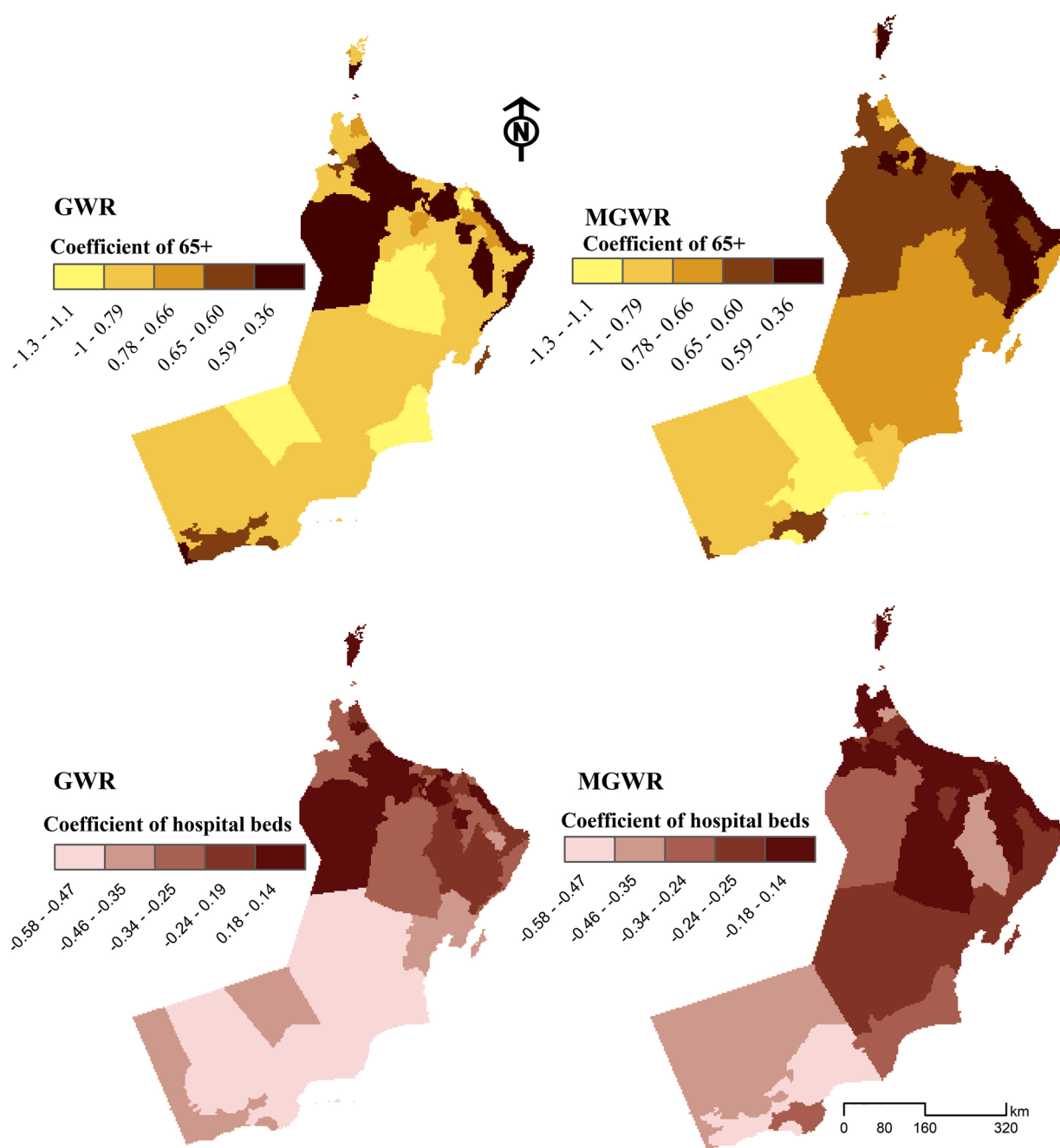


Fig. 3. The effects of populations aged 65+ (above) and hospital beds (below) in describing COVID-19 incidence rates utilizing GWR (left) and MGWR (right) models across the Omani Wilayats.

some Wilayats in the east, such as Sur and Jaalan Bani Bu Hassan. Noteworthy, the larger local R^2 distribution pattern is associated with Wilayats characterized by large population size and a higher number of health facilities, whereas the smaller local R^2 coefficients ($R^2 < 0.5$) of Al-Wusta and Dhofar governorates reveals a poor model fit, probably due to the significant local variability of their spatial and demographic characteristics.

The local model performance was demonstrated by plotting the observed against predicted values of the COVID-19 rate (Fig. 6). Although the goodness-of-fit of both local models was higher than the global model, the MGWR provided slightly richer estimates for the COVID-19 incidence rates compared to OLS and GWR. The summary statistics for the local coefficients of MGWR are provided in Table 5, including the mean, standard deviation, median, maximum, and bandwidth. Overall, the global models produced more generalized spatial

patterns that hindered local variation, while the GWR and MGWR produced more accurate estimates of disease rate by taking into account local characteristics and spatial heterogeneity.

In this research, a cross-validation process was adopted to spatially determine how many zones (Wilayats) were needed to solve the local regression model structure. Consequently, AICs was employed to form the adaptive number and identify the best appropriate number of neighbouring polygons using the adaptive kernel method (Brunsdon et al., 1996). Moreover, the calibration of the local models (GWR & MGWR) rendered a matrix of optimal bandwidths (Table 6) which is considered as a spatial scale of the model processing. For example, the explanatory variables of population aged 65+ (BW:49 neighbours) and Hospital beds (BW:49 neighbours) operated on a local scale in the MGWR compared to their process in the GWR (BW: 54 & 57 respectively).

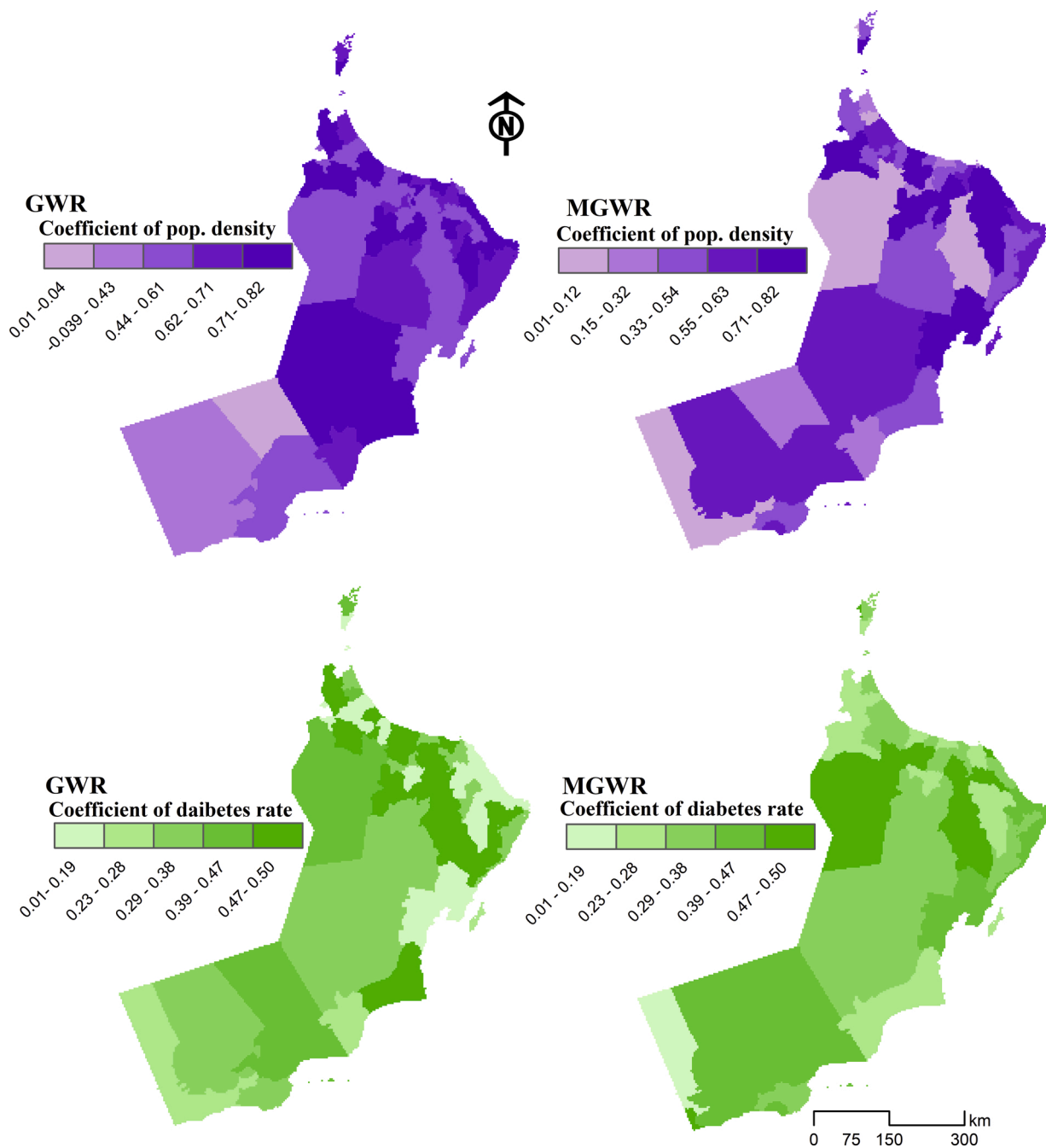


Fig. 4. The effects of population density (above) and diabetes (below) in describing COVID-19 incidence rates utilizing GWR (left) and MGWR (right) models across the Omani Wilayats.

The residuals are the differences between the observed and predicted values of COVID-19 rates. Considering the spatial distribution of residuals is an important key indicator of model structure and performance. Fig. 7a & b illustrates a comparison between the residuals of GWR and MGWR models. The light colors show low residuals while the darker colors reveal either overestimation or underestimation. Overall, both maps show random distribution with strong consistency across the study area. Similarly, Fig. 7c & d exhibit fitted residuals against the predicted values of the dependent variables in both models and it seems that the distribution demonstrates a random pattern of over-under estimation which indicates properly specified local models.

5. Discussion

The current COVID-19 pandemic has evolved rapidly into one of the most threatening and devastating public health crises in recent history. While most COVID-19 studies have been undertaken from a medical perspective or focused on epidemiological evolution, there is a growing literature applying spatial analysis and disease mapping, particularly in the developed world (e.g. Chen et al., 2020; Mollalo, Rivera et al., 2020; Mollalo, Vahedi et al., 2020). Nevertheless, spatial modeling of disease spread at subnational boundaries, especially in developing countries is still rare. Spatial modeling of disease outbreaks is important not only to assess where and why hotspots and clusters are located but also to provide explicit perceptions of the spatial variation in disease incidence and transmission. Likewise, little work has been conducted regarding

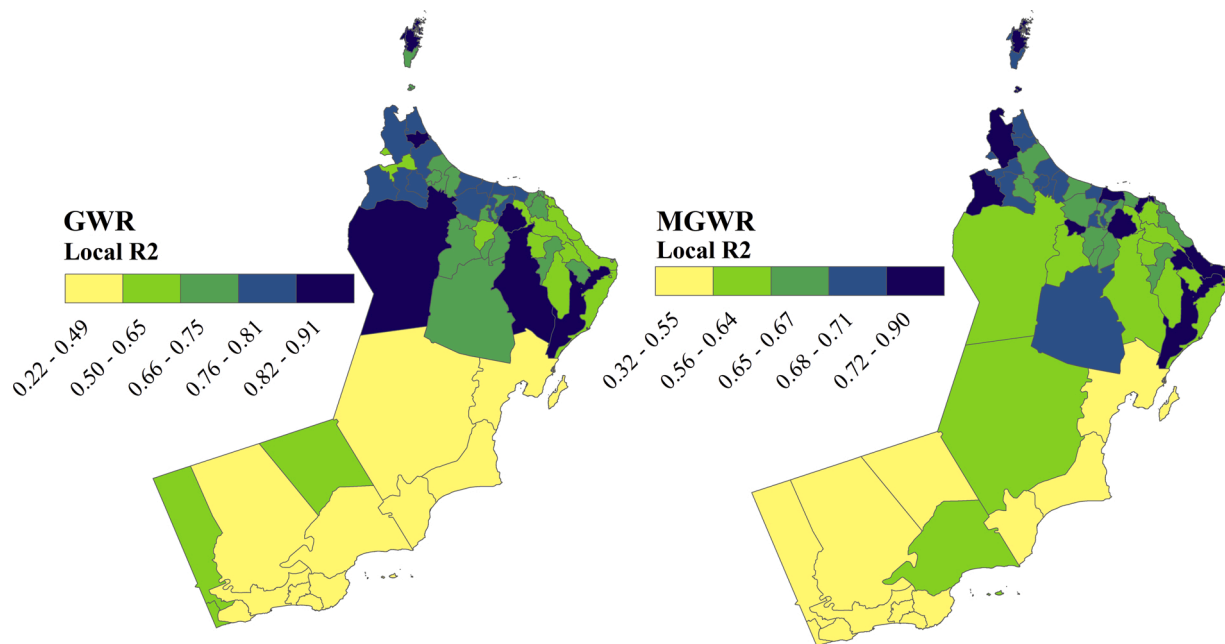


Fig. 5. Spatial distribution of local R^2 of GWR and MGWR models for COVID-19 incidence rate associated with the significant covariates across the Omani Wilayats.

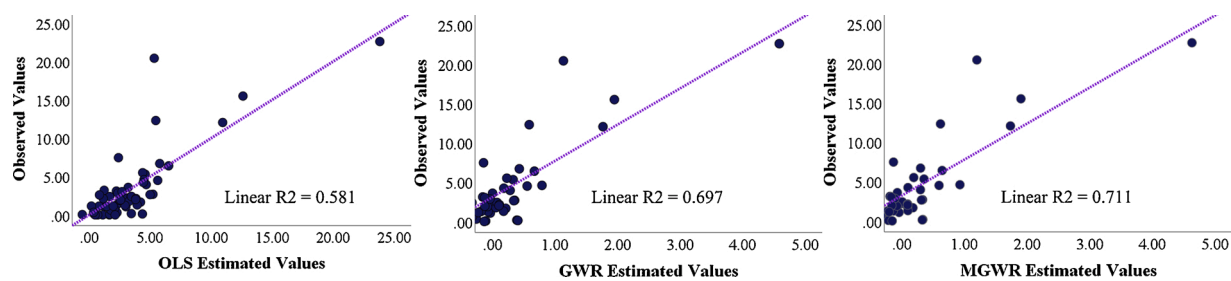


Fig. 6. Observed values of COVID-19 versus estimated values of OLS, GWR, and MGWR models.

Table 5
Summary statistics for MGWR parameter estimates.

Variable	Mean	STD	Min	Median	Max
Intercept	-0.037	0.008	-0.054	-0.037	-0.021
Population 65+	0.475	0.131	0.887	0.414	0.355
N. of hospital beds	-0.209	0.117	-0.456	-0.154	-0.124
Population density	0.752	0.003	0.745	0.752	0.760
Diabetes rate (per 1000)	0.245	0.021	0.190	0.255	0.259

Table 6
Multiscale bandwidth for the local models: GWR and MGWR.

Variable	GWR Bandwidth	MGWR Bandwidth
Intercept	60	60
Population 65+	54	49
N. of hospital beds	57	49
Population density	60	60
Diabetes rate (per 1000)	60	60

how sociodemographic determinants of the COVID-19 incidence rates vary geographically across the GCC states. Therefore, the overall goal of this research was to provide clear insights into the relationships between sociodemographic characteristics and disease incidence rates in Oman at the subnational level.

Methodologically, several advanced geospatial techniques offer an opportunity to drastically increase the accuracy of spatial estimates of

disease incidence rates at the local geographical scale. In this research, while global models (OLS, SLM, SEM) performed well, local (GWR, MGWR) based on an ensemble of local regressions, provided a parsimonious quantitative representation of the sociodemographic determinants that may influence COVID-19 incidence rates. The outputs of MGWR overcame the disadvantages and drawbacks of global modeling where the relationship between variables is constant across the study area.

The key findings from this research were that a set of sociodemographic and health variables were found to impact on COVID-19 incidence rate and that these factors vary geographically. More specifically, our results indicated a considerable spatial heterogeneity in COVID-19 incidence rates along the urban-rural gradient. The higher rates of disease incidence within the Muscat governorate were specifically across urban Wilayats and associated with higher population densities, consistent with previous studies that reported the positive influence of overcrowding and population densities on transmission rates (Sigler et al., 2020; Sirkeci & Yucesahin, 2020). In contrast, low incidence rates of COVID-19 in North Al-Batnah Wilayats were associated with low influences of explanatory factors such as population age, density, and diabetes, which may be explained by their more rural, semi-urban and less crowded residential environment.

Accounting for the impact of demographic characteristics on disease incidence risk, population aging is fundamental and contributes significantly to the spread of COVID-19 (Sun, Lu, Xu, Sun, & Pan, 2020; Tian et al., 2020). The findings demonstrated that the elderly population (aged 65+) was an influential factor in explaining spatial variation in

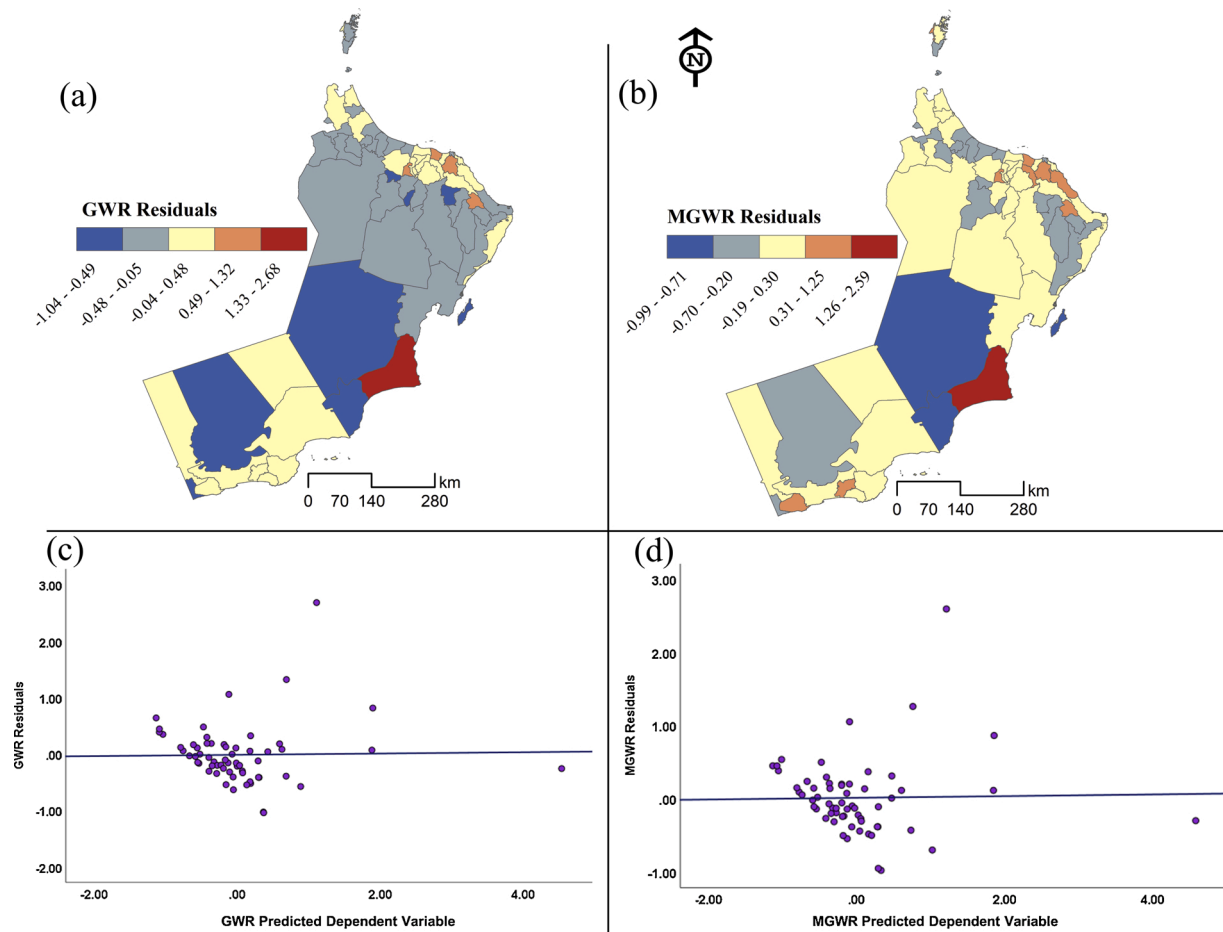


Fig. 7. Spatial distribution of the MGWR and GWR residuals.

disease incidence, particularly in the Wilayats located in the northwest and northeast of the country. These areas are characterized by a small proportion of immigrants and, thus a lower percentage of young people, consequently, as the disease may progress faster in the elderly than in the young, the incidence rates were higher among the Omani elderly.

There is a debate about the relationship between underlying diabetes and COVID-19 incidence rates, with some studies reporting that diabetes is a risk factor for COVID-19 infection (e.g. Guo et al., 2020; Shah & Hux, 2003), while other studies indicating that the evidence remains controversial regarding whether diabetes increases vulnerability and influences the consequences from infections (Fadini et al., 2020; Li et al., 2020). In this research, COVID-19 incidence was associated with diabetes, specifically in the central northern part of the Muscat governorate, with a low local association between infection rates and diabetes in southern areas. The geographical difference in diabetes rates was considered as an important indicator for explaining variation in disease incidence across the subnational zones of the study area. Ultimately, areas with large elderly populations and high diabetes rates were more at risk of disease prevalence.

The findings of the local models (GWR and MGWR) pointed to the substantial impact of the health system and medical care on disease incidence. Although two variables (physicians and nurse practitioners) were not significant predictors, hospital beds had a strong negative relationship with the disease rate. Wilayats of Muscat and Al-Batnah governorates located in the northern part of the country showed large coefficient estimates compared to other Wilayats in the central and southern parts. This distribution pattern seems to be reasonable as in the north, most settlements are urban neighborhoods with an increased number of public and private hospitals.

In Oman, the government has attempted to control the spread of disease, with preventative measures constituting the first line of defense policy. Additionally, to prevent the geographical spread of COVID-19, several spatial procedures have been implemented, such as applying quarantine procedures to Mutrah Wilayat from 1st April 20220 due to the higher rates of infection among residents, followed by the complete isolation of Muscat governorate from 10th to 22nd April. However, in reality, overcoming this epidemic requires many integrated strategies that extend to all social, economic, logistical, and humanitarian aspects as well as geographical measures. Furthermore, COVID-19 infection rates cannot be studied geographically through global models, but rather disease prevalence should be assessed in the context of possible spatial dependence, autocorrelation, and non-stationary relationships.

This analysis is an attempt to provide a greater understanding of the sociodemographic determinants of COVID-19, delving deeper into where these factors are likely to influence disease incidence. Subsequently, this research has various policy implications. First, the outcome of this analysis can serve as a spatial guideline for decision-makers to formulate COVID-19 mitigation strategies across Oman and other surrounding countries. Second, quantifying the sociodemographic determinants of COVID-19 incidence rates on the subnational scale not only can provide spatially explicit information about disease infection drivers but also can help scale up intervention pathways to identify localities at high risk of infection. Third, although the output of our analysis showed that specific demographic characteristics, such as immigrant classes and nationalities, were not significantly influential parameters on COVID-19, further research should include additional explanatory variables to model spatially the disease incidence rates. Finally, similar local modeling utilizing GWR and MGWR can be applied

elsewhere within the GCC region or in other developing countries within a different spatial framework, including various determinants, particularly ecological and environmental parameters, to identify spatial aspects of susceptibility, vulnerability, exposure and risk factors of COVID-19 infections in local communities.

Furthermore, the findings may provide effective and clear guidelines regarding developing subnational plan to mitigate disease incidence. For instance, according to the results of this study, policy makers can establish spatial monitoring frameworks based on the major socio-demographic indicators and risk factors influencing disease transmission across local communities. Likewise, a risk assessment plan can be developed which targets lessening inequalities in healthcare systems particularly reducing the gaps in health facilities coverages and accessibility to hospitals between communities in urban areas, major cities and small villages and rural settlements. In addition, any spatial action plan at subnational scale should be developed to enhance community preparedness and resilience in mitigating COVID-19 incidence rates.

The outcomes of the local models substantially can help effectively to map and identify vulnerable populations and social groups that are at high risks of disease infection such as elderly people, people with long term illness, and patients in hospitals. Besides, several spatial actions can also be taken by the Omani decision makers to reduce the disease transmission through intervention, preventive and mitigation measures that reduce introducing or reintroducing the virus from high-infected places to low-infected or non-infected communities. Also, disease incidence risks and spatial clusters of high rates and their associations to sociodemographic covariates and other social customs are essential and should be taken into account. For instance, in most of the Omani rural and Bedouin communities, traditional events, gathering and social celebrations continue which may accelerate infection risks and disease transmission.

Due mainly to data availability, most COVID-19 studies to-date have been conducted at the area level, most commonly at the national level. Similarly, the sub-national level of the analysis was selected based on data availability. However, in comparison to such national studies, our research has the advantage of 'zooming in' to reveal spatial variation at the subnational scale. Nevertheless, this is an ecological study utilizing aggregated data of disease incidence rate. As such, it is acknowledged that our findings are necessarily influenced by scaling and zoning configuration choices implicit in the data. For this reason, it was not possible to integrate data on the personal characteristics of infected individuals or data at a finer spatial scale than the sub-national level. Another potential limitation is that the infection rate was calculated based on confirmed cases only, while suspected cases were not considered due to lack of availability and uncertainty.

6. Conclusion

Currently, at the time of writing, despite a range of prevention strategies, particularly social distancing, lockdowns and stay-at-home restrictions, COVID-19 is still widespread globally and is yet to be brought fully under control. Therefore, there is a need for more research on spatial modeling of disease transmission and dynamics at the community-level to identify possible drivers that may affect infection rates, and investigate the intricate connections between those factors. The importance of capturing the spatial variability of local COVID-19 rates in the developed world is increasingly recognized. However, in most developing countries, datasets regarding disease prevalence are limited. Consequently, modeling the spatial variation in the incidence rates and determinants of disease is challenging.

This research utilized a spatial modeling framework to explore the relationship between incidence rates and a set of sociodemographic and health covariates in the Omani subnational administrative zones. Global (OLS, SLM, SEM) and local (GWR, MGWR) regression models were applied and compared to explain the spatial variation in infection rate. Overall, while the global models did not adequately represent the

observed spatial variation, the goodness-of-fit of local models was greater and the models were able to capture local patterns, especially MGWR which provided the best overall fit. Moreover, and compared to using GWR alone, the MGWR was useful for modelling spatial variation in COVID-19 incidence rates at local geographic scales. In addition, incorporating the MGWR model lessens the effects of the common problem of aggregated data, particularly Modifiable Areal Unit Problem (MAUP) (Fotheringham & Wong, 1991; Openshaw, 1979). Accordingly, the spatial variability of this model across the Omani Wilayats depicted various spatial patterns of COVID-19 rates in response to significant sociodemographic parameters.

To the best of our knowledge, this is the first study to model COVID-19 incidence rates at subnational boundaries in the GCC states. Hence, the outcomes of this analysis could allow a closer focus on COVID-19 prevalence and its sociodemographic covariates to further mitigate incidence rates and support controlling the disease not only in Oman or the GCC states but also across other developing nations.

Declaration of Competing Interest

The authors report no declarations of interest.

References

- Ahmadi, M., Sharifi, A., Dorosti, S., Ghouschi, S. J., & Ghanbari, N. (2020). Investigation of effective climatology parameters on COVID-19 outbreak in Iran. *Science of the Total Environment*, Article 138705.
- Anselin, L. (1988). Spatial econometrics: Methods and models. In *Studies in operational regional science* (Vol. 4). Dordrecht: Springer Netherlands.
- Anselin, L. (2003). Spatial externalities, spatial multipliers, and spatial econometrics. *International Regional Science Review*, 26(2), 153–166.
- Arab-Mazar, Z., Sah, R., Rabaan, A. A., Dhama, K., & Rodriguez-Morales, A. J. (2020). Mapping the incidence of the COVID-19 hotspot in Iran—Implications for travellers. *Travel Medicine and Infectious Disease*, 34, 101630.
- Asna-ashary, M., Farzanehan, M. R., Feizi, M., & Sadati, S. M. (2020). COVID-19 outbreak and air pollution in Iran: A panel VAR analysis. Retrieved from.
- Azarafza, M., Azarafza, M., & Tanha, J. (2020). COVID-19 infection forecasting based on deep learning in Iran. medRxiv.
- Brunsdon, C., Fotheringham, A., & Charlton, M. (2002). Geographically weighted summary statistics—A framework for localised exploratory data analysis. *Computers, Environment and Urban Systems*, 26(6), 501–524.
- Brunsdon, C., Fotheringham, A. S., & Charlton, M. E. (1996). Geographically weighted regression: A method for exploring spatial nonstationarity. *Geographical Analysis*, 28(4), 281–298.
- Buerhaus, P. I., Auerbach, D. I., & Staiger, D. O. (2020). Older clinicians and the surge in novel coronavirus disease 2019 (COVID-19). *JAMA*, 323(18), 1777–1778.
- Buheji, M., da Costa Cunha, K., Beka, G., Mavric, B., de Souza, Y., da Costa Silva, S. S., ... Yein, T. C. (2020). The extent of covid-19 pandemic socio-economic impact on global poverty. A global integrative multidisciplinary review. *American Journal of Economics*, 10(4), 213–224.
- Caprarelli, G., & Fletcher, S. (2014). A brief review of spatial analysis concepts and tools used for mapping, containment and risk modelling of infectious diseases and other illnesses. *Parasitology*, 141(5), 581–601.
- Charlton, M., Fotheringham, S., & Brunsdon, C. (2009). Geographically weighted regression. *White paper. National centre for geocomputation*. National University of Ireland Maynooth.
- Chen, Y., Jiao, J., Bai, S., & Lindquist, J. (2020). Modeling the spatial factors of COVID-19 in New York City. Available at SSRN 3606719.
- Choi, K., Denice, P., Haan, M., & Zajacova, A. (2020). Studying the social determinants of COVID-19 in a data vacuum. *UCLA CCPR Population Working Papers*.
- Civelek, M. E., & Xiarewana, B. (2020). Effects of covid-19 on China and the world economy: Birth pains of the post-digital ecosystem.
- DiMaggio, C., Klein, M., Berry, C., & Frangos, S. (2020). Blacks/African Americans are 5 times more likely to develop COVID-19: Spatial modeling of New York City ZIP code-level testing results. medRxiv.
- Ebrahim, S. H., & Memish, Z. A. (2020). Saudi Arabia's drastic measures to curb the COVID-19 outbreak: Temporary suspension of the Umrah pilgrimage. *Journal of Travel Medicine*, 27(3). taaa029.
- Edgell, H. S. (2006). *Arabian deserts: Nature, origin and evolution*. Springer Science & Business Media.
- Fadini, G., Morieri, M., Longato, E., & Avogaro, A. (2020). Prevalence and impact of diabetes among people infected with SARS-CoV-2. *Journal of Endocrinological Investigation*, 1.
- Fotheringham, A. S., & Oshan, T. M. (2016). Geographically weighted regression and multicollinearity: Dispelling the myth. *Journal of Geographical Systems*, 18(4), 303–329.
- Fotheringham, A. S., & Wong, D. W. (1991). The modifiable areal unit problem in multivariate statistical analysis. *Environment and planning A*, 23(7), 1025–1044.

- Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2003). *Geographically weighted regression: The analysis of spatially varying relationships*. John Wiley & Sons.
- Fotheringham, A. S., Charlton, M. E., & Brunsdon, C. (1998). Geographically weighted regression: A natural evolution of the expansion method for spatial data analysis. *Environment and Planning A*, 30(11), 1905–1927.
- Fotheringham, A. S., Yang, W., & Kang, W. (2017). Multiscale geographically weighted regression (MGWR). *Annals of the American Association of Geographers*, 107(6), 1247–1265.
- Garzanti, E., Andò, S., Vezzoli, G., & Dell'era, D. (2003). From rifted margins to foreland basins: Investigating provenance and sediment dispersal across desert Arabia (Oman, UAE). *Journal of Sedimentary Research*, 73(4), 572–588.
- Ge, X.-Y., Pu, Y., Liao, C.-H., Huang, W.-F., Zeng, Q., Zhou, H., ... Zhou, P.-C. (2020). Evaluation of the exposure risk of SARS-CoV-2 in different hospital environment. *Sustainable Cities and Society*, 61, Article 102413.
- Guo, L., Ma, Z., & Zhang, L. (2008). Comparison of bandwidth selection in application of geographically weighted regression: A case study. *Canadian Journal of Forest Research*, 38(9), 2526–2534.
- Guo, W., Li, M., Dong, Y., Zhou, H., Zhang, Z., Tian, C., & Du, K. (2020). Diabetes is a risk factor for the progression and prognosis of Covid-19. *Diabetes/Metabolism Research and Reviews*, e3319.
- Gupta, A., Banerjee, S., & Das, S. (2020). Significance of geographical factors to the COVID-19 outbreak in India. *Modeling Earth Systems and Environment*, 1.
- Gupta, S., Kumar Patel, K., Sivaraman, S., & Mangal, A. (2020). Global epidemiology of first 90 days into COVID-19 pandemic: Disease incidence, prevalence, case fatality rate and their association with population density, urbanisation and elderly population. *Journal of Health Management*, 22(2), 117–128.
- Hutcheson, G. D. (2011). Ordinary least-squares regression. *L. Moutinho and GD Hutcheson, the SAGE dictionary of quantitative management research* (pp. 224–228).
- 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), 96–102.
- Khanijahani, A. (2020). *County-level proportions of Black and Hispanic populations, and socioeconomic characteristics in association with confirmed COVID-19 cases and deaths in the United States*. medRxiv.
- Kumar, P., Hama, S., Omidvarborna, H., Sharma, A., Sahani, J., Abhijith, K., & Tiwari, A. (2020). Temporary reduction in fine particulate matter due to “anthropogenic emissions switch-off” during COVID-19 lockdown in Indian cities. *Sustainable Cities and Society*, 62, Article 102382.
- Li, B., Yang, J., Zhao, F., Zhi, L., Wang, X., Liu, L., ... Zhao, Y. (2020). Prevalence and impact of cardiovascular metabolic diseases on COVID-19 in China. *Clinical Research in Cardiology*, 109(5), 531–538.
- Liang, L.-L., Tseng, C.-H., Ho, H. J., & Wu, C.-Y. (2020). Covid-19 mortality is negatively associated with test number and government effectiveness. *Scientific reports*, 10(1), 1–7.
- Mansour, S. (2017). Spatial concentration patterns of South Asian low-skilled immigrants in Oman: A spatial analysis of residential geographies. *Applied Geography*, 88, 118–129.
- Megahed, N. A., & Ghoneim, E. M. (2020). Antivirus-built environment: Lessons learned from Covid-19 pandemic. *Sustainable Cities and Society*, 61, Article 102350.
- Ministry of Regional Municipalities, W. M. (2008). *Water resources in Oman. Sultanate of Oman*. http://www.omanws.org.om/images/publications/5465_Water Atlas E.pdf.
- Ministry of Health (MOH), (2020), Oman, Statistical Data, <https://www.moh.gov.om/en/home>.
- Mollalo, A., Rivera, K. M., & Vahedi, B. (2020). Artificial neural network modeling of novel coronavirus (COVID-19) incidence rates across the continental United States. *International Journal of Environmental Research and Public Health*, 17(12), 4204.
- Mollalo, A., Vahedi, B., & Rivera, K. M. (2020). GIS-based spatial modeling of COVID-19 incidence rate in the continental United States. *Science of the Total Environment*, Article 138884.
- Montgomery, D., Peck, E., & Vining, G. (2012). *Wiley series in probability and mathematical statistics* (2nd edn.). New York: John Wiley & Sons, 1992. Introduction to linear regression analysis.
- Muhammad, S., Long, X., & Salman, M. (2020). COVID-19 pandemic and environmental pollution: A blessing in disguise? *Science of the Total Environment*, Article 138820.
- Muniyappa, R., & Gubbi, S. (2020). COVID-19 pandemic, coronaviruses, and diabetes mellitus. *American Journal of Physiology-Endocrinology and Metabolism*, 318(5), E736–E741.
- National Centre for Statistics and Information (NCSI), Oman. (2020). <https://www.ncsi.gov.om/Pages/AllIndicators.aspxzzzz>.
- Openshaw, S. (1979). Alternative methods of estimating spatial interaction models and their performance in short-term forecasting. *Exploratory and explanatory statistical analysis of spatial data* (pp. 201–225). Springer.
- Oshan, T. M., Li, Z., Kang, W., Wolf, L. J., & Fotheringham, A. S. (2019). mgwr: A Python implementation of multiscale geographically weighted regression for investigating process spatial heterogeneity and scale. *ISPRS International Journal of Geo-Information*, 8(6), 269.
- Oud, J. H., & Folmer, H. (2008). A structural equation approach to models with spatial dependence. *Geographical Analysis*, 40(2), 152–166.
- Pohlmann, J. T., & Leitner, D. W. (2003). A comparison of ordinary least squares and logistic regression (1). *The Ohio Journal of Science*, 103(5), 118–126.
- Qiu, Y., Chen, X., & Shi, W. (2020). Impacts of social and economic factors on the transmission of coronavirus disease 2019 (COVID-19) in China. *Journal of Population Economics*, 1.
- Qu, G., Li, X., Hu, L., & Jiang, G. (2020). *An imperative need for research on the role of environmental factors in transmission of novel coronavirus (COVID-19)*. ACS Publications.
- Rahman, M. A., Zaman, N., Asyhari, A. T., Al-Turjman, F., Bhuiyan, M. Z. A., & Zolkipli, M. (2020). Data-driven dynamic clustering framework for mitigating the adverse economic impact of Covid-19 lockdown practices. *Sustainable Cities and Society*, 62, Article 102372.
- Saadat, S., Rawtani, D., & Hussain, C. M. (2020). Environmental perspective of COVID-19. *Science of the Total Environment*, Article 138870.
- Sannigrahi, S., Pilla, F., Basu, B., & Basu, A. S. (2020). *The overall mortality caused by covid-19 in the European region is highly associated with demographic composition: A spatial regression-based approach*. arXiv preprint arXiv:2005.04029.
- Sannigrahi, S., Pilla, F., Basu, B., Basu, A. S., & Molter, A. (2020). Examining the association between socio-demographic composition and COVID-19 fatalities in the European region using spatial regression approach. *Sustainable Cities and Society*, 62, Article 102418.
- Scala, A., Flori, A., Spelta, A., Brugnoli, E., Cinelli, M., Quattrociocchi, W., ... Pammolli, F. (2020). *Time, space and social interactions: Exit mechanisms for the COVID-19 epidemics*. arXiv preprint arXiv:2004.04608.
- Shah, B. R., & Hux, J. E. (2003). Quantifying the risk of infectious diseases for people with diabetes. *Diabetes Care*, 26(2), 510–513.
- Sigler, T., Mahmuda, S., Kimpton, A., Loginova, J., Wohland-Jakhar, P., Charles-Edwards, E., ... Corcoran, J. (2020). The socio-spatial determinants of COVID-19 diffusion: The impact of globalisation. *Settlement Characteristics and Population*, 11, 1–26. <https://doi.org/10.21203/rs.3.rs-33615/v1>
- Sirkeci, I., & Yucesahin, M. M. (2020). Coronavirus and migration: Analysis of human mobility and the spread of COVID-19. *Migration Letters*, 17(2), 379–398.
- Statistical Year Book (2018). <https://www.ncsi.gov.om/Elibrary/Pages/LibraryContentDetails.aspx?ItemID=GxJqSZUD0v4K7T%2FPJp13A%3D%3D>.
- Sun, P., Lu, X., Xu, C., Sun, W., & Pan, B. (2020). Understanding of COVID-19 based on current evidence. *Journal of Medical Virology*, 92(6), 548–551.
- Sun, C., & Zhai, Z. (2020). The efficacy of social distance and ventilation effectiveness in preventing COVID-19 transmission. *Sustainable Cities and Society*, 62, Article 102390.
- Tian, S., Hu, N., Lou, J., Chen, K., Kang, X., Xiang, Z., & Liu, D. (2020). Characteristics of COVID-19 infection in Beijing. *Journal of Infection*, 80(4), 401–406.
- Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography*, 46(sup1), 234–240.
- United Nation. (2019). *World mortality highlight*. <https://www.un.org/en/development/desa/population/publications/pdf/mortality/WMR2019/WMR2019.Highlights.pdf>.
- Velraj, R., & Haghigat, F. (2020). The contribution of dry indoor built environment on the spread of coronavirus: Data from various Indian states. *Sustainable Cities and Society*, 62, Article 102371.
- Wang, A.-L., Zhong, X., & Hurd, Y. (2020). *Comorbidity and sociodemographic determinants in COVID-19 mortality in an US urban healthcare system*. medRxiv.
- Ward, M. D., & Gleditsch, K. S. (2018). *Spatial regression models sage publications university of essex, England*, 155.
- Wolf, L. J., Oshan, T. M., & Fotheringham, A. S. (2018). Single and multiscale models of process spatial heterogeneity. *Geographical Analysis*, 50(3), 223–246.
- World Health Organization. (2020a). *Population-based age-stratified seroepidemiological investigation protocol for coronavirus 2019 (COVID-19) infection, 26 May 2020*. Retrieved from.
- World Health Organization (WHO). (2020b). *Report of the WHO-China joint mission on coronavirus disease 2019 (COVID-19)*. <https://www.who.int/docs/default-source/coronavirus/who-china-joint-mission-on-covid-19-final-report.pdf>.
- Yu, H., Fotheringham, A. S., Li, Z., Oshan, T., Kang, W., & Wolf, L. J. (2020). Inference in multiscale geographically weighted regression. *Geographical Analysis*, 52(1), 87–106.
- Zhou, C., Su, F., Pei, T., Zhang, A., Du, Y., Luo, B., ... Zhu, Y. (2020). COVID-19: Challenges to GIS with big data. *Geography and Sustainability*, 1(1), 77–87.
- Zhu, Y., & Xie, J. (2020). Association between ambient temperature and COVID-19 infection in 122 cities from China. *Science of the Total Environment*, Article 138201.
- Zia, K., & Farooq, U. (2020). *COVID-19 outbreak in Oman: Model-driven impact analysis and challenges*. medRxiv.