

# Demographic and socioeconomic determinants of COVID-19 across Oman - A geospatial modelling approach

Khalifa M. Al Kindi,<sup>1</sup> Adhra Al-Mawali,<sup>2</sup> Amira Akharusi,<sup>3</sup> Duhai Alshukaili,<sup>4</sup> Noura Alnasiri,<sup>1,5</sup> Talal Al-Awadhi,<sup>1</sup> Yassine Charabi,<sup>1,5</sup> Ahmed M. El Kenawy<sup>1,6</sup>

<sup>1</sup>Geography Department, College of Arts & Social Sciences, Sultan Qaboos University, Muscat, Oman;

<sup>2</sup>Director/Centre of Studies & Research, Ministry of Health, Muscat, Sultanate of Oman;

<sup>3</sup>Physiology Department, College of Medicine & Health Sciences, Sultan Qaboos University, Muscat, Oman;

<sup>4</sup>University of Technology & Applied Sciences, Nizwa, Oman; <sup>5</sup>Center for Environmental Studies &

Research, Muscat, Oman; <sup>6</sup>Department of Geography, Mansoura University, Mansoura, Egypt

## Abstract

Local, bivariate relationships between coronavirus 2019 (COVID-19) infection rates and a set of demographic and socioeconomic variables were explored at the district level in Oman. To limit multicollinearity a principal component analysis was conducted, the results of which showed that three components together could explain 65% of the total variance that were therefore subjected to further study. Comparison of a generalized linear model

(GLM) and geographically weighted regression (GWR) indicated an improvement in model performance using GWR (goodness of fit=93%) compared to GLM (goodness of fit=86%). The local coefficient of determination ( $R^2$ ) showed a significant influence of specific demographic and socioeconomic factors on COVID-19, including percentages of Omani and non-Omani population at various age levels; spatial interaction; population density; number of hospital beds; total number of households; purchasing power; and purchasing power per km<sup>2</sup>. No direct correlation between COVID-19 rates and health facilities distribution or tobacco usage. This study suggests that Poisson regression using GWR and GLM can address unobserved spatial non-stationary relationships. Findings of this study can promote current understanding of the demographic and socioeconomic variables impacting the spatial patterns of COVID-19 in Oman, allowing local and national authorities to adopt more appropriate strategies to cope with this pandemic in the future and also to allocate more effective prevention resources.

Correspondence: Khalifa M. Al Kindi, Geography Department, College of Arts & Social Sciences, Sultan Qaboos University, Muscat, Oman. E-mail: k.alkindi@squ.edu.om

Key words: COVID-19; geographically weighted regression; generalized linear model; geographical information systems; Oman; socioeconomic determinants.

Acknowledgements: we sincerely thank the NCSI, Oman, for affording the demographic data employed in this research. Our gratefulness extends to the Omani MOH for affording the daily real-time data of COVID-19 via the Tarassud+ App (TA+). We also thank the MSD in Oman for providing the Social charity index.

Funding: this research is funded by The Ministry of the Higher Education, Research and Innovation, Sultanate of Oman (TRC/CRP/SQU/COVID-19/20/02).

Conflict of interests: the authors declare no potential conflict of interests.

See online Appendices for additional materials.

Received for publication: 6 February 2021.

Revision received: 13 April 2021.

Accepted for publication: 14 April 2021.

©Copyright: the Author(s), 2021  
Licensee PAGEPress, Italy  
Geospatial Health 2021; 16:985  
doi:10.4081/gh.2021.985

This article is distributed under the terms of the Creative Commons Attribution Noncommercial License (CC BY-NC 4.0) which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## Introduction

The coronavirus 2019 (COVID-19) pandemic has had an immense global influence on physical and human environments (Ciotti *et al.*, 2019; Weissleder *et al.*, 2020) threatening more than 220 countries and regions around the world (Bonaccorsi *et al.*, 2020; Karaye & Horney, 2020; Liu *et al.*, 2020b; Verma & Prakash, 2020), the Sultanate of Oman (Oman) among them (Wiwanitkit, 2020). According to the Omani Ministry of Health (MOH) (<https://covid19.moh.gov.om/>), there have been 160,018 confirmed cases of COVID-19 in Oman, with 1681 deaths from 3 January 2020 to 2 April 2021.

Governments, health-related organisations and virology experts in several countries have attempted to improve our understanding of COVID-19 and its relationships with demographic, environmental, and socioeconomic variables (Ahasan *et al.*, 2020; Arora *et al.*, 2020; Bagal *et al.*, 2020; Franch-Pardo *et al.*, 2020; Gupta *et al.*, 2020; Mollalo *et al.*, 2020; Pourghasemi *et al.*, 2020; Pramanik *et al.*, 2020). In India, Roy *et al.* (2020) implemented geospatial analysis using autoregressive integrated moving average methods. Specifically, geographic information systems (GIS) can support risk analysis by employing epidemiological, ecological, climate and socioeconomic data (Nakada & Urban, 2020; Shakil *et al.*, 2020). Spatial statistics, such as local bivariate rela-



tionship (LBR) analysis (Song *et al.*, 2020), geographically weighted regression (GWR) (Iyanda *et al.*, 2020), ordinary least squares (OLS) (Sun *et al.*, 2020a) and generalised linear modelling (GLM) (Liu *et al.*, 2020a) have been widely used to detect areas at risk for COVID-19. Various studies have used existing social susceptibility index data and adapted methods to study the spatial trend of COVID-19 emergence across different sociocultural environments (Acharya & Porwal, 2020; Karaye & Horney, 2020). Numerous studies have verified the dependency between the geographic distribution of socioeconomic factors and the spatiotemporal evolution of the disease using spatiotemporal regression models (Bray *et al.*, 2020; Kadi & Khelifaoui, 2020).

The spatial distribution of the COVID-19 outbreaks in Oman has varied considerably from one wilayat (district) to another, thereby posing further difficulties to monitoring the COVID-19 spreading in the country on a detailed spatial scale (Al-Kindi *et al.*, 2020). Only few studies of COVID-19 spatial modelling have been carried out in Oman (Mansour *et al.*, 2020), which demands further research focused on the socio-economic and environmental impact of the infection, which may have a significant effect on this pandemic as outlined by (Bashir *et al.*, 2020). We believe that socioeconomic factors are vital for understanding of its distribution and

spread in Oman (Khalatbari-Soltani *et al.*, 2020; Lin *et al.*, 2020) and that their determination can play a crucial role for resource allocation.

The main objective of this study was to analyse the demographic and socioeconomic variables affecting the COVID-19 infections in Oman. We studied first the connection between these infections and a single known factor to build a relationship model and continued by developing a more detailed predictive model to determine the relative impact of various candidate socioeconomic factors.

## Materials and methods

### Study area

Oman is situated between latitudes 16°40' and 20°20'N and longitudes 51°50' and 59°40'E in the south-eastern corner of the Arabian Peninsula (Figure 1). The total area is 309,500 km<sup>2</sup>, with a shoreline extending about 165 km from the Strait of Hormuz in the North to Yemen in the South. The country is divided into 11 governorates and 61 districts (wilayats) and has a population of

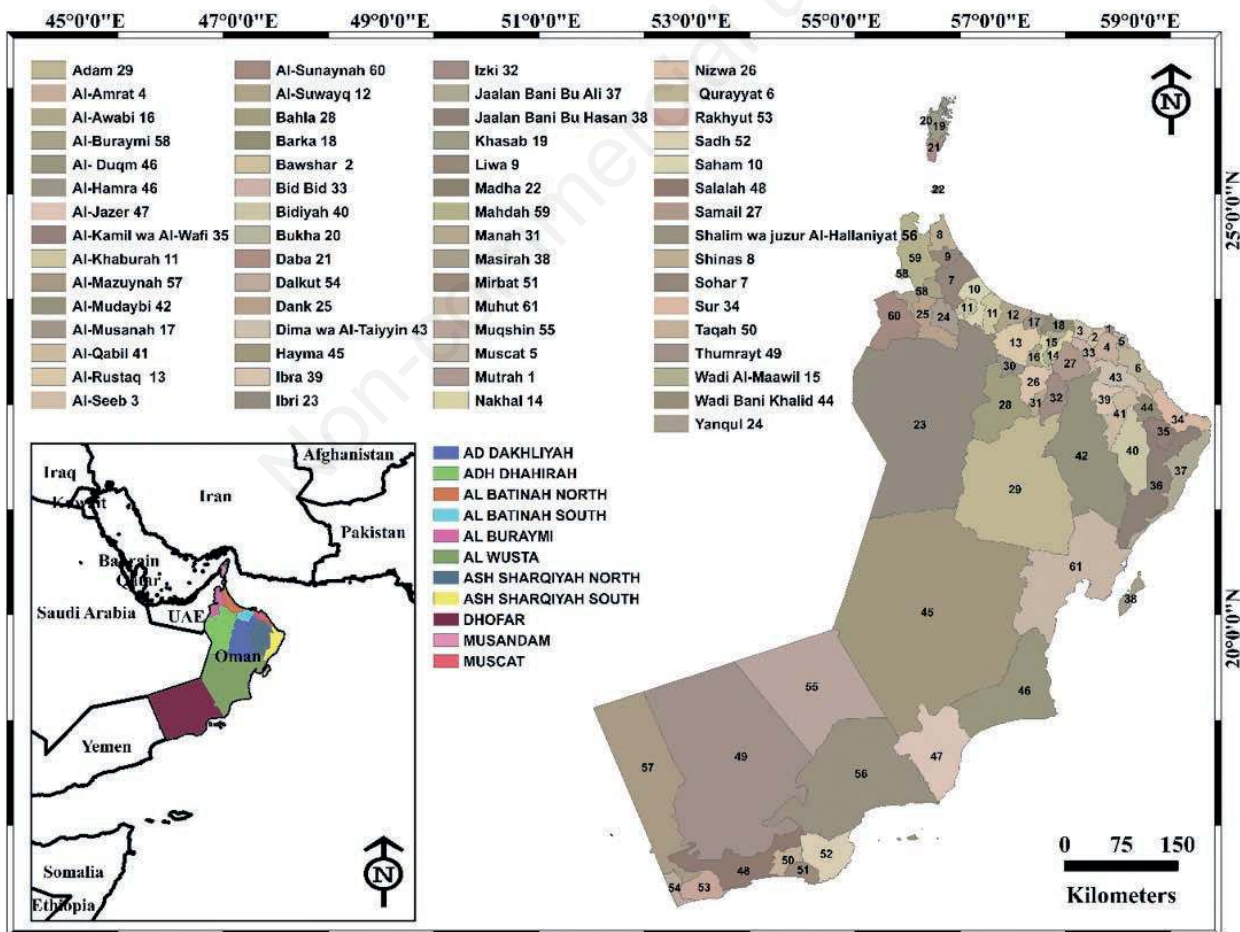


Figure 1. Area of study, including Oman’s location and the distribution in the study area of the 11 governorates and 61 wilayats.

4,617,927 people according to a 2020 national census (<https://ncsi.gov.om/>). With around 253,000 people, As-Seeb is the most populated wilayat (9.5% of the total Omani population), while the largest population density is registered in the Muscat Governorate (355.4/km<sup>2</sup>). The Bowsher Wilayat has the largest expatriate population (356,000 people), which amounts to 78% of its total population (<https://data.gov.om/>).

### COVID-19 data

Data about Omani and non-Omani COVID-19 cases for population groups aged 0-14, 15-59 and  $\geq 60$  years were obtained from the MOH and quantified at the sub-national level. The data for the period April 30 to August 30, 2020 were selected as base period since it corresponds to the time when the prevalence of the infection was the most pronounced in Oman.

To obtain real-time COVID-19 data, Tarassud+ App (TA+), a prototype system established by the MOH was used. This App shows the COVID-19 country status plus recommendations, personal information, declarations and other metadata related to COVID-19. TA+ is updated daily with information on deaths and existing, new and recovered cases (Ming *et al.*, 2020; Waheed & Shafi, 2020). While TA+ has an analytics platform to extract updated information, the MOH is responsible for all official announcements (Al Fannah *et al.*, 2020).

### Demographic and socioeconomic data

Data of the variables collected from different national authorities in Oman are displayed in Table 1. Healthcare resources including the number of physicians per wilayat, including the TA+ data, were collected from the MOH. Other datasets were obtained directly from local and national authorities, mainly the National Center for Statistics and Information (NCSI), while others were computed at national and wilayat levels using a GIS approach. Specifically, the Enrich tool in ArcGIS Pro 2.7 (ESRI, Redlands, CA, USA) was used to generate the geospatial dataset. The purchasing power index per capita was computed using data for disposable income. These enriched geospatial and population data created via the Enrich tool were validated against data collected from the NCSI ([www.ncsi.gov.om/](http://www.ncsi.gov.om/)) (Table 1).

Overall, 24 demographic and socioeconomic variables were obtained (Table 1). All geospatial datasets were projected onto the World Geodetic System (WGS) 1984 Universal Transverse Mercator (UTM) zone 40 North.

### Calculation of health risk and service metrics

#### Tobacco risk

As COVID-19 primarily is a respiratory disease (Gattinoni *et al.*, 2020), a tobacco index (TI) dataset was used to identify this risk. The data were obtained from the MOH and based on the World Health Organisation (WHO) STEPwise approach to

**Table 1. Variables and data sources used in this study and their main attributes.**

Socioeconomic variable	Definition	Format	Tool	Source
PD	Population density	Polygon/GIS shapefile	Enrich	NCSI
TH	Total number of households	Polygon/GIS shapefile	Enrich	NCSI
AHS	Average household size	Polygon/GIS shapefile	Enrich	NCSI
SII-1	Spatial interaction index 1	GIS Network datasets	Spatial analysis	NCSI
SII-2	Spatial interaction index 2	GIS database	Gravity model	NCSI
SCI	Social charity index	Excel sheets	MSD	MSD
TPPI	Total purchasing power index	Polygon/GIS shapefile	Enrich	NCSI
PPI	Purchasing power index (%)	Polygon/GIS shapefile	Enrich	NCSI
RCD	Relative cases distance	Polygon/GIS shapefile		NCSI
TP	Total population	Polygon/GIS shapefile		NCSI
N-OC (0-14)	Non- Omani cases (0-14)	Excel sheets	TA+	MOH
N-OC (15-59)	Non- Omani cases (15-59)	Excel sheets	TA+	MOH
N-OC (60+)	Non- Omani (60+)	Excel sheets	TA+	MOH
OC (0-14)	Omani cases (0-14)	Excel sheets	TA+	MOH
OC (15-59)	Omani cases (15-59)	Excel sheets	TA+	MOH
OC (60+)	Omani cases (60+)	Excel sheets	TA+	MOH
OFC	Omani female cases	Excel sheets	TA+	MOH
OMC	Omani male cases	Excel sheets	TA+	MOH
TO	Total Omani	Excel sheets	TA+	MOH
TN-O	Total non-Omani	Excel sheets	TA+	MOH
TI	Tobacco index	Excel sheets	TA+	MOH
Doctors	Doctors (%)	Excel sheets	TA+	MOH
Nurses	Nurses (%)	Excel sheets	TA+	MOH
HB	Hospital beds (no.)	Excel sheets	TA+	MOH
COVID-19	Monthly COVID-19 data	Mobile phone	TA+	MOH

NCSI, National Center for Statistics and Information; MOH, Ministry of Health; MSD, Ministry of Social Development.

Surveillance (STEPS), a standardized surveillance tool through which countries can collect, analyse and disseminate core information on non-communicable diseases.

### Spatial interaction index

Flow estimations (also referred to as spatial interactions), between locations is of interest for transport. These flows enable the demand for transport services to be assessed (existing or potential). They include mobility modes, such as work trips, relocation, tourism, the use of public facilities, information or resource transfer, retail markets, international exchange and freight distribution. In this study, the spatial layer of interaction was generated from road connectivity based on ArcGIS analysis. While this approach offers a basis for location communication, it does not allow detailed studies of the changing spatial interaction patterns that occur due to fear, closures, policy changes, *etc.* The most commonly used spatial interaction models, the Gravity Index (GI) and the Network Identify tool in ArcGIS 10.7, were used to estimate a spatial interaction index based and road connectivity network as potential for human movement (Song *et al.*, 2019). The resulting layer contains a GI field, in which higher values correspond to higher weights of spatial interaction and influence (Table 1).

### Relative case distance index

This is based on the chance of exposure to COVID-19 and displayed as a map exhibiting various levels of this risk. The relative case distance area represents a view from the middle of each municipality to all close areas with COVID-19 cases. To support each person's integrity, case identities are not revealed, only the nearest road intersection to at-risk areas. We measured the distance to cases from the municipality polygon centroids rather than from population centres. To accomplish this, we first created polygons for each wilayat and included random points in them as representation of the percentage of the total population therein. Then, we computed the mean centres for all these random points associated with the polygons and took all municipalities close to a large number of known cases as being at increased risk of exposure (Table 1).

### Impact on COVID-19 prevalence of demographic and socioeconomic variables

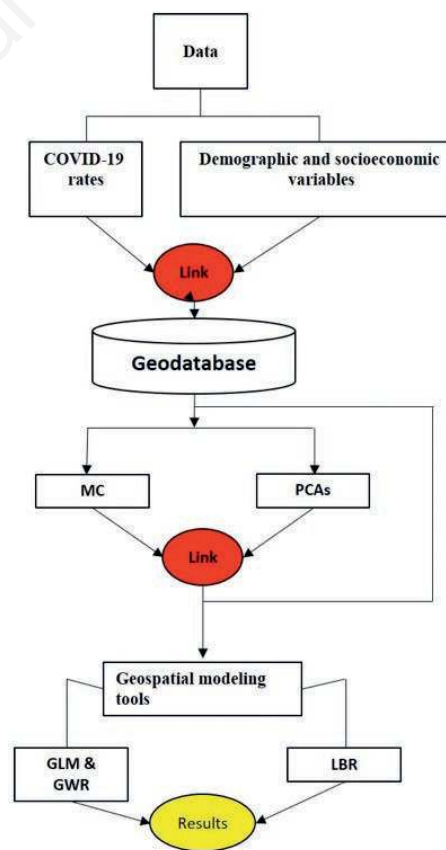
A flowchart displayed as shown in Figure 2 shows how we modelled the spatial relationships between COVID-19 prevalence and the 24 demographic and socioeconomic variables that were substantially multicollinear. A regression model examining the effect of these variables on COVID-19 was constructed to ensure their independency. We first turned all the correlated factors into artificial, uncorrelated principal components (PCs), the analysis of which (PCA) would clarify the maximum amount of original data variance (Abdi & Williams, 2010) while still retaining the most significant information about the data structure.

To solve this, we applied what is called data rotation using the VARIMAX method (<https://www.sciencedirect.com/topics/nursing-and-health-professions/varimax-rotation>). The outcome revealed six components that accounted for 90% of the variance, but for simplicity and ease of interpretation, we only kept the first three components (PC1, PC2 and PC3), which explained 65% of the variance Table S1 (Appendix A).

The LBR approach was used to quantify the strength of connotation amid the COVID-19 disease rate and the various demographic and socioeconomic variables, thereby developing a relationship model to determine if the relationships found were consistent across

the study area. This can be done by randomising the data, since the presence of significant, meaningful correlations between two variables investigated in a data subset would increase the entropy considerably (Guo, 2010).

Multivariate clustering (MC) utilises unsupervised machine learning techniques to detect natural data clusters (Asante & Kreamer, 2015; Pászto *et al.*, 2020). MC is a toolset that can be used to identify and map the locations of statistically significant spatial outliers (Peeters *et al.*, 2015). As it discovers spatial patterns (Garcia, 2020; Nipperess *et al.*, 2008), the MC method was used to understand how the COVID-19 outbreaks occurred and spread among all the wilayats in Oman from April to August 2020. The clustering efficacy was calculated applying the Calinski Harabsz pseudo-F statistic to investigate the ratio of variation between clusters (Belvisi *et al.*, 2020). By default, the MC method uses a K-means algorithm (Singhal & Seborg, 2005). The GLM and GWR techniques were used to detect and model the associations between the number of infected people and the different socioeconomic variables (PC1, PC2 and PC3) across the country (Figure 2). The GLM technique offers a global model for understanding or predicting early incidence and infection rates and generates an individual regression equation to describe such a process. GWR, on the other hand, offers a local model for predicting and understanding early incidence and infection rates by implementing a regression equation (Sarra & Nissi, 2016). GLM possesses OLS regression functionality with three models



**Figure 2.** Flowchart of the modelling of the spatial relationships between COVID-19 and demographic and socioeconomic variables. MC, multivariate clustering; PCA, principal component analysis; GLM, generalized linear model; GWR, geographically weighted regression; LBR, local bivariate relationship.

(Gaussian, logistic and Poisson) that permits its use for a broader approach (Yang *et al.*, 2013). Fair model use, however, depends on the existence of the dependent variable. GWR generates an equation for each characteristic in the dataset, calibrating each one by utilising the target features and neighbours. It includes spatial non-stationary or spatially varying relationships in the regression by allowing the coefficient to change with location (Windle *et al.*, 2010). Similar to GLM, GWR offers three types of regression models: count, binary and continuous. The non-Gaussian (continuous) zero-inflated distribution of the outcome variables (PC1, PC2 and PC3) and the discretization method allow Poisson regression to model the variable. Geographically weighted Poisson regression (GWPR) is useful in the study of disease distributions as it can be compared with the global OLS model results to examine which approach is the most appropriate. The spatial autocorrelation tool was used to ensure that the residuals (over- and under-predictors) were random across the study area. Golden search (Wheeler & Páez, 2010) calculates the distance or number of neighbours with the lowest corrected Akaike information criterion (AIC<sub>c</sub>). This tool and the golden section search method (Li *et al.*, 2019) was used to identify an optimal distance or number of neighbours on the basis of the characteristics of the data. The GWR model offers two kernel choices in the local weighting scheme parameter: Bi-square and Gaussian. The former truncates the kernel using a distance or several neighbours, while the latter assigns a weight of 1 to the regression feature (feature *i*). The importance of the neighbouring elements (*j* features) decreases gradually as the regression feature distance increases. In this study, the bi-square was used to run the model. To visualise how the associations between the socioeconomic variables and the COVID-19 infection levels varied across space, GWPR was used to create coefficient maps (Li *et al.*, 2013; Yu & Peng, 2019). GWPR is computed as follows:

$$y_i \sim \text{Poisson}[N_i \exp(\beta_0(u_i, v_i) + \sum_k \beta_k(u_i, v_i)x_{k,i})] \quad (1)$$

where  $\beta_0(u_i, v_i)$  is the intercept variable explicit to site *i*;  $\beta_k(u_i, v_i)$  the coefficient of predictor variables  $x_k$  at location *i*;  $N_i$  the inverse variable at each location; and  $(u_i, v_i)$  the *x-y* coordinate at each site. Accordingly, to validate the hypothesis and its applicability to the investigation of the regional differences, GLM was adopted also to scrutinise the connection amid the COVID-19 infection rate and different combinations of socioeconomic variables (PC1, PC2 and PC3) of two governorates: Muscat and Al-Batinah North. All analyses were carried out using ArcGIS Pro version 2.7.

## Results

### Local bivariate relationships

The direction and strength of correlation amid the COVID-19 infection rate and the different demographic and socioeconomic variables in Oman are illustrated in Figure 3, while the values of the adjusted  $R^2$  (Adj  $R^2$ ) are listed in Table 2. As indicated, positive significant correlations were observed between COVID-19 prevalence and 14 of the 24 variables investigated (Table 2). In comparison, significantly negative spatial connotations were detected between the COVID-19 infection rate and the relative cases distance (Adj  $R^2 = -60.66$ ). A negative but not significant relationships were identified amid the COVID-19 infection rate and percentages of doctors ( $P=0.85$ ) and nurses ( $P=0.57$ ). No relationship was detected between the COVID-19 infection rate and the tobacco index, number of hospital beds, percent average household size,

**Table 2. Summary of the local bivariate associations between COVID-19 infection rate and individual socioeconomic variables.**

Variable	Adj $R^2$ positive linear (%)	Adj $R^2$ negative linear (%)	Concave (%)	Convex (%)	Mean entropy (%)	P
1-Total population	80.33*	0	19.67	0	1.4227	0.0050
2-Total Omani	85.25*	0	0	14.75	1.2652	0.0050
3- Total non-Omani	75.41*	0	24.59	0	1.4434	0.0050
4-Tobacco index	0	0	0	0	1.7573	0.8300
5-Spatial interaction index 1	54	0	0	0	1.4204	0.1650
8-Hospital beds	0	0	0	0	1.6942	0.0654
9-Population density	73.77*	0	0	0	1.332	0.0050
10-Total hospital	88.52*	0	11.48	0	1.2376	0.0052
11-Average households size (%)	0	0	0	0	1.1669	0.2017
12-Spatial interaction index 2	73.77*	0	1.64	0	1.2974	0.0050
13-Social charity index	44.26*	0	0	0	1.3695	0.0350
14-Total purchasing power index	88.52*	0	11.48	0	1.2496	0.0050
15-Purchasing power index (%)	0	0	0	0	1.3596	0.0600
16-Relative cases distance	0	60.66*	39.34	0	1.2214	0.0050
17-Non-Oman cases aged (0-14)	100*	0	0	0	1.2186	0.0050
18-Non-Oman cases aged (15-59)	77.05*	0	22.95	0	1.2904	0.0050
19-Non-Omani cases aged ( $\geq 60$ )	86.00*	0	0	0	1.2341	0.0050
20-Omani cases aged (0-14)	90.00*	0	0	9	1.2224	0.0050
21-Omani cases aged (15-59)	90.16*	0	0	9.84	1.0542	0.005
22-Omani cases aged ( $\geq 60$ )	95.08*	0	0	4.9	1.1667	0.0050

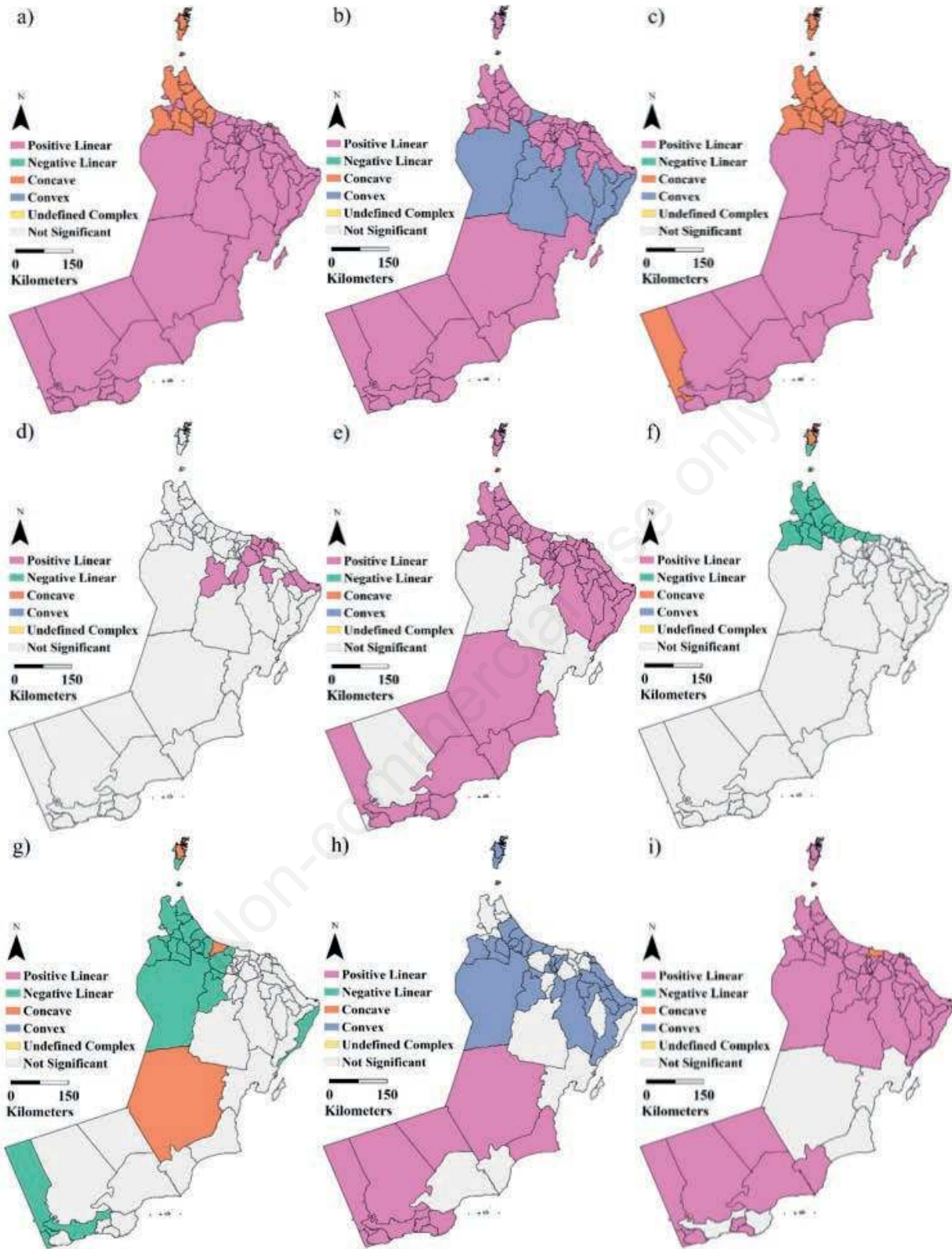


Figure 3. Maps showing the types of association between the COVID-19 infection rate and individual socioeconomic variables. a) total population; b) total Omani population; c) total non-Omani population; d) tobacco index; e) spatial interaction index-1; f) doctors %; g) nurses %; h) number of hospital beds; i) population density.

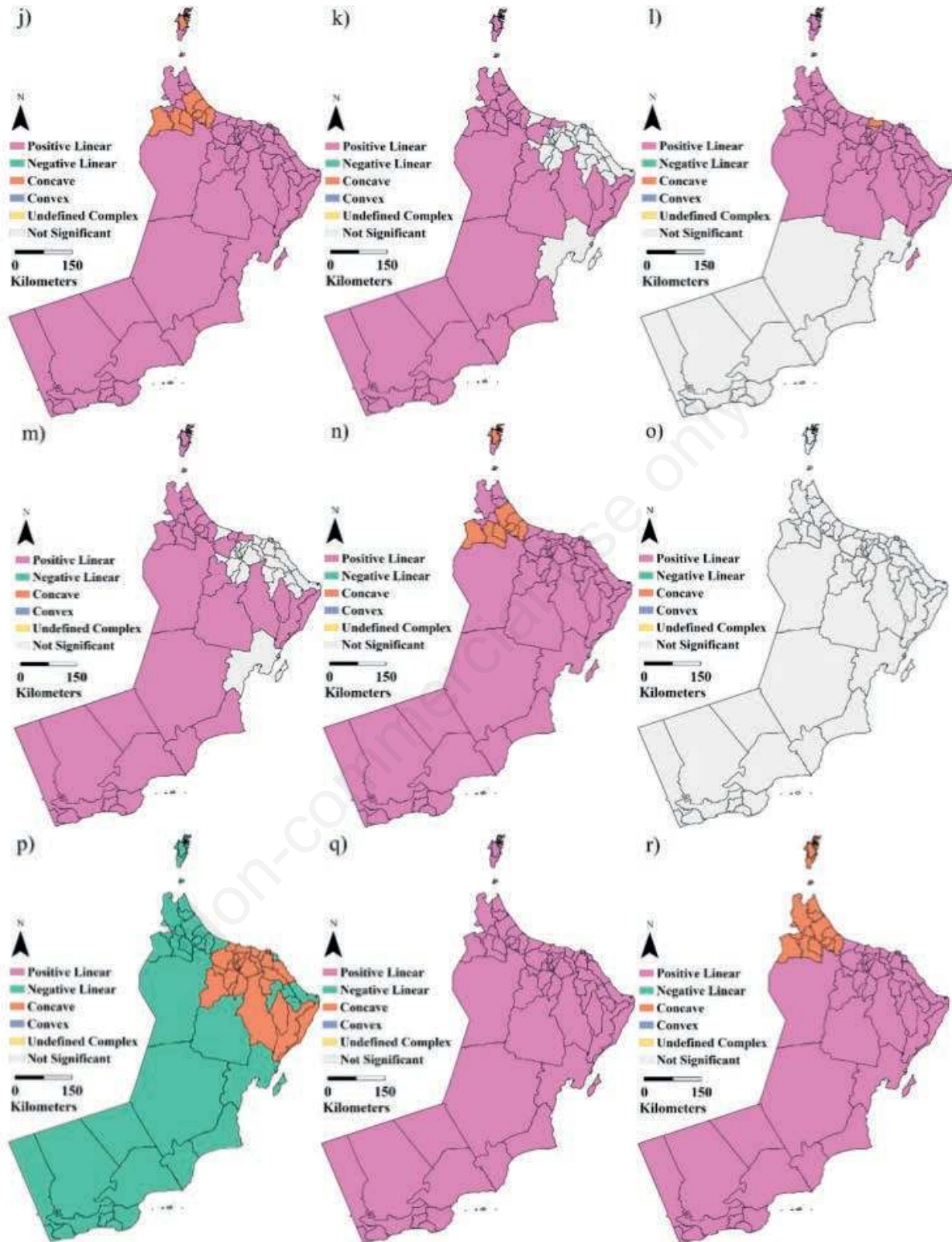


Figure 3. Maps showing the types of association between the COVID-19 infection rate and individual socioeconomic variables. j) total number of households; k) average household size %; l) spatial interaction index-2; m) social charity index; n) total purchasing power index; o) purchasing power index; p) relative cases distance. q) non-Omani cases aged 0-14; r) non-Omani cases aged 15-59.

percent purchasing power index and male and female Omani case variables (Figure 3 and Table 2). The coefficients maps produced throughout the LBR model revealed where the variable had the most significant impact of the regression across the country (Appendix B, S1).

**Principal component analysis**

Figure 4 illustrates the spatial pattern of the first three PCs. For PC1, the highest loads were located in the wilayats, particularly Mutrah, Bowsher and As-Seeb in Muscat Governorate and Salahah in Dhofar Governorate (Figure 4a), while the lowest ones were distributed in coastal wilayats such as Shinas, Saham and Al-Suwayq in Al-Batinah North Governorate and internally in wilayat Bid-Bid in Al-Dakhliyah Governorate (Figure 4a). For PC2, the highest loads were concentrated in the most populated districts, especially

in As-Seeb in Muscat Governorate, while the lowest were found in Khasab in Musandam Governorate (Figure 4b). For PC3, the highest were in Muscat in the Muscat Governorate and in Huma in Al-Wusta Governorate. In contrast, the lowest were seen in many wilayats, including Al-Sunaynah in Al-Buraymi Governorate, As-Seeb in Muscat Governorate and Al-Mazuynah and Shalim Wa Juzur Al-Hallaniyat in Dhofar Governorate (Figure 4c).

The highest eigenvectors of the six components are summarised in Table S1 (Appendix A). The highest factor loads for PC1 were 0.94 for the non-Omani cases aged  $\geq 60$ , 0.92 for the non-Omani cases aged 0-14, 0.92 for the spatial interaction index-2, 0.91 for the population density, 0.90 for the total non-Omani population, 0.89 for the non-Omani cases aged 15-59, 0.73 for the number of hospital beds, 0.72 for the total number of households, 0.70 for the Total purchasing power and 0.70 for the purchasing

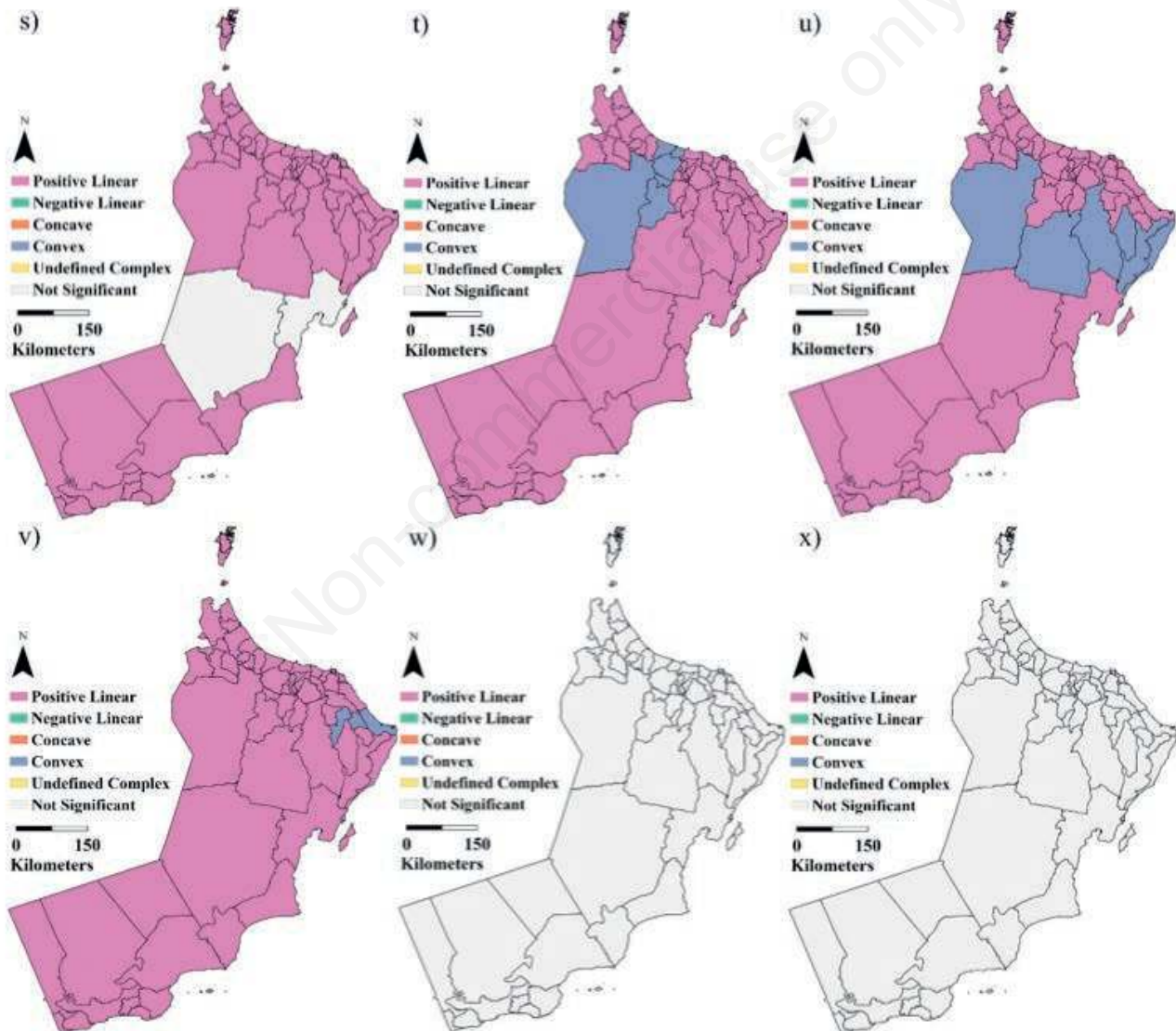
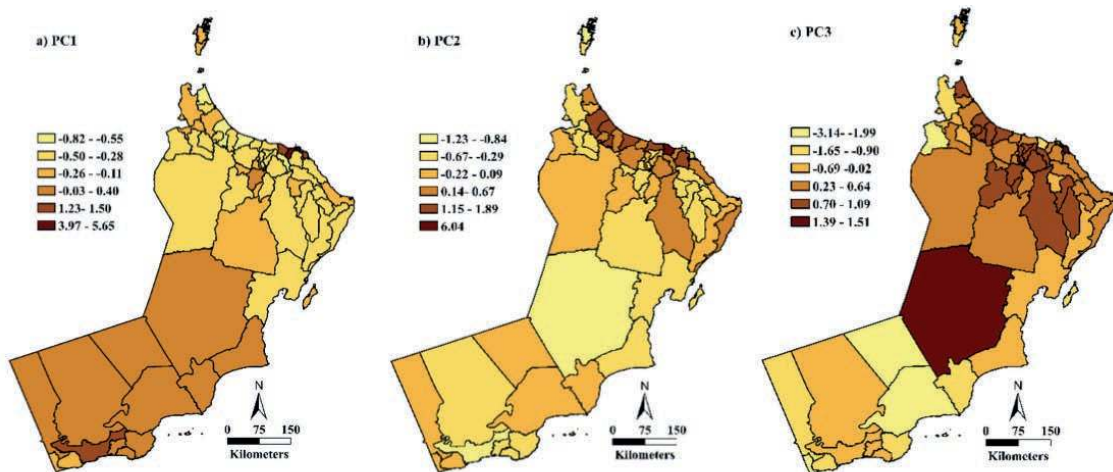


Figure 3. Maps showing the types of association between the COVID-19 infection rate and individual socioeconomic variables. s) non-Omani cases aged  $\geq 60$ ; t) Omani cases aged 0-14; u) Omani cases aged 15-59; v) Omani cases aged  $\geq 60$ ; w) female Omani cases; and x) male Omani cases.





**Figure 4.** Score value maps for each variable based on the illustrated values highlighting the spatial variances for each factor based on principal component analyses. a) PC1, b) PC2 and c) PC3.

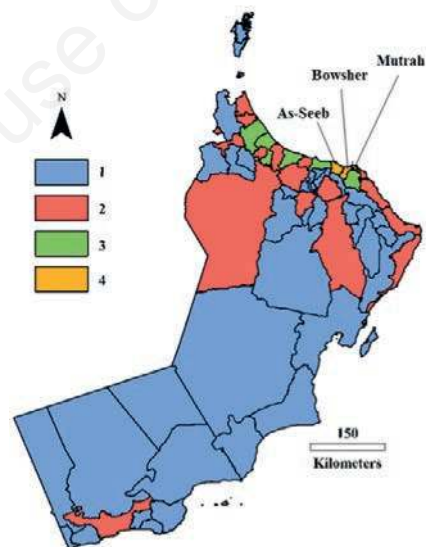
power per mile. The highest factor loads for PC2 were 0.93 for the Omani cases aged 15-59, 0.95 for the total Omani cases, 0.919 for the Omani cases aged 0-14 and 0.86 for the Omani cases aged  $\geq 60$ .

### Multivariate clustering

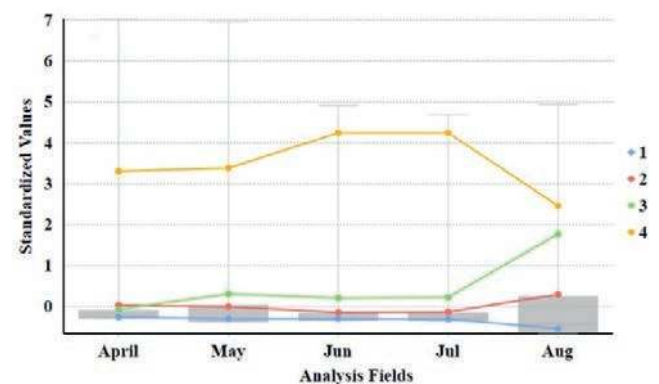
The MC-based results based on K-means and K-medoids analysis are summarised in Table S2 (Appendix A). Oman's districts are grouped according to the monthly registered COVID-19 cases from April to August 2020. The highest  $R^2$  value was found for the July variable. The  $R^2$  value indicates how much variance was retained after the clustering phase in the actual rate data (Vishwakarma *et al.*, 2017). Out of the 61 wilayats examined, three (0.3%; 925 km<sup>2</sup> total area) were identified as high-risk ones; five (2.0%; 6,308 km<sup>2</sup> total area) as medium-risk ones; 15 (21.3%; 65,858 km<sup>2</sup> total area) as being at low risk; and 38 (75.0%; 232,409 km<sup>2</sup> total area) were identified as being at very low risk (Figure 5). The wilayats depicted as cluster 4 in Figure 6 were found to be high spatial clusters. The COVID-19 spatiotemporal and spatial patterns were similar during much of the study period. By applying MC analysis, we were able to identify wilayats in high-risk areas (cluster 4) based on the monthly records from April 30 to August 30, 2020 (Figure 6). These wilayats were As-Seeb, Muttrah and Bowsher. The boxplots created for each cluster and variable show how the cluster values were related to those of other clusters (Figure 7). The map and the parallel boxplot charts in Figures 5, 6 and 7, respectively, summarise the areas with various risk levels identified in the different locations (wilayats) for different epidemic periods (monthly) using the MC tool. The boxplot shows the connections between the multivariate clustering outcome and the PCA results. These connections were used to run global and local regression models (Figure 8).

### The generalized linear model

The variance inflation factor (VIF) values were  $<1.1$  indicating that they were all under the ESRI-defined threshold of 7.5; thus, the independent variable distributions were not redundant or biased by multicollinearity. Note that we could not obtain good VIF results before applying PCA with the data used in this study. The Joint Wald and Koenker (probability  $>$ chi-square) statistics (Pick & Nishida, 2015) were significant, with P-values of approx-



**Figure 5.** The four spatial multivariate clusters based on the data regarding the confirmed case attributes.



**Figure 6.** Boxplots summarising the four multivariate clusters and their associated variables (April-August 2020).

imately 0. The significant positive association between the COVID-19 infection rate and the socioeconomic variables (PC1, PC2 and PC3) are illustrated in Table 3. GLM had an  $AIC_c$  value  $>25,174$  explaining 86% of the goodness of fit for the associations between the COVID-19 infection rate and the socioeconomic factors. The generalized linear regression (GLR) model results were also appraised for spatial autocorrelation showing a z-score of  $-0.817$  and Moran's  $I = -0.075$  at  $P=0.41$ . This indicated that the model's resulting pattern was highly random and as such, crucial independent variables were present. The residuals of the Poisson regression GLM exhibited negative spatial autocorrelations.

### Geographically weighted regression

The GWR model had an  $AIC_c$  value  $>13,379$  which explained 93% of the goodness of fit for the associations between the COVID-19 disease rate and the socioeconomic variables. The

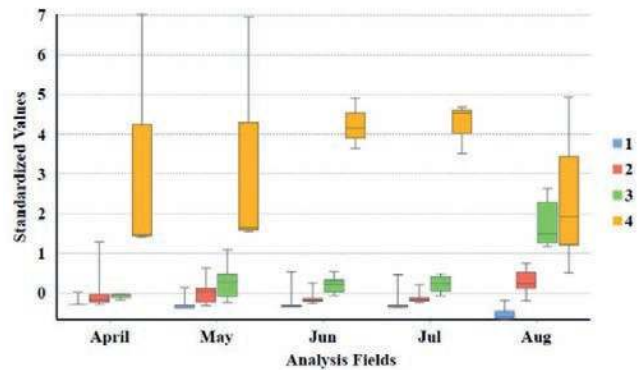


Figure 7. Parallel boxplot display of each multivariate cluster and corresponding variables.

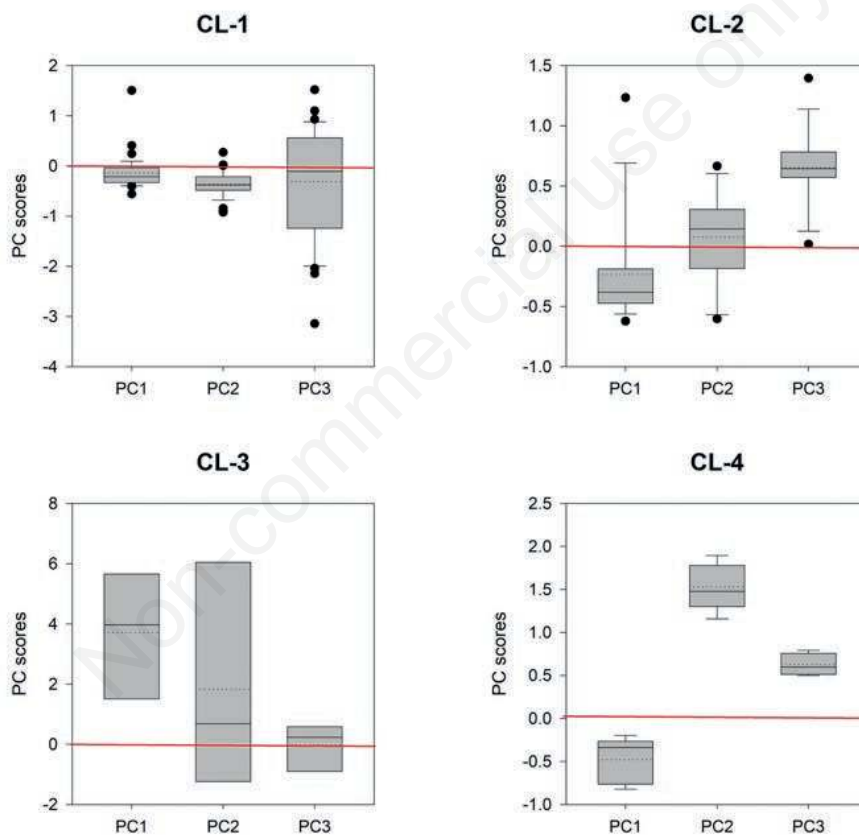


Figure 8. Differences of principal component (PC) scores as a function of the four main spatial clusters (CL).

Table 3. Statistically significant variables based on generalized linear regression.

Variable	Coefficient	t-statistic	z-statistic	Probability	VIF
Intercept	6.3136	0.0072	880	0.0001*	1.00
PC1	0.5224	0.0018	289	0.0001*	1.00
PC2	0.5986	0.0024	252	0.0001*	1.00
PC3	1.0018	0.0084	118	0.0001*	1.00

Coefficient, strength and type of relationship between each explanatory variable and the dependent variable. VIF, variance inflation factor where values  $>7.5$  indicates redundancy among the explanatory variables.

GWR model had a lower  $AIC_c$  value implying model complexity and a higher  $R^2$ , which gave a better fit than GLR for the observed results (Table 4). Therefore, the extent of GWR calculations showed the non-stationery of the relations between the socioeconomic variables and the COVID-19 infection rate in the study area. The spatial autocorrelation results reported a  $z$ -score of  $-0.5044$  and Moran's  $I = -0.0530$  at  $P=0.6139$ . This indicates that the over and underpredictions resemble those of the GLR model; they thus appear random and suggest a properly specified model. Figure 9 shows the bivariate colours of the quantitative relationship of the spatial associations between the fixed variable (COVID-19 infection rate) and the exploratory factors (PC1, PC2 and PC3). High associations between the COVID-19 infection rate and PC1 were found in the following wilayats: Bowsher, Mutrah, As-Seeb and Muscat in Muscat Governorate; Nizwa in Al-Dakhliyah Governorate; Salalah in the Dhofar Governorate and Al-Duqm in Wusta Governorate. In contrast, a low association between COVID-19 infection rate and PC1 was found in Daba in Musandam Governorate and Al-Sunaynah in Al-Brahimi Governorate (Figure 8). Figure 10 exhibits the local coefficient assessments for each of the explanatory variables, presented from high to low. The positive significant associations between COVID-19 infection rate and PC1 varied substantially, and the coefficients ranged from 0.1327 to 0.2374 (Figure 10a). Positive associations were also identified between COVID-19 infection rate and PC2, with coefficients ranging from 0.330 to 0.4017 (Figure 10b). High relational values were observed in the most northern portion of the study area, while the lowest values were found in the southern part. Therefore, positive associations were detected between the infection rate and PC3, with coefficients ranging from 0.528 to 0.551 (Figure 10c). The goal of these maps was to examine the spatial changes of the disease occurrence based on socioeconomic variables. The distribution of locally weighted  $R^2$  between the perceived, equipped values and the deviance residuals displayed,

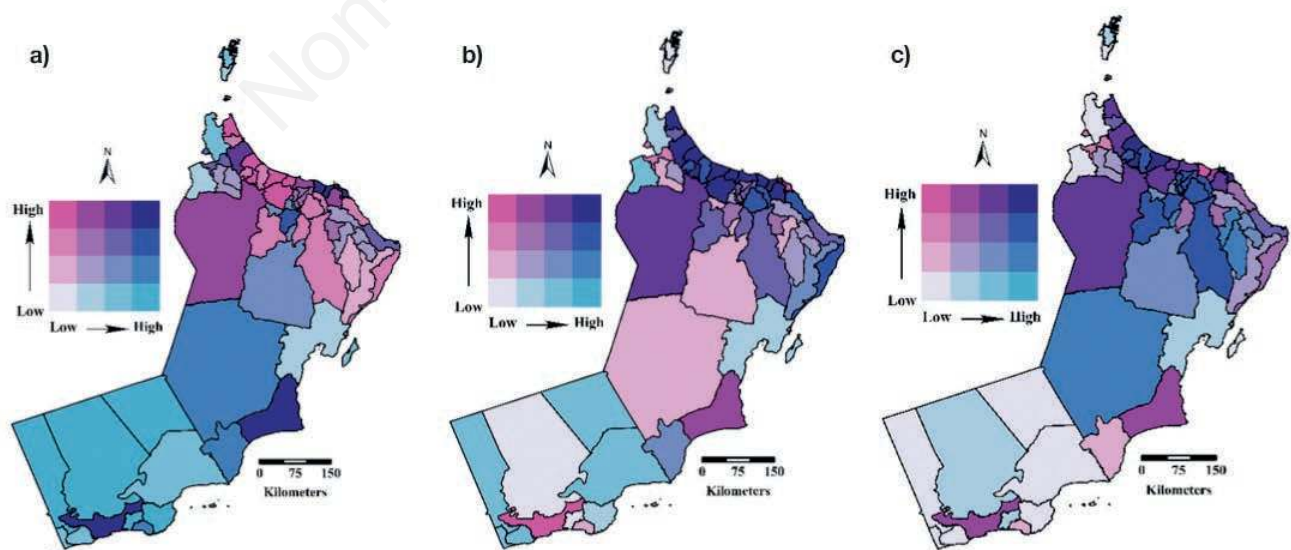
which prototype had a higher symmetry of dependent-variable variance estimated by the regression model (Figure 11a and b). The locally adapted strengths changed spatially, as seen in the map of local  $R^2$  (Figure 11a), which indicates that the high-occurrence areas were better suited than the low-occurrence ones. The local  $R^2$  values, which suggest the impacts of the elected socioeconomic variables on the COVID-19 cases, were perceived to be the highest, with the  $R^2$  ranging from 0.92 to 0.93 in As-Seeb, Bowsher, Mutrah, Muscat, Al-Amrat, Barka, Bid-Bid, Nakhal, Wadi Al-Maawil, Al-Awabi, Al-Musanah, Al-Suwayq, Al-Rustaq, Qurayyat, Wadi Bani Khalid, Sur, Al-Qabil, Izki, Nizwa, Manah and Al-Hamra (Figure 11). The statistics presented in the residuals map show that the residuals followed a normal distribution, with a mean of  $-2.411$  and a standard deviation (SD) of 14.71. Most of the locations had standardised residuals close to 0. The GWR model made fewer over- and underestimations (with the standardised residuals more than one SD away) than GLM. On the basis of the results, we utilised GWR to assemble a predictive map for classifying areas with potential future outbreaks (Figure 11c).

Although global and local regression were used to examine the associations between different combinations of the socioeconomic factors and the COVID-19 infection rate for the whole Oman, GLM was also used to model the regional differences (specific model for each governorate). The global model explained 95% and

**Table 4. Comparison of local and global models.**

Model	Deviance explained (%)	$AIC_c$
Generalized linear Poisson regression	0.86	25174
Geographically weighted Poisson regression	0.93	13379

$AIC_c$ , corrected Akaike information criterion.



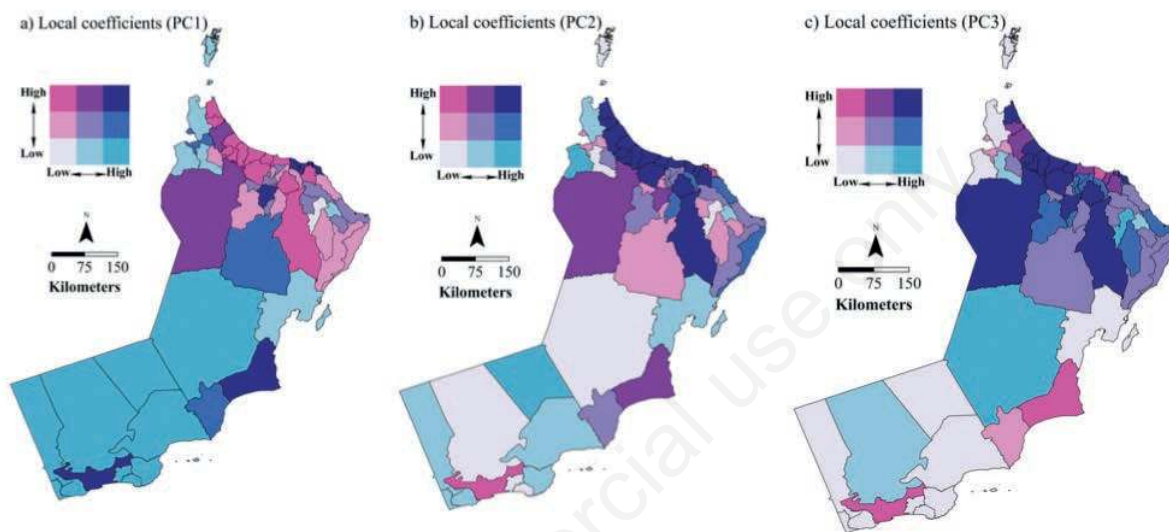
**Figure 9. Local  $R^2$  relationships between the COVID-19 infection rate and the three main components. a) COVID-19 infection rate and PC1; b) COVID-19 infection rate and PC2; c) COVID-19 infection rate and PC3.**

99% of the impacts of the socioeconomic variables on the COVID-19 infection levels of the Muscat and Al-Batinah North governorates, respectively (Table 5). The coefficient values displayed in Appendix A, S3 show that PC1 and PC2 had strong correlations with the COVID-19 infection levels. Figure 12 shows the predictive maps for classifying areas with potential future outbreaks of the Muscat and Al-Batinah North governorates, respectively.

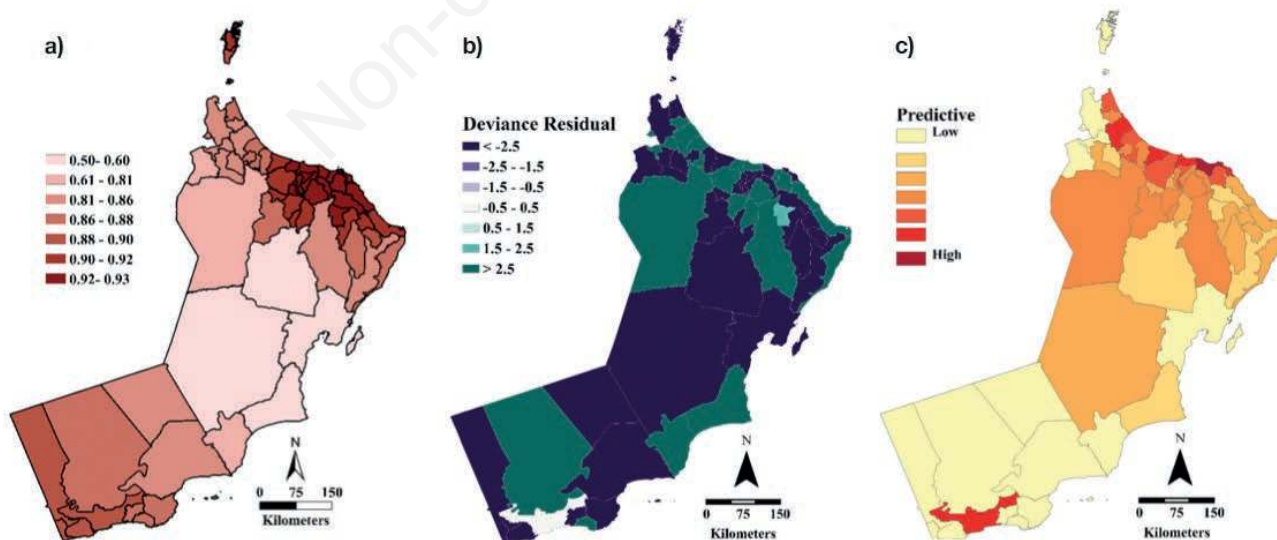
**Table 5. Summary of hypothesis validation and its applicability to global and local models of regional difference.**

Region	Deviance explained	AIC <sub>c</sub>
Muscat	0.95	6711
Al-Batinah North	0.99	107

AIC<sub>c</sub>, corrected Akaike information criterion.



**Figure 10. Local coefficients for the geographically weighted regression model. a) coefficients between COVID-19 infection rate and principal component (PC) 1 (ranging from 0.1327 to 0.2374); b) coefficients between COVID-19 infection rate and PC2 (ranging from 0.330 to 0.4017); c) coefficients between COVID-19 infection rate and PC3 (ranging from 0.528 to 0.551).**



**Figure 11. COVID-19 infection rate in the study area including deviance and predictive information. a) Local percentage deviance in the geographically weighted regression based on a prediction model of COVID-19 infection rate in the study area; b) Deviance residuals showing which model had a higher proportion of dependent-variable variance accounted for by the regression model; c) Predictive information on the possible COVID-19 outbreak areas.**

## Discussion

Although GWR, GLM and OLS are potent tools for modelling the relationships between such variables and predicting the geographic distribution of viruses (Pirdavani *et al.*, 2014; Su *et al.*, 2017), these models are affected mainly by multicollinearity (Bager *et al.*, 2017; Wheeler & Tiefelsdorf, 2005), and as such, PCA was applied in this study to reduce data dimensions, keeping only the most significant and non-correlated components.

The LBR results revealed that certain variables, age in particular, contributed significantly to the increased COVID-19 infection rate in Oman during the study period (Table 2). This is hardly surprising as it is known that younger people infected by COVID-19 generally have mild symptoms and therefore move around unhindered by the disease (Ludvigsson, 2020).

The large number of non-Omani residents had a meaningful influence on the COVID-19 infection rate in Oman. The expatriate population, numbering around 2/5 of the total population, largely live in more crowded situations than the nationals and have been found to have a comparatively higher rate of COVID-19 infection rates. In accordance with similar studies worldwide (Kadi & Khelfaoui, 2020; Sannigrahi *et al.*, 2020; Sun *et al.*, 2020b) our study discovered a positive association between the COVID-19 infection rate and variables, such as population density, number of households, spatial interaction and social charity in the study area (Figure 2 and Appendix B, S1).

The results of the multiple clustering model revealed that out of the 61 wilayats examined in this study only three were high risk and five at medium risk, while 15 were at low risk and as many as 38 were at a very low risk. The MC results also indicated that the clustering digressed in August (Appendix A, S2). One reason for this may be complete closure between the Sultanate governorates enacted between July 25 and August 8, 2020. A second reason may be related to the decision to prohibit movement in all the sultanate governorates from 9:00 p.m. to 5:00 A.M. from August 8 August

15, 2020. An earlier analysis showed that the geographical distribution of COVID-19 events in Oman varies greatly in time and space from one wilayat to another (Al-Kindi *et al.*, 2020).

LBR was used in this study to examine the correlations between the number of infected people and the socioeconomic variables for developing a correlation model. Based on this correlation model, we constructed a more complex predictive model for assessing the relative impacts of the candidate socioeconomic variables. While the GWR model was used to define a spatial model for the dependent and independent variables to represent a hedonic model with geographically varying weights, GLM was used to establish the relational attributes of explanatory variables globally. Both models were affected by multicollinearity, so we were unable to introduce or run more than 40% of the candidate explanatory variables before applying PCA.

Comparisons of regression models are essential to understand factors such as socioeconomic status that may impact the COVID-19 infection rate in the study area (Mollalo *et al.*, 2020). According to the GLM results, most of the introduced variables have the potential to perform a significant role in increasing or decreasing the COVID-19 infection rate in the study area. This approach is therefore useful for understanding what is happening in a location as some socioeconomic factors can have global impacts without spatial variation. The GWR model had an  $AIC_c$  value  $>2086$  and explained 92% of the goodness of fit. The overall  $AIC_c$  result for the GWR model was less than for GLM. The lower  $AIC_c$  statistic in the observed results indicates that the GWR model would be preferable (Tables 4 and 5). The GWR coefficients are also an effective means to visualise the spatial variation of the association between a descriptive variable and a target variable. In this comparison, the GWR model was found to be better suited to capturing spatial variations in the number of infected people in relation to socioeconomic status. The study results reveal positive coefficient relationships between COVID-19 infection rate and the socioeconomic factors in the study area (Figure 9). The coefficient maps

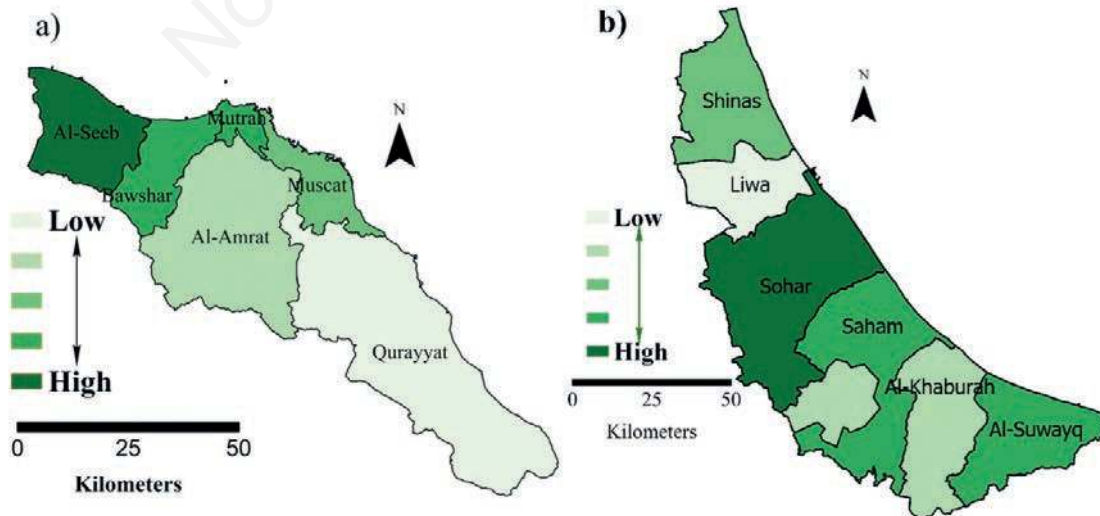


Figure 12. Predictive details of probable COVID-19 presence in the epidemic areas in two governorates. a) Muscat; b) Al-Batinah North.



created with GWR show where each variable had the strongest impact on the regression across the study area, ranging from strong positive relationships to weak positive relationships (Figure 10). The maps of standard residuals, predictive values and  $R^2$  values permitted us to recognize the confluence of factors that are most favourable to survival and prevent the growth of COVID-19 (Figure 11). These results suggest that Poisson regression GWR and GLM can address unobserved spatial non-stationary relationships.

Like various studies worldwide (Ahasan *et al.*, 2020; Asante & Mills, 2020; Cao *et al.*, 2020; Kwok *et al.*, 2020), our study proved that the spatial relationship tools have the power to discover associations amid COVID-19 infection rate and socioeconomic variables and predicting the future of COVID-19 at various scales. The models built in this study can be used against the spread of COVID-19 in Oman for various control management purposes.

Spatial modelling was used to detect relationships between COVID-19 and demographic and socioeconomic factors. GLM and GWR were adopted to detect the relationships between COVID-19 rates and demographic and socioeconomic variables in Oman. The GWR model showed an enhancement in model performance, as designated by lower  $AIC_c$  and goodness-of-fit compared with the GLM. The local  $R^2$  values indicated the influence of the elected socioeconomic factors on COVID-19 cases. Understanding the variables related to high COVID-19 rates can help authorities in Oman make predictions about future outbreaks and allocate prevention resources more effectively. However, it is noteworthy that data at a higher spatial resolution are needed to provide a more detailed spatial assessment of the response of the Omani community to COVID-19 under different environmental and socioeconomic conditions. Also, more sophisticated data analysis and management tools and techniques (*e.g.*, artificial intelligence and deep-learning algorithms) are needed to isolate the joint impact that can be introduced by some variables, allowing the relative importance of each independent variable to be defined at a more detailed spatial scale (*e.g.*, urban *vs.* rural areas). In this context, the use of the advanced techniques use here can be an avenue for future research to enhance current understanding of the spatial and temporal patterns of COVID-19 pandemic in Oman. Climatic variations and environmental variables may also perform critical roles in intensifying the expansion, development, and forgiveness of COVID-19. These prototypes might be worthy of forecasting the spatial distribution and densities of the number of infected people under predominating conditions at each season's commencement.

## Conclusions

Knowledge of the factors associated with elevated concentrations of COVID-19 would assist the Omani authorities to make predictions for potential pandemics and to carry out pandemic-preventive facilities. Spatial statistics provide powerful tools to detect relationships amid COVID-19 and demographic and socioeconomic factors. The use of advanced techniques can be an avenue for research work in the future to enhance current understanding of the spatial and temporal patterns of COVID-19 pandemic.

## References

- Abdi H, Williams LJ, 2010. Principal component analysis. *Wiley Interdiscip Rev Comput Stat* 2:433-59.
- Acharya R, Porwal A, 2020. A vulnerability index for the management of and response to the COVID-19 epidemic in India: an ecological study. *Lancet Glob Health* 8:e1142-51.
- Ahasan R, Alam MS, Chakraborty T, Hossain MM, 2020. Applications of GIS and geospatial analyses in COVID-19 research: a systematic review. *F1000Res* 9:1379.
- Al-Kindi KM, Alkharusi A, Alshukaili D, Al Nasiri N, Al-Awadhi T, Charabi Y, El Kenawy AM, 2020. Spatiotemporal Assessment of COVID-19 Spread over Oman Using GIS Techniques. *Earth Syst Environ* 4:1083-98.
- Al Fannah J, Al Harthy H, Al Salmi Q, 2020. COVID-19 pandemic: learning lessons and a vision for a better health system. *Oman Med J* 35:e169.
- Arora P, Kumar H, Panigrahi BK, 2020. Prediction and analysis of COVID-19 positive cases using deep learning models: a descriptive case study of India. *Chaos Solitons & Fractals* 139:110017.
- Asante J, Kreamer D, 2015. A new approach to identify recharge areas in the Lower Virgin River Basin and surrounding basins by multivariate statistics. *Math Geosci* 47:819-42.
- Asante LA, Mills RO, 2020. Exploring the socio-economic impact of COVID-19 pandemic in marketplaces in urban Ghana. *Africa Spectr* 55:170-81.
- Bagal DK, Rath A, Barua A, Patnaik D, 2020. Estimating the parameters of susceptible-infected-recovered model of COVID-19 cases in India during lockdown periods. *Chaos Solitons & Fractals* 140:110154.
- Bager A, Roman M, Algedih M, Mohammed B, 2017. Addressing multicollinearity in regression models: a ridge regression application. *J Soc Econom Stat* 6:30-45.
- Bashir MF, Benjiang M, Shahzad L, 2020. A brief review of socio-economic and environmental impact of Covid-19. *Air Qual Atmos Health* 13:1403-9.
- Belvisi D, Pellicciari R, Fabbri A, Costanzo M, Pietracupa S, De Lucia M, Modugno N, Magrinelli F, Dallochio C, Ercoli TJN, 2020. Risk factors of Parkinson disease: Simultaneous assessment, interactions, and etiologic subtypes. *Neurology* 95:e2500-8.
- Bonaccorsi G, Pierri F, Cinelli M, Flori A, Galeazzi A, Porcelli F, Schmidt AL, Valensise CM, Scala A, Quattrociochi W, 2020. Economic and social consequences of human mobility restrictions under COVID-19. *Proc Natl Acad Sci U S A* 117:15530-5.
- Bray I, Gibson A, White J, 2020. Coronavirus disease 2019 mortality: a multivariate ecological analysis in relation to ethnicity, population density, obesity, deprivation and pollution. *Publ Health* 185:261-3.
- Cao Y, Hiyoshi A, Montgomery S, 2020. COVID-19 case-fatality rate and demographic and socioeconomic influencers: worldwide spatial regression analysis based on country-level data. *BMJ Open* 10:e043560.
- Ciotti M, Angeletti S, Minieri M, Giovannetti M, Benvenuto D, Pascarella S, Sagnelli C, Bianchi M, Bernardini S, Ciccozzi MJC, 2019. COVID-19 outbreak: an overview. *Chemotherapy* 64:215-23.
- Franch-Pardo I, Napoletano BM, Rosete-Verges F, Billa L, 2020. Spatial analysis and GIS in the study of COVID-19. A review. *Sci Total Environ* 739:140033.

- Garcia AAdSG, 2020. Clustering of longitudinal data: application to COVID-19 data. Dissertation 20120-10.2. University of Porto, Portugal. Available from: <https://repositorio-aberto.up.pt/handle/10216/129753>
- Gattinoni L, Chiumello D, Caironi P, Busana M, Romitti F, Brazzi L, Camporota L. 2020. COVID-19 pneumonia: different respiratory treatments for different phenotypes? *Intensive Care Med* 46:1099-102.
- Guo D, 2010. Local entropy map: a nonparametric approach to detecting spatially varying multivariate relationships. *International J Geogr Inf Sci* 24:1367-89.
- Gupta A, Banerjee S, Das S, 2020. Significance of geographical factors to the COVID-19 outbreak in India. *Model Earth Syst Environ* 6:2645-53.
- Iyanda AE, Adeleke R, Lu Y, Osayomi T, Adaralegbe A, Lasode M, Chima-Adaralegbe NJ, Osundina AM. 2020. A retrospective cross-national examination of COVID-19 outbreak in 175 countries: a multiscale geographically weighted regression analysis (January 11-June 28, 2020). *J Infect Public Health* 13:1438-45.
- Kadi N, Khelfaoui M, 2020. Population density, a factor in the spread of COVID-19 in Algeria: statistic study. *Bull Natl Res Cent* 44:1-7.
- Karaye IM, Horney JA. 2020. The impact of social vulnerability on COVID-19 in the US: an analysis of spatially varying relationships. *Am J Prev Med* 59:317-25.
- Khalatbari-Soltani S, Cumming RG, Delpierre C, Kelly-Irving M. 2020. Importance of collecting data on socioeconomic determinants from the early stage of the COVID-19 outbreak onwards. *J Epidemiol Community Health* 74:620-3.
- Kwok CYT, Wong MS, Chan KL, Kwan M-P, Nichol JE, Liu CH, Wong JYH, Wai AKC, Chan LWC, Xu Y, 2020. Spatial analysis of the impact of urban geometry and socio-demographic characteristics on COVID-19, a study in Hong Kong. *Sci Total Environ* 144455 [Epub ahead of print].
- Li Z, Fotheringham AS, Li W, Oshan T, 2019. Fast Geographically Weighted Regression (FastGWR): a scalable algorithm to investigate spatial process heterogeneity in millions of observations. *Int J Geogr Inf Sci* 33:155-75.
- Li Z, Wang W, Liu P, Bigham JM, Ragland DR, 2013. Using geographically weighted Poisson regression for county-level crash modeling in California. *Saf Sci* 58:89-97.
- Lin Y, Zhong P, Chen T, 2020. Association between socioeconomic factors and the COVID-19 outbreak in the 39 well-developed cities of China. *Front Public Health* 8:546637.
- Liu J, Zhou J, Yao J, Zhang X, Li L, Xu X, He X, Wang B, Fu S, Niu T, 2020a. Impact of meteorological factors on the COVID-19 transmission: A multi-city study in China. *Sci Total Environ* 138513 [Epub ahead of print].
- Liu Q, Sha D, Liu W, Houser P, Zhang L, Hou R, Lan H, Flynn C, Lu M, and Hu T. 2020b. Spatiotemporal patterns of COVID-19 impact on human activities and environment in mainland China using nighttime light and air quality data. *Remote Sens* 12:1576.
- Ludvigsson JF, 2020. Systematic review of COVID-19 in children shows milder cases and a better prognosis than adults. *Acta Paediatr* 109:1088-95.
- Mansour S, Al Kindi A, Al-Said A, Al-Said A, Atkinson P, 2020. Sociodemographic determinants of COVID-19 incidence rates in Oman: geospatial modelling using multiscale geographically weighted regression (MGWR). *Sustain Cities Soc* 102627 [Epub ahead of print].
- Ming LC, Untong N, Aliudin NA, Osili N, Kifli N, Tan CS, Goh KW, Ng PW, Al-Worafi YM, Lee KS, 2020. Mobile health apps on COVID-19 launched in the early days of the pandemic: content analysis and review. *JMIR mHealth uHealth* 8:e19796.
- Mollalo A, Vahedi B, Rivera KM. 2020a. GIS-based spatial modeling of COVID-19 incidence rate in the continental United States. *Sci Total Environ* 728:138884.
- Nakada LYK, Urban RC, 2020. COVID-19 pandemic: environmental and social factors influencing the spread of SARS-CoV-2 in S o, Brazil. *Environ Sci Pollut Res Int* 28:1-7.
- Nipperess DA, Andersen AN, Pik AJ, Bramble R, Wilson P, Beattie AJ, 2008. The influence of spatial scale on the congruence of classifications circumscribing morphological units of biodiversity. *Divers Distrib* 14:917-24.
- Pászto V, Burian J, Mackù K, 2020. COVID-19 data sources: evaluation of map applications and analysis of behaviour changes in Europe's population. *Geografie (Utrecht)* 125:171-209.
- Peeters A, Zude M, Käthner J, Ünlü M, Kanber R, Hetzroni A, Gebbers R, Ben-Gal A, 2015. Getis-Ord's hot-and cold-spot statistics as a basis for multivariate spatial clustering of orchard tree data. *Comput Electron Agric* 111:140-50.
- Pick JB, Nishida T, 2015. Digital divides in the world and its regions: a spatial and multivariate analysis of technological utilization. *Technol Forecast Soc Change* 91:1-17.
- Pirdavani A, Bellemans T, Brijs T, Wets G, 2014. Application of geographically weighted regression technique in spatial analysis of fatal and injury crashes. *J Transport Eng* 140:04014032.
- Pourghasemi HR, Pouyan S, Heidari B, Farajzadeh Z, Shamsi SRF, Babaei S, Khosravi R, Etemadi M, Ghanbarian G, Farhadi A, 2020. Spatial modeling, risk mapping, change detection, and outbreak trend analysis of coronavirus (COVID-19) in Iran (days between February 19 and June 14, 2020). *Int J Infect Dis* 98:90-108.
- Pramanik M, Udmale P, Bisht P, Chowdhury K, Szabo S, Pal I, 2020. Climatic factors influence the spread of COVID-19 in Russia. *Int J Environ Health Resh* [Epub ahead of print].
- Roy S, Bhunia GS, Shit PK, 2020. Spatial prediction of COVID-19 epidemic using ARIMA techniques in India. *Model Earth Syst Environ* [Epub ahead of print].
- Sannigrahi S, Pilla F, Basu B, Basu AS, Molter A. 2020. Examining the association between socio-demographic composition and COVID-19 fatalities in the European region using spatial regression approach. *Sustain Cities Soc* 62:102418.
- Sarra AL, Nissi E, 2016. Geographically weighted regression analysis of cardiovascular diseases: evidence from Canada Health Data. In: Di Battista T., Moreno E., Racugno W. (Eds.), *Topics in Methodological and Applied Statistical Inference. Studies in Theoretical and Applied Statistics*. Springer, Cham., Berlin, Germany, pp 191-203.
- Shakil MH, Munim ZH, Tasnia M, Sarowar S, 2020. COVID-19 and the environment: A critical review and research agenda. *Sci Total Environ* [Epub ahead of print].
- Singhal A, Seborg DE, 2005. Clustering multivariate time-series data. *J Chemom* 19:427-38.
- Song L, Cao Y, Zhou W, Kuang X, Luo G, 2019. Study on the variation of arable land use and management countermeasures under rapid urbanization: the application of a gravity model in a regional perspective. *Environ Monit Assess* 191:120.
- Song W, Wang C, Chen W, Zhang X, Li H, Li J, 2020. Unlocking



- the spatial heterogeneous relationship between Per Capita GDP and nearby air quality using bivariate local indicator of spatial association. *Resour Conserv Recycl* 160:104880.
- Su S, Gong Y, Tan B, Pi J, Weng M, Cai Z, 2017. Area social deprivation and public health: Analyzing the spatial non-stationary associations using geographically weighed regression. *Soc Indic Res* 133:819-32.
- Sun F, Matthews SA, Yang T-C, Hu M-H, 2020a. A spatial analysis of the COVID-19 period prevalence in US counties through June 28, 2020: where geography matters? *Ann Epidemiol* 52:54-59.e1.
- Sun Z, Zhang H, Yang Y, Wan H, Wang Y, 2020b. Impacts of geographic factors and population density on the COVID-19 spreading under the lockdown policies of China. *Sci Total Environ* 746:141347.
- Verma A, Prakash S, 2020. Impact of covid-19 on environment and society. *J Glob Biosci* 9:7352-63.
- Vishwakarma S, Nair PS, Rao DS, 2017. A Comparative Study of K-means and K-medoid Clustering for Social Media Text Mining. *Int J Adv Sci Res Engineering Trends* 2:297-301.
- Waheed A, Shafi J, 2020. Successful role of smart technology to combat COVID-19. pp 772-777 in 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC): IEEE.
- Weissleder R, Lee H, Ko J, Pittet MJ, 2020. COVID-19 diagnostics in context. *Sci Transl Med* 12:eabc1931.
- Wheeler DC, Páez A, 2010. Geographically weighted regression. In: Fischer M., Getis A. (Eds.), *Handbook of applied spatial analysis*. Springer, Berlin, Heidelberg, Germany, pp 461-486.
- Wheeler D, Tiefelsdorf M, 2005. Multicollinearity and correlation among local regression coefficients in geographically weighted regression. *J Geogr Syst* 7:161-87.
- Windle MJ, Rose GA, Devillers R, Fortin M-J, 2010. Exploring spatial non-stationarity of fisheries survey data using geographically weighted regression (GWR): an example from the Northwest Atlantic. *ICES J Mar Sci* 67:145-54.
- Wiwanitkit V, 2020. COVID-19 infection in Oman. *Oman Med J* 35:e186.
- Yang T-C, Shoff C, Matthews SA, 2013. Examining the spatially non-stationary associations between the second demographic transition and infant mortality: a Poisson GWR approach. *Spatial Demogr* 1:17-40.
- Yu H, Peng Z-R, 2019. Exploring the spatial variation of ridesourcing demand and its relationship to built environment and socioeconomic factors with the geographically weighted Poisson regression. *J Transport Geogr* 75:147-63.