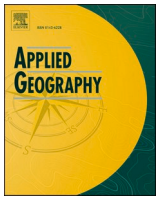
Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Applied Geography 134 (2021) 102523

Contents lists available at ScienceDirect

Applied Geography

journal homepage: www.elsevier.com/locate/apgeog



The main factors influencing COVID-19 spread and deaths in Mexico: A comparison between phases I and II

Francisco Benita [a,*], Francisco Gasca-Sanchez [b]

[a] Engineering Systems and Design, Singapore University of Technology and Design, Singapore
[b] Business School, Department of Economics, Universidad de Monterrey, Mexico

ARTICLE INFO

Keywords:
Mexico lockdown
Cluster analysis
Municipalities
Socio-demographic determinants
Regression

ABSTRACT

This article investigates the geographical spread of confirmed COVID-19 cases and deaths across municipalities in Mexico. It focuses on the spread dynamics and containment of the virus between Phase I (from March 23 to May 31, 2020) and Phase II (from June 1 to August 22, 2020) of the social distancing measures. It also examines municipal-level factors associated with cumulative COVID-19 cases and deaths to understand the spatial determinants of the pandemic. The analysis of the geographic pattern of the pandemic via spatial scan statistics revealed a fast spread among municipalities. During Phase I, clusters of infections and deaths were mainly located at the country's center, whereas in Phase II, these clusters dispersed to the rest of the country. The regression results from the zero-inflated negative binomial regression analysis suggested that income inequality, the prevalence of obesity and diabetes, and concentration of fine particulate matter (PM 2.5) are strongly positively associated with confirmed cases and deaths regardless of lockdown.

1. Introduction

As of May 1, 2021, Mexico ranks fourth globally in the COVID-19 death rates, behind only by the U.S., Brazil, and India. The country has a total of 2,344,755 positive cases and 216,907 deaths, which currently represent one of the highest lethality rates (10.81 %) worldwide. Like other Latin American countries, Mexico is framed by growing poverty levels, large income inequality, and disparities in access to healthcare services (Dávila-Cervantes & Agudelo-Botero, 2019; Hessel et al., 2019; Mesenburg et al., 2018). Furthermore, the prevalence of hypertension, diabetes, and other conditions that might shape obesity in the country has a more significant role in the lethality of the disease. The strong relationship between overweight or obese individuals and severe hospitalization outcomes has been well documented (Hernández-Garduño, 2020; Popkin et al., 2020). Hence, individuals with excessive body fat or major cardiometabolic problems (ranging from hypertension to cardiovascular disease and diabetes) could be presented with severe or even lethal complications. In this vain, Mexico faces a significant risk as it has been in the second spot in the global ranking of adult obesity, just behind the U.S., and the highest worldwide in childhood overweight and obesity, for years now (OECD (Organisation for Economic Co-operation and Development), 2019).

Similarly, an enormous amount of scientific literature has focused on the possible links between climate factors, such as air temperature, relative humidity or exposure to air pollution, and transmission rates and lethality. With respect to air temperature, which is easy to measure and spatially explicit climate data, results are controversial as there is no clear evidence that a temperature rise reduces case counts of COVID-19 (Briz-Redón & Serrano-Aroca, 2020; Lin et al., 2020a, 2020b). Interestingly, these studies pointing to significant positive association between air pollution and lethality rates. This has been explained by the contribution of air pollution to respiratory tract infections, pulmonary disease, or diabetes burden (Lubrano et al., 2020; Ma et al., 2020). With vastly different climates among its states and municipalities, Mexico has displayed differentiated effects on confirmed COVID-19 cases (Méndez-Arriaga, 2020). Some other studies have shown that the underlying socio-economic conditions could correlate with the dynamics of the pandemic. For instance, Bambra et al. (2020) documented how the exacerbated existing social inequalities in the country have translated to higher infection rates and mortality among the most disadvantaged communities. One can also hypothesize that the relative number of COVID-19 infections and deaths in Mexican metropolitan areas, like other highly densely populated cities in the world (Sun et al., 2020a, 2020b, 2020c), would be larger than non-metro areas.

* Corresponding author.
E-mail addresses: francisco_benita@sutd.edu.sg (F. Benita), francisco.gasca@udem.edu (F. Gasca-Sanchez).

https://doi.org/10.1016/j.apgeog.2021.102523
Received 18 December 2020; Received in revised form 11 July 2021; Accepted 23 July 2021
Available online 27 July 2021

On top of the above-mentioned socio-economic, health, and climate factors associated with the dynamics of the disease, physical distancing and lockdown measures restricting mobility add an extra layer of complexity to understanding the unfolding of the pandemic. This has caused regionally differentiated impacts calling for a geographical approach to health policy responses. In Mexico, "Phase I" of temporary mobility restriction efforts started with the National Safe and Social Distancing Program (in Spanish *Jornada Nacional de Sana Distancia*) and comprised 14 weeks between March 23 and May 31, 2020. In Phase I, residents were exhorted to (i) stay at home (unenforced), (ii) wash their hands repeatedly, and (iii) avoid touching their face. With 56 % of informal employment, the possibility to "stay home" was limited to a small sub-group of Mexicans. In fact, reports of automobile traffic and tracking of mobile phones and social media activities indicated that most people were ignoring Phase I mobility restrictions (Macip, 2020).

This study investigates the main socio-economic, health, climate, and mobility factors associated with COVID-19 infections and deaths. We explore zero-inflated negative binomial (ZINB) regression models. Similarly, this work also attempts to understand the transmission dynamics of current and future emerging geographical clusters by using scan statistical methods.

The remainder of this paper is structured as follows. Section 2 presents our methodological framework. Section 3 shows the main results and discussion of the empirical findings. Finally, Section 4 concludes with policy recommendations and future work.

2. Data and methods

2.1. Data sources

Table 1 presents the data sources and the summary statistics for all variables used in this study. We consider all the 2459 municipalities in 32 States. Here Phase II is defined as the following 14 weeks after the end of Phase I (i.e., between June 1 and August 22, 2020). The column "VIF" indicates the variance inflation factor to determine the multicollinearity problem and avoid overfitting, that is, (if needed) drop highly correlated variables during the econometric strategy (see Section 2.3).

2.1.1. Measures: COVID-19 cases and deaths

Monthly counts of COVID-19 cases and deaths are the outcome variables of interest. The cumulative monthly number of confirmed cases and deaths in each municipality until August 22, 2020, was calculated to construct both variables. During the social distancing program, a total of 1493 (60.71 %) municipalities reported at least one confirmed case, and 1216 (50.5 %) reported at least one death. Three months after the social distancing program, 2096 (85.2 %) and 1551 (63 %) municipalities reported at least one confirmed case and death, respectively.

2.1.2. Measures: potential factors influencing positive COVID-19 cases and deaths

Socio-economic: a dummy of *metropolitan area*, the log *of population density*, *proportion of males*, *Gini of income inequality*, and *proportion of poverty*. As shown in Table 1, 36 ± 95 % (mean ± standard deviation) of the municipalities belong to metropolitan areas. Meanwhile, the average

Table 1
Summary statistics.

	Phase I[a]			Phase II[a]			Data Source
Variable	Mean	Std. Dev.	VIF[c]	Mean	Std. Dev.	VIF[c]	
Monthly counts of COVID—19 cases (cumulative)	14.84	123.93		126.54	570.29		1
Monthly counts of COVID—19 deaths (cumulative)	2.84	17.88		12.83	57.02		1
Socio—economic							
Metropolitan area {0–1}	0.36	0.95	1.87	–	–	1.89	2
Population density (people per sq. km)[b]	319.83	1269.56	2.70	–	–	2.72	3
Proportion of males	0.95	0.06	2.53	–	–	2.53	3
Income inequality (Gini) [0–1]	0.39	0.04	1.07	–	–	1.07	4
Proportion of people in poverty [0–1]	0.65	0.22	2.86	–	–	2.81	4
Health							
Proportion of obesity [0–1][c]	0.32	0.09	3.35	–	–	3.53	5
Proportion of hypertension [0–1][c]	0.19	0.04	1.97	–	–	1.96	5
Proportion of diabetes [0–1][c]	0.10	0.02	1.63	–	–	1.62	5
Fast-food outlets per 1000 people	4.03	3.08	1.26	–	–	1.21	6
Climate							
Avg. monthly max. temp. (°F)[b]	91.04	7.11	2.60	85.29	9.13	4.49	7
Avg. monthly min. temp. (°F)[b]	56.65	9.03	2.22	59.66	10.39	3.07	7
Monthly cumulative rainfall (mm)[b]	53.28	81.24	1.90	187.36	142.37	1.18	7
Annual PM 2.5 (metrics tons)[b]	249.89	560.66	1.31	–	–	1.33	8
Commuting							
Public buses per 1000 people	1.47	4.85	1.17	–	–	1.16	3
Private cars per 1000 people	109.80	137.38	1.71	–	–	1.69	3
Proportion of essential activities [0,1]	0.68	0.11	1.09	–	–	1.09	6

Notes: [a]Values reported as "—" in column "Phase II" correspond to the same values as the ones in column "Phase I"[b.] We take the log of this variable for estimation purposes in our econometric strategy.[c]In adults aged 20 years or older; VIF above 5 (*10*) indicates moderate (*severe*) collinearity among variables (Dormann et al., 2013).

Data sources.

1. SS (Secretaría de Salud).
2. INEGI 2015 (Instituto Nacional de Estadística y Geografía-Encuesta nacional intercensal).
3. INEGI 2019 (INEGI-Vehículos de motor registrados en circulación).
4. CONEVAL 2015 (Consejo Nacional de Evaluación de la Política y Desarrollo Social-La pobreza en México 2015).
5. ENSANUT 2018 (Encuesta Nacional de Salud y Nutrición).
6. DENUE 2019 (Directorio Nacional de Unidades Económicas).
7. CONAGUA (Comisión Nacional del Agua- Servicio Meteorológico Nacional).
8. SEMARNAT 2016 (Secretaría de Medio Ambiente y Recursos Naturales).

Gini income inequality per municipality is 0.39 ± 0.04, whereas in the average municipality, 65 ± 22 % of the population lives below the poverty line. On average, municipalities have 319.83 ± 1269.56 inhabitants per square kilometer of land. However, its standard deviation is considerably large, indicating that data points tend to be far from the mean value.

Health: the proportions of obesity, hypertension, and diabetes, and fast-food outlets per *1000 people*. Here, our variables measuring the presence of obesity, hypertension, and diabetes correspond to adults aged 20 years or older, but cases and deaths were not exclusively concentrated on such age groups. There are two reasons behind our choice. First, all three variables are the most comprehensive that can provide at the municipal level (no other proxies are available). Second, as of April 22, 2020, which is the end of Phase II according to this study, 94.8 % (*99.1 %*) of the confirmed cases (*deaths*) correspond to adults aged 20 years and older. Moreover, as of May 1, 2021, the percentage of cases (*deaths*) has remained virtually unchanged at 96.7 % (*99.8 %*). The proportion of obesity is high; that is, on average, 32 ± 9 % of the population aged 20 years and above suffer from obesity. As mentioned previously, the relationship between obesity and respiratory disease is largely known (Jiao et al., 2015; Richardson et al., 2015; Walton et al., 2009), and we expect its positive relationship with COVID-19 deaths. However, the magnitude of the coefficient estimates associated with obesity and the coefficients of the other comorbidities are region-dependent.

Climate: log *of average monthly maximum* and *minimum temperature,* log *of cumulative monthly rainfall, and* log *of annual PM 2.5 concentration*. Here, to ensure that the log transformation does not disrupt our analysis, we have converted temperature into Fahrenheit degrees (see Section 2.3) to avoid negative values in Celsius degrees during cold months. Long-term exposure to PM 2.5 adversely affects respiratory and cardiovascular systems and may increase lethality rates of COVID-19 (Wu et al., 2020). This could be particularly true for municipalities in metro areas. Méndez-Arriaga (2020) has documented that climate conditions in the Mexican states with low temperatures and poor precipitation (in mm), such as Tlaxcala, Guanajuato, and Michoacán, favored local infections in Phase I.

Commute: *public buses* per *1000 people, private vehicles automobiles* per *1000 people,* and *the proportion of economic units operating essential activities*. These proxies for human mobility behavior are expected to be positively correlated with the spread of the virus (see Benita, 2021), but its association with deaths remains unknown. The variable proportion of essential activities include health-related companies, public security, and fundamental sectors of the economy (utilities, food and beverages, agriculture, fisheries and livestock, industrial chemicals, storage, courier services, private security, telecommunications, emergency services, funeral services, transportation of individuals, logistics, and governmental activities). The average percentage of economic units that operate essential activities during Phase I is 68 ± 11 %.

2.2. Spatial scanning statistical analysis

Kulldorf's (1997, 2011) spatial scan statistics is used to identify the size of emerging geographical clusters of cumulative disease cases and deaths during Phases I and II. Numerous scientists have successfully used this spatial scanning to detect and evaluate geographical disease clusters (Kihal-Talantikite et al., 2013; Rao et al., 2017). Moreover, recent spatial scan statistics applications in the context of COVID-19 can be found in the works of Desjardins et al. (2020) or Kim and Castro (2020) for the U.S. and South Korean cases, respectively. This purely spatial scan statistics imposes a circular window to the geographical cluster and allows its centroid to move across locations.

The null hypothesis of our Kulldorff spatial scan statistic states that the outcome variable of interest, namely, COVID-19 cases and deaths, is randomly distributed across space, and the expected count is proportional to the population at risk. The alternative hypothesis states that at least one circle has a higher relative risk than the outside. Note that the relative risk (RR) is a non-negative number representing how much more common disease is inside a geographical cluster than the outside. More precisely, for each circle, one can compute a likelihood, and the circle with the maximum likelihood function among all radius sizes at all possible point locations is called the most-likely cluster. Other clusters with significant log-likelihood ratios (*LLR*) can be seen as secondary potential clusters.

We consider the discrete Poisson model, per phase and per outcome variable, where we assume that the outcome variable follows a Poisson distribution according to the population of the geographic region. Then, we detect potential clusters by calculating the following likelihood ratio for each circle: $(c/e)^c((C-c)/(C-e))^{C-c}I()$. Here, C is the total number of counts (i.e., cases or deaths), c and e are the observed and expected number of counts within a circle, respectively. The binary indicator $I()$ facilitates the identification of high-risk clusters as it equals 1 when $c > e$, and 0 otherwise. Each *LLR* has a *p-value* estimated via 999 Monte Carlo Simulations. If *p-value* < 0.05, then the geographical cluster has a significantly higher risk of COVID-19 infections (deaths) than that outside the area.

2.3. Econometric strategy

Analogous to our spatial scanning statistical method, we elaborate separate datasets in which municipalities are the unit of analysis for each phase and outcome variable. The aim is to propose a model that uses the predictors mentioned above to approach the outcome variable and check whether a statistically significant influence exists or not. Both outcome variables, that is, cumulative cases and cumulative deaths, are counting variables with zeros at some dates. To model the counting variables, researchers use many approaches such as the Poisson model, quasi-Poisson model, negative binomial (NB) model, zero-inflated Poisson (ZIP), ZINB, and Hurdle models. Particularly, zero-inflated models provide the flexibility to model zero counts as they correct for overdispersion (e.g., excessive zeros in the count data). Note that we suspect overdispersion because in Phases I and II, a great variability (disproportion among municipalities; see Section 2.1.1) regarding total cases and deaths exists.

For each dataset, the econometric strategy follows the model presented by McLaren (2021):

$$\log(E[y_{it}]) = \beta_0 + \sum_m \beta_m D_{mt} + \sum_s \beta_s X_{s,it} + \varepsilon_{it}, \quad (1)$$

where $y = 0, 1, 2, \ldots$ is the outcome variable; i stands for municipality and t for the month. D_{mt} is a dummy variable that takes the value of 1 if $m = t$, and 0 otherwise. β_s are the coefficients associated with the s predictor variables.

2.4. Data processing

Most consulted datasets provide information at the municipal level; however, records of air temperature and rainfall were retrieved from 1520 meteorological stations along the country. To append each municipality with its respective climate data, we have implemented the spatial interpolation of kriging using the Geostatistical Analyst Toolbox in ArcGIS 10.4.1. Next, the variables from all eight data sources enlisted in Table 1 were merged using the National Institute of Statistics, Geography and Informatics (NEGI) geostatistical code as each municipality's unique identifier.

We then construct four main datasets, namely, cumulative cases in Phase I, cumulative cases in Phase II (also containing the cases in Phase I), cumulative deaths in Phase I, and cumulative deaths in Phase II (also containing the dates in Phase I). All the main results derived from this study correspond to each one of the four datasets. The spatial scan statistic was implemented in the Kulldorff (2011) "SaTScan" software version 9.4.2. Then, we used ArcGis to visualize the RR of positive

COVID-19 cases and deaths in high-risk cluster areas. The econometric strategy was conducted using various R software packages.

Multicollinearity among predictors was assessed by determining the VIF of the independent variables using the "olsrr" package (Hebbali, 2020). To unveil the association between the outcome variables and socio-economic, health, climate, and commute predictors, we have explored numerous possible regression models (i.e., Poisson, ZIP, NB, and ZINB) using the following three steps. First, we estimate the classic Poisson model that is adequate when its conditional variance is equal to its conditional mean in the counting process. This assumption can be validated using the quasi-Poisson model, which can also check for overdispersion. Second, in the presence of overdispersion, NB and ZINB are more flexible, attractive models because they allow the relaxation of the strong assumptions regarding the relationship between the mean and variance. To decide whether NB is preferred over ZINB (e.g., if the splitting mechanism of the ZINB is rejected or not), we use the well-known Vuong (1989) test. Third, if the ZINB is preferred, we should only compare ZINB with ZIP. To do so, we check the statistical significance of the dispersion parameter. If this parameter is statistically significant, ZIP model is rejected, and the best estimates correspond to the ZINB model. Poisson, quasi-Poisson, and NB models were implemented using the "MASS" package (Ripley et al., 2013), whereas ZIP and ZINB models were conducted with the "psc" package (Jackman et al., 2015).

3. Results

3.1. Spatial location of the pandemic in Mexico during phases I and II

Fig. 1(a) depicts the geographical location and size of the pandemic during Phase I. At the beginning of the outbreak, the most-likely cluster of COVID-19 cumulative positive cases was located in Mexico City metropolitan area, whereas secondary clusters were also situated at the center of the country with spatial overlaps. More precisely, the most-likely cluster was composed of 21 municipalities (see Table 2), accounting for 29,784 positive cases in a radius of 23.24 km. Recall that Mexico City metropolitan area has a total population of more than 20 million inhabitants. Like New York in the U.S., this high-density area was the first to be hit by the pandemic. During this period, the risk of positive COVID-19 cases of the most-likely cluster was about 4.5 ($RR = 4.49$) times higher than that in other clusters.

The effectiveness of social distancing measures imposed by local and federal governments could be appreciated in Fig. 1(b). As shown in the figure, most-likely and secondary clusters shifted toward the southern and eastern regions. As expected, these dynamics indicate the rising spread of the virus in less populated areas of the States of Tlaxcala or Guanajuato. However, the results of spatial clusters of COVID-19 deaths that show an opposite pattern are interesting. That is, municipalities clustered at high-risk areas during Phase I were located along different

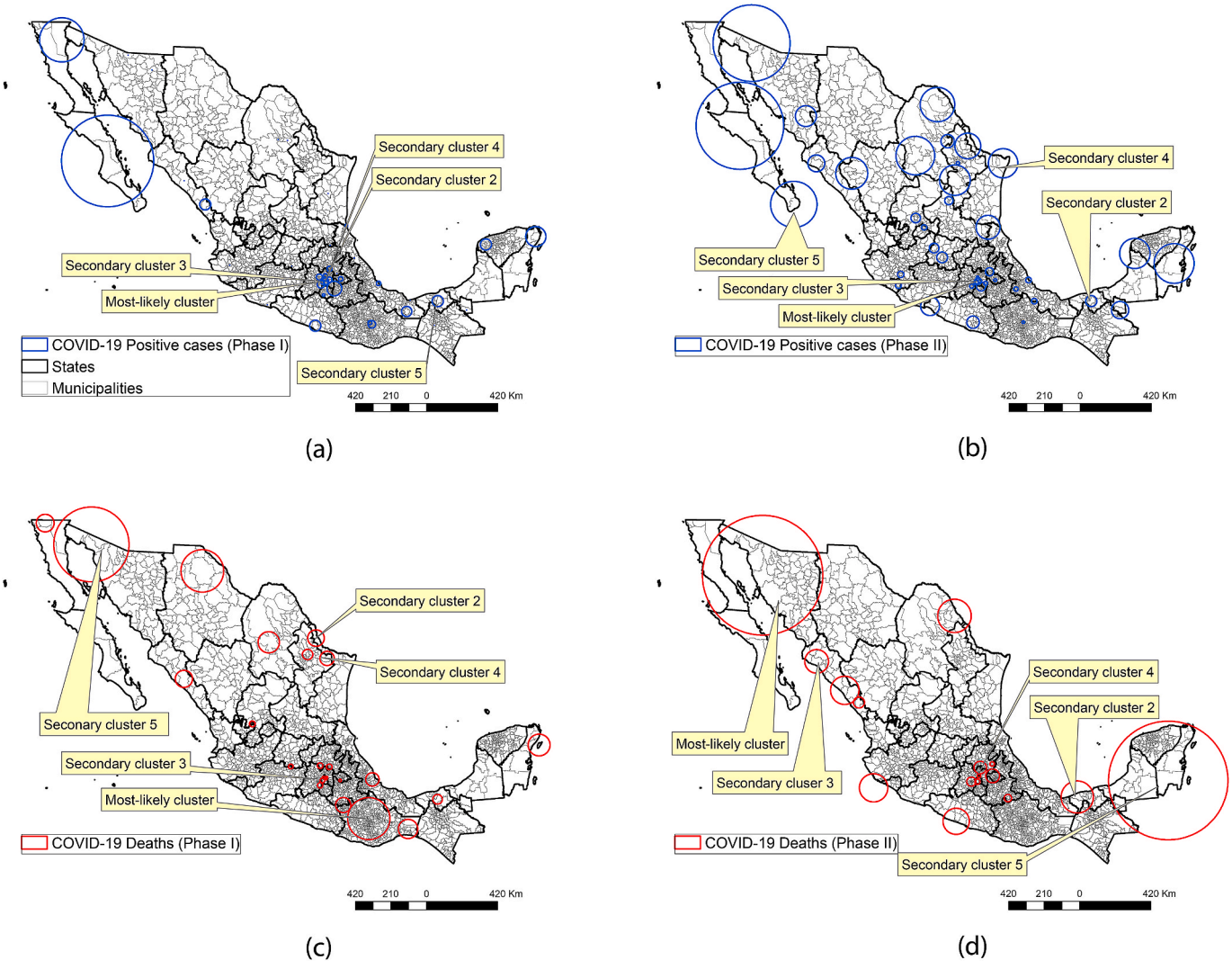


Fig. 1. Spatial location of COVID-19 high-risk areas; Cluster of positive cases during (a) Phase I and (b) Phase II; Clusters of deaths during (c) Phase I and (d) Phase II.

Table 2
Spatial COVID-19 clusters detected by SaTScan.

Cluster	The Lat	Lon	Radius (km)	Municipalities	*LLR*	Observed	Expected	*RR*	*p−value*
Phase I (positive cases)									
Most−likely cluster	19.3267	−99.1504	23.27	21 (MCMA)	18142.29	29,784	8813.94	4.49	<0.01
Sec. cluster 2	19.4314	−99.1491	8.80	7 (MCMA)	7182.68	10,916	2776.27	4.32	<0.01
Sec. cluster 3	19.2451	−99.0904	12.44	5 (MCMA)	7156.70	10,204	2446.21	4.56	<0.01
Sec. cluster 4	19.1652	−99.3593	22.37	18 (MCMA, Morelos)	2094.09	4977	1717.70	3.00	<0.01
Sec. cluster 5	18.0091	−92.8602	32.76	4 (Tabasco)	1744.04	2859	748.92	3.91	<0.01
Phase II (positive cases)									
Most−likely cluster	19.3267	−99.1504	20.04	16 (MCMA)	10362.04	62,662	34359.67	1.96	<0.01
Sec. cluster 2	18.0091	−92.8602	32.76	4 (Tabasco)	7682.62	13,065	3524.42	3.79	<0.01
Sec. cluster 3	19.269	−99.2684	13.94	5 (MCMA)	5332.78	19,680	8617.41	2.34	<0.01
Sec. cluster 4	28.812	−101.395	101.83	11 (Coahuila)	3057.02	6718	2174.98	3.12	<0.01
Sec. cluster 5	27.2861	−113.231	258.68	4 (BCS)	2150.18	8551	3848.19	2.25	<0.01
Phase I (deaths)									
Most−likely cluster	17.1637	−96.7337	124.40	415 (Oaxaca)	698.19	1072	270.32	4.20	<0.01
Sec. cluster 2	26.9041	−99.4827	49.00	3 (Nuevo León, Tamaulipas)	354.83	96	0.89	108.87	<0.01
Sec. cluster 3	19.3492	−99.0568	6.60	2 (MCMA)	213.29	635	249.29	2.62	<0.01
Sec. cluster 4	17.9299	−98.1292	44.69	50 (Puebla, Oaxaca)	207.27	197	29.55	6.75	<0.01
Sec. cluster 5	32.4409	−116.269	51.85	2 (BCN)	203.07	572	217.47	2.70	<0.01
Phase II (deaths)									
Most−likely cluster	29.9518	−112.094	353.85	69 (Sonora, BCN)	725.29	3179	1514.84	2.18	<0.01
Sec. cluster 2	18.1702	−93.6468	95.32	21 (Tabasco, Veracruz)	339.11	1879	974.13	1.97	<0.01
Sec. cluster 3	25.5346	−108.558	70.17	5 (Sinaloa)	337.37	943	354.92	2.69	<0.01
Sec. cluster 4	19.5041	−99.1159	10.78	6 (Mexico State)	322.86	2873	1740.26	1.70	<0.01
Sec. cluster 5	18.6862	−88.5025	352.29	128 (Campeche, Yucatan, QR)	210.36	2641	1740.59	1.55	<0.01

Notes: BCN, Baja California; BCS, Baja California Sur; MCMA, Mexico City metropolitan area; QR, Quintana Roo.

latitudes and longitudes (States of Baja California, Nuevo León, Tamaulipas, Oaxaca, and Puebla) without a visible spatial pattern. Alternatively, the high-risk areas during Phase II were concentrated around Mexico City metro area (States of Mexico, Mexico City, Tlaxcala, Puebla, and Hidalgo). The center of the most-likely cluster of COVID-19 deaths during Phase II was located in Sonora State, with centroid coordinates 29.9518 latitude and −112.094 longitude (*LLR* = 698.19 *p-value* < 0.01). This circular area covered 69 municipalities with a radius of 353.85 km, and it included municipalities sharing a border with the U.S. The total number of COVID-19 deaths in the cluster is 3,179, and the fatality risk is about two times (*RR* = 2.18) higher than outside this area.

We also performed Kulldorff's space–time scan statistical analysis to detect possible space–time clusters as recently suggested elsewhere (Desjardins et al., 2020). The advantage under this setting is that it locates high-risk clusters not only in space but also in time. The results of the space–time analysis (not presented here) were of little interest to this work as they lead to large-sized most-likely cluster areas and large-sized secondary cluster areas, thereby making the detection of high-risk

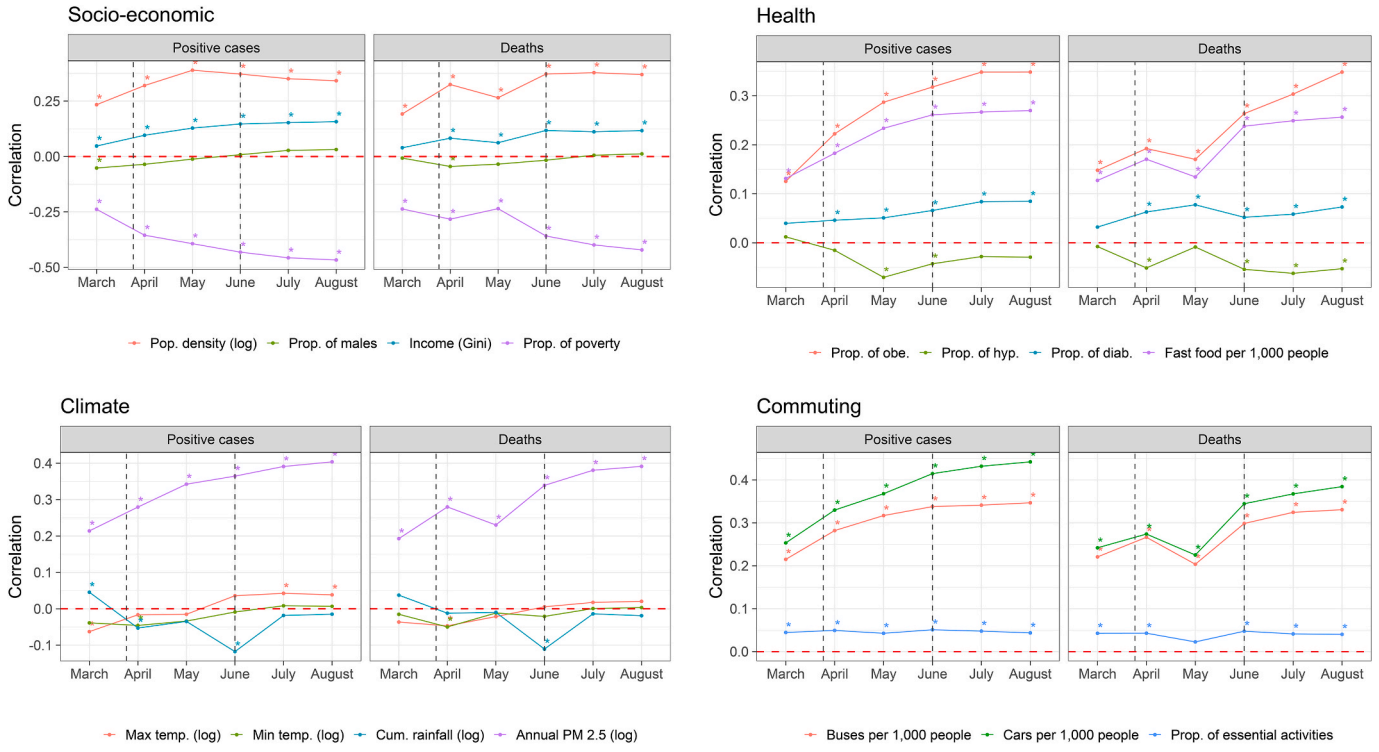


Fig. 2. Evolution of the Kendall rank correlation between cumulative monthly new positive COVID-19 cases and new deaths using data from 2459 Mexican municipalities. Vertical dashed lines in black represent the start and end of Phase I. "*" indicates *p-value* < 0.01.

clusters far less clear than the simple spatial scan framework. Possible reasons for these results are the irregular size and spatial shape of the municipalities and the time span of this study.

Lastly, the spatial distribution of positive COVID-19 cases and deaths is probably shaped by climate and social processes related to the municipalities' growth and social inequalities. This motivates our further analysis regarding the examination of the main drivers of COVID-19 disease spread, which we discuss in the next section.

3.2. Correlation analysis between municipal-level characteristics and COVID-19 cases and deaths

Fig. 2 serves as an initial understanding of the association between output variables and predictors. As noted in the recent literature review of Briz-Redón and Serrano-Aroca (2020), many of the early studies focusing on the association between several climate, socio-economic, or mobility factors, and the number of cumulative or daily COVID-19 cases have been based, solely or partially, on simple correlation analysis (Pearson's, Spearman's rank, or Kendall's rank correlation). Shapiro Wilk's test was applied to evaluate the normality of the data, but we found evidence of non-normal distribution. Therefore, the Kendall rank correlation was tested. Recall that this is a non-parametric rank-order correlation that does not make any assumptions about the data distribution. Moreover, Kendall rank correlation has been recently used to investigate the effects of meteorological (Bashir et al., 2020) or socio-economic parameters on COVID-19 cases.

By visual inspection, we notice that population density (log), the proportion of adults aged 20 years or older with obesity, annual PM 2.5 (log), and cars per 1000 people are strongly positively correlated with cumulative monthly COVID-19 cases and deaths. Our further regression analysis shall produce refined insights and stronger conclusions.

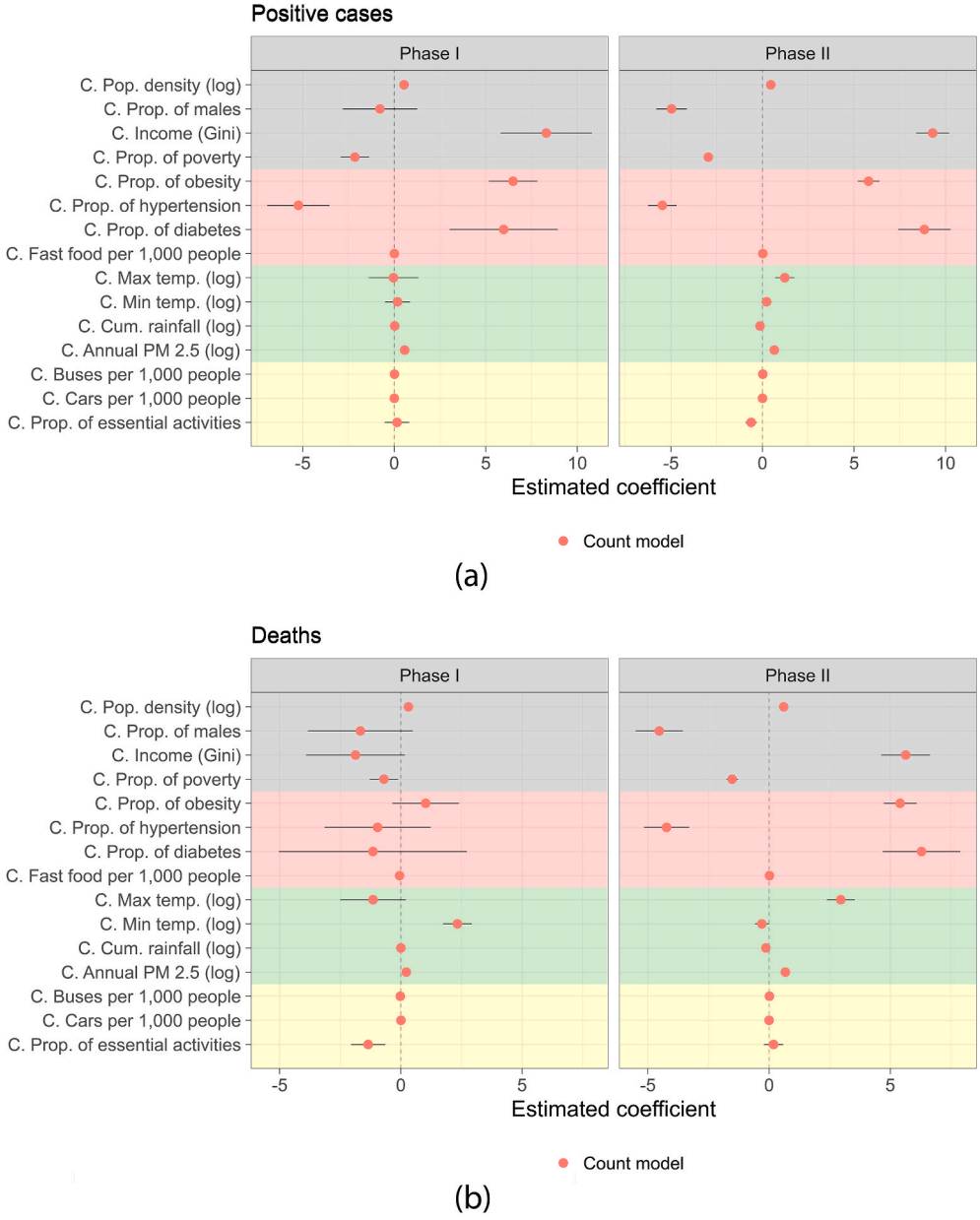


Fig. 3. The effect size of predictor coefficients in the count model of the zero-inflated negative binomial. Black lines represent the 95 % confidence interval. Variables in gray, red, green, and yellow correspond to socio-economic, health, climate, and commuting factors, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

3.3. Main determinants of COVID-19 cumulative cases and deaths in phases I and II

We start by noticing that all VIF values presented in Table 1 do not exceed the accepted threshold of 5 (Dormann et al., 2013). Thus, we discard the presence of significant multicollinearity among the variables. The Vuong test detected significant overdispersion for both of our output variables, thus strongly favoring a ZINB regression. A combination of a process (i.e., logistic regression in this study) that predicts some zero counts and another one (i.e., NB) that models the counts is used in the ZINB regression. The logistic analysis of the ZINB model predicts the probability of *y* falling into the zero group (e.g., no positive cases (*no deaths*) in municipality *i*), whereas the NB part models the expectation of *y* conditional on a non-zero value; that is, the expected count of the number of positive cases (*deaths*) in municipality *i*.

For ease of synthesizing our results, Fig. 3 shows the effect size of the estimated coefficients related to the count part of the model (e.g., NB), and Tables 3 and 4 summarize the full results. Recall that the count model describes associations between predictors and COVID-19 cases (*deaths*) among municipalities with at least one reported COVID-19 case (*death*). During both Phases I and II, we observed a large and significant positive association between municipal-level cumulative positive cases and the variables proportion of obesity, proportion of hypertension, proportion of diabetes, and income inequality (Fig. 3(a)). Although positive and statistically significant, the estimated coefficients of the proportion of essential activities, population density (log), average minimum temperature (log), and annual PM 2.5 concentration (log) are considerably smaller.

Concerning COVID-19 fatalities, population density (log) and air pollution are positively associated with deaths. The statistical significance of both variables is s not surprising, as we could hypothesize such results from the strong positive correlation shown in Fig. 1. The estimated coefficient of average maximum temperature goes from negative (Phase I) to positive (Phase II) indicating sample dependency.

Socio-economic and health-related variables played an important role in the evolution of the monthly cumulative number of COVID-19

Table 3
Incidence rate ratios (IRR) of cumulative COVID-19 cases.

	Phase I		Phase II	
	IRR	p-value	*IRR*	p-value
Count Model				
Intercept	4.92E-06	<0.001	7.54E-4	<0.001
Month 2	14.42	<0.001	2.532	<0.001
Month 3	111.26	<0.001	3.54	<0.001
Socio-economic				
Metropolitan area	1.11	0.266	1.23	<0.001
Pop. Density (log)	1.71	<0.001	1.58	<0.001
Prop. of males	0.46	0.451	0.01	<0.001
Income inequality (Gini)	4045.05	<0.001	10,899.25	<0.001
Prop. of people in poverty	0.12	<0.001	0.05	<0.001
Health				
Prop. of obesity	658.14	<0.001	327.83	<0.001
Prop. of hypertension	0.01	<0.001	0.01	<0.001
Prop. of diabetes	394.17	<0.001	6932.12	<0.001
Fast-food per 1000 people	1.01	0.652	1.02	0.005
Climate				
Avg. monthly max. temp. (log)	0.96	0.952	3.38	<0.001
Avg. monthly min. temp. (log)	1.2	0.606	1.25	0.102
Monthly cumulative rainfall (log)	1.03	0.29	0.87	<0.001
Annual PM 2.5 (log)	1.77	<0.001	1.91	<0.001
Commuting				
Public buses per 1000 people	1.01	0.058	1.02	<0.001
Private cars per 1000 people	1.01	<0.001	1.01	<0.001
Prop. of essential activities	1.16	0.659	0.54	<0.001
Zero-Inflated Model				
Intercept	2.44E-06	0.17	0.04	0.806
Month 2	0.04	<0.001	0.41	0.045
Month 3	0.01	<0.001	0.27	0.004
Socio-economic				
Metropolitan area	0.12	0.002	1.68E-05	0.922
Pop. Density (log)	0.89	0.327	1.46	0.014
Prop. of males	3.07 E+09	<0.001	0.48	0.826
Income inequality (Gini)	0.08	0.655	7.17E-05	0.016
Prop. of people in poverty	129.78	0.13	8.51 E+04	<0.001
Health				
Prop. of obesity	0.03	0.322	1.57	0.886
Prop. of hypertension	1289.48	0.05	3.61 E+15	<0.001
Prop. of diabetes	0.07	0.748	1.98E-17	<0.001
Fast-food per 1000 people	1.06	0.282	1.07	0.082
Climate				
Avg. monthly max. temp. (log)	0.11	0.59	0.08	0.499
Avg. monthly min. temp. (log)	1.61	0.808	0.87	0.923
Monthly cumulative rainfall (log)	1.06	0.587	1.39	0.185
Annual PM 2.5 (log)	0.6	0.001	0.54	0.001
Commuting				
Public buses per 1000 people	1.01	0.824	0.01	0.001
Private cars per 1000 people	1.01	0.004	0.94	<0.01
Prop. of essential activities	40.81	0.164	0.59	0.521
Observations	7390		7390	

Table 4
Incidence rate ratios (IRR) of COVID-19 deaths.

	Phase I		Phase II	p-value
	IRR	p-value	*IRR*	
Count Model				
Intercept	0.01	0.068	1.37E-07	<0.001
Month 2	5.35	<0.001	2.66	<0.001
Month 3	20.58	<0.001	3.69	<0.001
Socio-economic				
Metropolitan area	1.32	0.014	1.05	0.360
Pop. Density (log)	1.37	<0.001	1.83	<0.001
Prop. of males	0.19	0.131	0.01	<0.001
Income inequality (Gini)	0.15	0.072	280.35	<0.001
Prop. of people in poverty	0.5	0.02	0.22	<0.001
Health				
Prop. of obesity	2.78	0.143	220.77	<0.001
Prop. of hypertension	0.38	0.391	0.01	<0.001
Prop. of diabetes	0.32	0.561	537.02	<0.001
Fast-food per 1000 people	1.01	<0.001	1.02	0.070
Climate				
Avg. monthly max. temp. (log)	0.32	0.096	19.24	<0.001
Avg. monthly min. temp. (log)	10.32	<0.001	0.74	0.041
Monthly cumulative rainfall (log)	1.001	0.972	0.88	<0.001
Annual PM 2.5 (log)	1.25	<0.001	1.96	<0.001
Commuting				
Public buses per 1000 people	0.98	<0.001	1.01	<0.001
Private cars per 1000 people	1.003	<0.001	1.002	0.001
Prop. of essential activities	0.26	<0.001	1.20	0.366
Zero-Inflated Model				
Intercept	39.26	0.597	1.23 E+32	0.487
Month 2	0.19	<0.001	0.38	0.115
Month 3	0.02	<0.001	0.33	0.072
Socio-economic				
Metropolitan area	0.13	<0.001	5.03E-09	0.935
Pop. Density (log)	0.88	0.205	6.22	<0.001
Prop. of males	896.22	0.09	2.13E-05	0.020
Income inequality (Gini)	5.18E-04	0.05	2.83	0.816
Prop. of people in poverty	5.5	0.047	1.10 E+06	<0.001
Health				
Prop. of obesity	0.01	0.005	1.52	0.908
Prop. of hypertension	42,347.92	<0.001	1.40 E+17	<0.001
Prop. of diabetes	1.99E-16	<0.001	2.10E-30	<0.001
Fast-food per 1000 people	0.88	0.003	0.63	<0.001
Climate				
Avg. monthly max. temp. (log)	0.11	0.239	0.02	0.421
Avg. monthly min. temp. (log)	16.97	0.002	5.88	0.282
Monthly cumulative rainfall (log)	0.94	0.407	0.49	0.012
Annual PM 2.5 (log)	0.45	<0.001	0.22	<0.001
Commuting				
Public buses per 1000 people	0.76	0.002	1.44	<0.001
Private cars per 1000 people	1.01	0.004	0.92	<0.001
Prop. of essential activities	2.3	0.279	2.94	0.430
Observations	7390		7390	

cases and deaths. For the count model results in Fig. 3, the estimated coefficient of population density (log) during Phase I equals 0.53 (0.31 for deaths), and it is statistically significant at the 99 % level. An alternative explanation for the estimated coefficients is the incidence rate ratios (*IRR*), also known as odds ratios in ZINB models, which are calculated by taking the exponent of the estimated coefficients in Fig. 3. The *IRR* of population density (log) (Tables 3 and 4) indicates that a unit increase in this variable increased it by 1.71 (= exp (0.53)) and 1.37 times among municipalities that reported cases of *COVID-19* infection and death, respectively, during Phase I. However, in Phase II, the magnitude of the estimated coefficient became larger in the case of deaths, indicating that more densely populated municipalities suffered from larger counts of fatalities. This result is also partially illustrated in Fig. 2. The finding agrees with recently published work using U.S. county-level data (Sun, Matthews, et al., 2020) or Iran province-level data (Ahmadi et al., 2020), but it is contrary to empirical evidence of the provincial regions in China (Sun, Zhang, et al., 2020).

Municipalities located in metropolitan areas were more-likely (*p-value* < 0.001) to suffer from infections in Phase II. Income inequality is revealed as the main factor associated with the spread of the virus and deaths. In Phase I, holding all other municipal-level variables at their mean values, positive cases increased by a factor of 4045.05 for each one-unit increase in the Gini index, whereas in Phase II, this factor increased to 10,899.25. Moreover, municipalities with disproportionately social inequalities (e.g., municipalities in the States of Oaxaca, Mexico, or Mexico City) suffered from a larger *IRR* of COVID-19 deaths. The result agrees with the evidence of the U.S. at the county-level (Millett et al., 2020) or Singapore at the planning area-level (Yi et al., 2021). The pronounced social disparities across states have shown that areas with the highest numbers of COVID-19 cases and deaths are the vulnerable geographical areas in terms of large social disadvantages.

Our findings also provide support for the strong correlation between poor air quality conditions at the municipal level and the COVID-19 prevalence and deaths. In Phase I (*Phase II*), an increase in 1 metric ton in PM2.5 (log) per year would increase the COVID-19 death rate by a factor of 1.25 (*1.96*). The finding agrees with relevant studies tackling the atmospheric determinants of population's vulnerability to COVID-19 in the U.S. (Bashir et al., 2020; Wu et al., 2020), Italy (Lubrano et al., 2020), China (Lin, Lau, et al., 2020), and some Latin American cities (Bolaño-Ortiz et al., 2020). All of them unanimously confirmed the correlation between atmospheric conditions and COVID-19 cases, making poor air quality an additional co-determinant of COVID-19 lethality. For precipitation, we observed that high cumulative rainfall (log) reduced virus transmission and deaths during Phase II only. The literature focusing on the role of the amount of rainfall and viral transmission/deaths, different from air pollution, has found inconclusive results. In some countries like the U.S. (Bashir et al., 2020), Indonesia (Tosepu et al., 2020), or Brazil (Rosario et al., 2020), rainfall is not associated with the dynamics of the virus.

The large magnitude of the coefficients associated with health-related explanatory variables during Phase I indicates that regions with a high prevalence of comorbidities, such as municipalities in Mexico, Sinaloa, Veracruz, or Yucatan, were hard hit by COVID-19. Note that during Phase I, comorbidities were not statistically associated with the cumulative count of deaths.

Lastly, commuting-related factors are weakly correlated with confirmed cases and deaths. For instance, after the social distancing was gradually lifted in Phase II, an increase in the proportion of essential activities was linked to a decrease (*IRR* = 0.54, *p-value* < 0.001) in the spread of the virus. The explanation for these apparently inconsistent effects is linked to the share of informal employment in Mexico, that is, 53.4 % in 2016 according to the International Labor Organization ILO (2018). Needless to say, the size of informal employment in the country is comparable to other Latin American nations (e.g., Argentina, 47.2 %; Brazil, 40.5 %; Chile, 40.1 %; and Ecuador, 59 %) or China (54.4 %).

4. Discussion

The spatial distribution of emerging spatial clusters of COVID-19 at the municipal level showed that after Phase I was lifted, the de-escalation process exhibited a rapid spread of the virus to northern and southern regions, particularly to those municipalities along the Mexico–U.S. border. We identified 23 clusters of COVID-19 infections during Phase I, and surprisingly, the total number of clusters increased to 40 in Phase II. Notwithstanding, we also found an opposite dynamic with respect to cumulative COVID-19 deaths. That is, the total number of clusters reduced from 24 to 16 between Phase I and Phase II, respectively. Moreover, their centroids moved away from Mexico City metropolitan area to Sonora and other bordering states. Although Phase II was characterized by a small number of clusters of COVID-19 deaths, they also reported a considerably larger number of cumulative deaths. For instance, the most-likely cluster during Phase I reported 1072 deaths, whereas it exhibited 3179 deaths in Phase II. Our spatial scan analysis illustrates the usefulness of the spatial scan statistic (SaTScan) as a tool for geographical disease surveillance.

Besides the physical restriction measures imposed by local and federal governments, other potential municipal-level factors associated with the dynamic of the disease, such as socio-economic inequalities, could be existing. The empirical evidence of Das et al. (2020) showed that COVID-19 spread grows exponentially in the Indian regions with enormous socio-economic deprivation. Similarly, the evidence comparing U.K. regions (Sun, Hu, & Xie, 2020) revealed important geographical disparities in risk and outcomes of the disease. That is, the risk of dying among those diagnosed with the infection was higher for those belonging to minority ethnic groups and those living in more deprived areas. Sannigrahi et al. (2020) also found strong evidence supporting the idea that spatial determinants of COVID-19 infections and deaths in Europe were mainly associated with poverty rates and income inequality.

Additional arguments explaining the geographical spread of the disease are presented by Geng et al. (2020), who found associations between the reopening of urban business centers in the U.S. and rapid spread of the virus across counties. Meanwhile, Sun, Matthews, et al. (2020) highlighted the noticeable spatial heterogeneity in the distribution of positive cases in the U.S., particularly when comparing central counties with those located at the east and west coasts. The authors illustrated a strong positive correlation between socio-economic factors (population density, social segregation, and differences by ethnicity) and infection rates. Moreover, Snyder and Parks (2020), Xie et al. (2020), and Karaye and Horney (2020) has also recently investigated the link between space–time variations in disease and socio-economic factors.

ZINB regression explicitly models the excess of zeros in the distribution of observed infections and fatalities. Hence, the results produced by the ZINB have enhanced the correlation analysis in Fig. 2. Recall that correlation analysis has been the most widely used technique to assess the association between climate or socio-economic variables and the number of cases (Briz-Redón & Serrano-Aroca, 2020). However, correlation analysis does not account for the possible presence of temporal trends, which can strongly affect the correlation value and yield artefactual associations. Our findings suggest that correlation analysis alone is not suitable to unveil the association between the outcome variables and potential predictors. For instance, Fig. 2 suggests strong evidence that fast-food restaurant per 1000 people was strongly and positively associated with infections and deaths in both phases. Nevertheless, after modeling excessive zeros, controlling by relevant factors, and including time trends, we determined that results from the econometric strategy do not support the notion that an increasing number of local fast-food restaurants is associated with higher rates of disease infections or lethality. However, a debate is ongoing on whether a higher density of unhealthy food options is associated with higher obesity and diabetes prevalence. Although the evidence is mixed (Frankenfeld et al., 2015;

Mejova et al., 2015), Mexico has recently implemented a front-of-pack labeling system to tackle obesity by reducing the purchase of unhealthy products.

Similarly, from Fig. 2, one can conclude that municipalities with high poverty rates tended to report fewer cases and deaths. Interestingly, in the early weeks of Phase I, this was the official discourse of some local government authorities by stating that this is a "rich person's disease," personified in groups of cosmopolitan, wealthy people who travel, become infected, and in turn infect others (Páez & Pérez, 2020). Estimates from the ZINB model do suggest a negative association between disease cases and poverty levels. However, the findings should be interpreted cautiously given the enormous effect of income inequality concentration on both infections and fatalities (see Tables 3 and 4 and Fig. 3). The result is consistent with other studies conducted in various Latin American countries, where income inequality is a general problem in this region. As documented by Bolaño-Ortiz et al., 2020, for some Latin American cities, including Mexico City, income inequality and poverty are positively associated with the spread of COVID-19.

All in all, the mechanisms in which social inequalities act on the spread and lethality of the virus could be related to the severe shortages in the healthcare sector of numerous Mexican municipalities, such as the lack of healthcare facilities in rural areas. This situation is placing great strain on the healthcare systems as patients must travel long distances to acquire treatments (Benita, 2019). Social inequalities among residents from urban areas could also contribute to the easy spread of this virus by having less fortunes with limited access to healthcare resources or the use of crowded public transportation with minimum safety standards and protocols to reduce viral transmission and spread. In fact, we observed a positive association between cumulative cases and public buses per 1000 people, which turned out to be significant at the 10 % and 1 % levels during Phases I and II, respectively.

In the context of Mexico, our results at the municipal-level indicate that, regardless of lockdown policies, metabolic disorders, such as obesity or diabetes, can exacerbate infections and deaths. Our findings agree with the findings of Lubrano et al. (2020), who demonstrated that the extremely high level of COVID-19 lethality in Northern Italy, compared to the rest of the country, can be explained by differences in air pollution and obesity. Sanchis-Gomar et al. (2020) also supported the above-mentioned link between obesity/air pollution and higher COVID-19 lethality. The authors detailed how adipose tissue may be vulnerable to COVID-19 infection; therefore, obese patients also have worse outcomes with COVID-19 infection, including respiratory failure, need for mechanical ventilation, and higher mortality (Almerie & Kerrigan, 2020; Flint & Tahrani, 2020; Palaiodimos et al., 2020). Similarly, diabetes has been associated with increased mortality risk (Papadokostaki et al., 2020; Pitocco et al., 2020) because people with diabetes could suffer from a loss of capacity to regulate immunity. With this body of knowledge pointing to the links between comorbidities associated with obesity and COVID-19 complications, the scenario in Mexico is not promising as the diabetes prevalence reached 15 % and figures of obesity and overweight surpass 60 % (OECD (Organisation for Economic Co-operation and Development), 2019).

Concerning the group of climate predictors, results from the ZINB regression indicate that low temperatures are positively correlated with deaths caused by COVID-19 during Phase I but not during Phase II. Note that such association is not apparent under the correlation analysis of Fig. 2. This finding could be explained by the fact that in the first months of the pandemic (i.e., March to May 2020), Mexico experienced its 2019–2020 winter, whereas Phase II coincided with the onset of the spring season. Remarkably, the evidence (see Briz-Redón and Serrano-Aroca (2020) and the references therein) tends to agree that low temperatures relate to the rapid spread of the virus. The mechanism behind this dynamic is that COVID-19 virus can stay alive for extended periods not only in cold, humid conditions but also in cold weather that favors indoor social activities, thereby posing additional driving forces of the contagion velocity. Lin, Wei, et al. (2020) have found similar results in the context of Chinese provinces. They pointed out that the transmission of the virus was higher in colder provinces.

Pollution level, proxied by the log of annual PM 2.5 concentration, was also one of the environmental variables with statistically significant effects on both infections and deaths. Particularly, we highlight the positive incidence rate ratios in deaths during and after social distancing. Here, our findings agree with the recently published literature emphasizing the critical role of environmental pollution in the quick-paced spread. Lin, Lau, et al. (2020) found that provinces in China with large amounts of carbon monoxide also reported high virus transmissibility. Moreover, Bolaño-Ortiz et al., 2020 considered PM2.5 measurements among major Latin American cities and reported similar results; that is, a strong link exists between air pollution and the spread of positive COVID-19 cases.

5. Concluding remarks

This study aimed to understand the influence of several socio-economic, health, climate, and mobility factors on the spatial distribution of COVID-19 transmission and deaths during the first wave of the epidemic in Mexico. We used publicly available municipal-level case data and geospatial information from multiple sources to perform spatial and statistical analyses. Emerging clusters of transmissions and fatalities were observed through the spatial scanning statistical analysis. Likewise, the econometric strategy revealed a significant influence of some socio-economic (e.g., income inequality), health (prevalence of obesity and hypertension), and climate (air temperature and air pollution) factors on this viral disease in Mexico.

Several policy recommendations can be drawn from the detection of spatial clusters of infections and deaths as policymakers and healthcare authorities could target regions that must be prioritized for interventions. For instance, substantial overlapping of secondary clusters (Fig. 1(a)) indicates an urgent need for regional measures for prevention, containment, and coordination to be implemented in combating the COVID-19 health crisis. Successful examples of localized policy interventions are detailed by Leal-Neto et al. (2020) and Dlamini et al. (2020) for the case of Brazil and Eswatini (Africa), respectively.

Likewise, some spatial clusters identified in this work are the municipalities from different states that are grouped together. This scenario is relevant while designing localized public health efforts. In Mexico, the federal government has delegated much of the responsibility for managing the public health system during the pandemic to the Federal States. The Federal States also have main duties and responsibilities related to the restriction of non-essential economic activities when the Epidemiological Risk Traffic Light (a bi-weekly monitoring system used to monitor and grade the use of public space according to virus transmission) turns into red.

In the context of emerging clusters of municipalities from different states, coordinated efforts among various government levels are critical. Past published research has shown the advantages of this type of coordinated public policy intervention not only to prevent the spread of sexual transmission diseases (Marotta, 2016) or Salmonella (Seixas et al., 2018) but also to provide better healthcare services to vulnerable populations suffering from cleft lip and palate (Gasca-Sanchez et al., 2019), obesity (Gartner et al., 2016), tuberculosis (Rao et al., 2017), and so forth.

Meanwhile, some experts say that the hardest hit countries had an aging population (Gardner et al., 2020), underdeveloped healthcare systems (Tanne et al., 2020), or unfavorable natural environment (Di Marco et al., 2020). However, unresolved structural problems of high poverty and income inequality appeared to have aggravated the spread of the pandemic in Mexico. Unless major policy changes are implemented to improve environmental sustainability and promote human development, more disadvantaged Mexican municipalities will become hotspots for virus circulation and reemergence. Making a trade-off between socio-economic development and infection control is a major

challenge facing not only in Mexico but in many other developing countries (Yin et al., 2021).

Some of the possible policy actions could include support for workers in the informal sector, assistance for the population in poverty, a decarbonization pathway that focuses on electrification and clean renewable energy, optimize the ownership and use of natural resources, or devolve rights and duties for ecosystem regulation to individuals or communities. Interestingly, the recent front-of-pack labeling system based on warning labels of unhealthy products seems to be an adequate government's strategy to fight obesity and ultimately COVID-19.

Finally, it is vital for government to continue with the important efforts of testing procedures, tracking of individuals, social distancing measures, and plans for hospital reconversion and immediate expansion to mitigate the dynamics of disease spread. Furthermore, sudden outbreaks can be brought under control with quick and decisive action supported by Geographic Information Systems and Big Data technologies. Data acquisition methods, and integration of data sources from various organizations can be used to develop COVID-19 geo-located prediction systems based on real-time operational data assimilation and parameter estimation (Li et al., 2020; Zhou et al., 2020). This study, therefore, may be useful for decision-making on health policy at the municipal-level as it not only allows the prompt detection of hotspots of transmission and death but also unveils its relationship with socio-economic, health, climate, and mobility factors.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Ahmadi, M., Sharifi, A., Dorosti, S., Ghoushchi, S. J., & Ghanbari, N. (2020). Investigation of effective climatology parameters on COVID-19 outbreak in Iran. *The Science of the Total Environment, 729*, 138705.

Almerie, M. Q., & Kerrigan, D. D. (2020). The association between obesity and poor outcome after COVID-19 indicates a potential therapeutic role for montelukast. *Medical Hypotheses, 143*, 109883.

Bambra, C., Riordan, R., Ford, J., & Matthews, F. (2020). The COVID-19 pandemic and health inequalities. *Journal of Epidemiology & Community Health, 74*, 964–968.

Bashir, M. F., Bilal, B. M., & Komal, B. (2020). Correlation between environmental pollution indicators and COVID-19 pandemic: A brief study in californian context. *Environmental Research, 187*, 109652.

Benita, F. (2019). A new measure of transport disadvantage for the developing world using free smartphone data. *Social Indicators Research, 145*(1), 415–435.

Benita, F. (2021). Human mobility behavior in COVID-19: A systematic literature review and bibliometric analysis. *Sustainable Cities and Society, 70*, 102916.

Bolaño-Ortiz, T. R., Camargo-Caicedo, Y., Puliafito, S. E., Ruggeri, M. F., Bolaño-Diaz, S., Pascual-Flores, R., Saturno, J., Ibarra-Espinosa, S., Mayol-Bracero, O. L., Torres-Delgado, E., & Cereceda-Balic, F. (2020). Spread of SARS-CoV-2 through Latin America and the caribbean region: A look from its economic conditions, climate and air pollution indicators. *Environmental Research, 191*, 109938.

Briz-Redón, Á., & Serrano-Aroca, Á. (2020). The effect of climate on the spread of the COVID-19 pandemic: A review of findings, and statistical and modelling techniques. Progress in physical geography: *Earth and Environment, 44*(5), 591–604.

Das, A., Ghosh, S., Das, K., Basu, T., Dutta, I., & Das, M. (2020). Living environment matters: Unravelling the spatial clustering of COVID-19 hotspots in Kolkata megacity, India. *Sustainable Cities and Society*. https://doi.org/10.1016/j.scs.2020.102577

Dávila-Cervantes, C. A., & Agudelo-Botero, M. (2019). Health inequalities in Latin America: Persistent gaps in life expectancy. *The Lancet Planetary Health, 3*(12), e492–e493.

Desjardins, M., Hohl, A., & Delmelle, E. (2020). Rapid surveillance of COVID-19 in the United States using a prospective space-time scan statistic: Detecting and evaluating emerging clusters. *Applied Geography, 118*, 102202.

Di Marco, M., Baker, M. L., Daszak, P., De Barro, P., Eskew, E. A., Godde, C. M., Harwood, T. D., Herrero, M., Hoskins, A. J., Johnson, E., & Karesh, W. B. (2020). Opinion: Sustainable development must account for pandemic risk. *Proceedings of the National Academy of Sciences, 117*(8), 3888–3892.

Dlamini, W. M., Dlamini, S. N., Mabaso, S. D., & Simelane, S. P. (2020). Spatial risk assessment of an emerging pandemic under data scarcity: A case of COVID-19 in Eswatini. *Applied Geography, 125*, 102358.

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., et al. (2013). Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography, 36*(1), 27–46.

Flint, S. W., & Tahrani, A. A. (2020). COVID-19 and obesity—lack of clarity, guidance, and implications for care. *The Lancet Diabetes & Endocrinology, 8*(6), 447–475.

Frankenfeld, C. L., Leslie, T. F., & Makara, M. A. (2015). Diabetes, obesity, and recommended fruit and vegetable consumption in relation to food environment subtypes: A cross-sectional analysis of behavioral risk factor surveillance system, United States census, and food establishment data. *BMC Public Health, 15*(1), 491.

Gardner, W., States, D., & Bagley, N. (2020). The coronavirus and the risks to the elderly in long-term care. *Journal of Aging & Social Policy, 32*, 1–6.

Gartner, D. R., Taber, D. R., Hirsch, J. A., & Robinson, W. R. (2016). The spatial distribution of gender differences in obesity prevalence differs from overall obesity prevalence among US adults. *Annals of Epidemiology, 26*(4), 293–298.

Gasca-Sanchez, F. M., Santos-Guzman, J., Elizondo-Dueñaz, R., Mejia-Velazquez, G. M., Ruiz-Pacheco, C., Reyes-Rodriguez, D., Vazquez-Camacho, E., Hernandez-Hernandez, J. A., Lopez-Sanchez, R. D. C., Ortiz-Lopez, R., & Olvera-Posada, D. (2019). Spatial clusters of children with cleft lip and palate and their association with polluted zones in the monterrey metropolitan area. *International Journal of Environmental Research and Public Health, 16*(14), 2488.

Geng, X., Gerges, F., Katul, G. G., Bou-Zeid, E., Nassif, H., & Boufadel, M. C. (2020). Population agglomeration is a harbinger of the spatial complexity of COVID-19. *Chemical Engineering Journal*. https://doi.org/10.1016/j.cej.2020.127702

Hebbali, A. (2020). *olsrr: Tools for building OLS regression models*. Cran R.

Hernández-Garduño, E. (2020). Obesity is the comorbidity more strongly associated for Covid-19 in Mexico. A case-control study. *Obesity Research & Clinical Practice, 14*, 375–379.

Hessel, P., Sayer, P., & Ruimallo-Herl, C. (2019). Educational inequalities in disability linked to social security coverage among older individuals in five Latin American countries. *Social Science & Medicine, 267*, 112378.

ILO (International Labour Organization). (2018). *Women and men in the informal economy: A statistical picture* (3rd ed.). Geneva: ILO.

Jackman, S., Tahk, A., Zeileis, A., Maimone, C., Fearon, J., Meers, Z., Simon, M., & Imports, M. A. S. S. (2015). Package 'pscl'. *Political Science Computational Laboratory*, 18(04.2017).

Jiao, J., Moudon, A. V., Kim, S. Y., Hurvitz, P. M., & Drewnowski, A. (2015). Health implications of adults' eating at and living near fast food or quick service restaurants. *Nutrition & Diabetes, 5*(7), e171.

Karaye, I. M., & Horney, J. A. (2020). The impact of social vulnerability on COVID-19 in the U.S.: An analysis of spatially varying relationships. *American Journal of Preventive Medicine, 59*(3), 317–325.

Kihal-Talantikite, W., Padilla, C. M., Lalloue, B., Rougier, C., Defrance, J., Zmirou-Navier, D., & Deguen. (2013). An exploratory spatial analysis to assess the relationship between deprivation, noise and infant mortality: An ecological study. *Environmental Health, 12*(109), 1–15.

Kim, S., & Castro, M. C. (2020). Spatiotemporal pattern of COVID-19 and government response in South Korea (as of May 31, 2020). *International Journal of Infectious Diseases, 98*, 328–333.

Kulldorf, M. (1997). A spatial scan statistic. Communications in statistics: *Theory and Methods, 26*(6), 1481–1496.

Kulldorff, M. (2011). *SatScan Software for spatial, temporal and space-time scan statistics*. version 9.4.2 ([software]).

Leal-Neto, O. B., Santos, F. A. S., Lee, J. Y., Albuquerque, J. O., & Souza, W. V. (2020). Prioritizing COVID-19 tests based on participatory surveillance and spatial scanning. *International Journal of Medical Informatics, 143*, 104263.

Lin, C., Lau, A. K., Fung, J. C., Guo, C., Chan, J. W., Yeung, D. W., Zhang, Y., Bo, Y., Hossain, M. S., Zeng, Y., & Lao, X. Q. (2020). A mechanism-based parameterisation scheme to investigate the association between transmission rate of COVID-19 and meteorological factors on plains in China. *The Science of the Total Environment, 737*, 140348.

Lin, S., Wei, D., Sun, Y., Chen, K., Yang, L., Liu, B., Huang, Q., Paoliello, M., Li, H., & Wu, S. (2020). Region-specific air pollutants and meteorological parameters influence COVID-19: A study from mainland China. *Ecotoxicology and Environmental Safety, 204*, 111035.

Li, X., Zhao, Z., & Liu, F. (2020). Big data assimilation to improve the predictability of COVID-19. *Geography and Sustainability, 1*(4), 317–320.

Lubrano, C., Risi, R., Masi, D., Gnessi, L., & Colao, A. (2020). Is obesity the missing link between COVID-19 severity and air pollution? *Environmental Pollution, 266*(Part 3), 115327.

Macip, R. F. (2020). The party is over: Cracking under Sana Distancia in Mexico. *Dialectical Anthropology, 44*(3), 243–250.

Marotta, P. (2016). Assessing spatial relationships between race, inequality, crime, and gonorrhea and chlamydia in the United States. *Journal of Urban Health, 94*(5), 683–698.

Ma, Y., Zhao, Y., Liu, J., He, X., Wang, B., Fu, S., Yan, J., Niu, J., Zhou, J., & Luo, B. (2020). Effects of temperature variation and humidity on the death of COVID-19 in Wuhan, China. *The Science of the Total Environment, 724*, 138226.

McLaren, J. (2021). Racial disparity in COVID-19 deaths: Seeking economic roots with census data. *The B.E. Journal of Economic Analysis & Policy*. https://doi.org/10.1515/bejeap-2020-0371

Mejova, Y., Haddadi, H., Noulas, A., & Weber, I. (2015). # foodporn: Obesity patterns in culinary interactions. In *Proceedings of the 5th international conference on digital health 2015* (pp. 51–58).

Méndez-Arriaga, F. (2020). The temperature and regional climate effects on communitarian COVID-19 contagion in Mexico throughout phase 1. *The Science of the Total Environment, 735*, 139560.

Mesenburg, M. A., Restrepo-Mendez, M. C., Amigo, H., Balandrán, A. D., Barbosa-Verdun, M. A., Caicedo-Velásquez, B., Carvajal-Aguirre, L., Coimbra, C. E., Jr., Ferreira, L. Z., del Pilar Flores-Quispe, M., & Flores-Ramírez, C. (2018). Ethnic group

inequalities in coverage with reproductive, maternal and child health interventions: Cross-sectional analyses of national surveys in 16 Latin American and caribbean countries. *The Lancet Global Health, 6*(8), e902–e913.

Millett, G. A., Jones, A. T., Benkeser, D., Baral, S., Mercer, L., Beyrer, C., Honermann, B., Lankiewicz, E., Mena, L., Crowley, J., Sherwood, J., & Sullivan, P. (2020). Assessing differential impacts of COVID-19 on Black communities. *Annals of Epidemiology, 47*, 37–44.

Oecd (Organisation for Economic Co-operation and Development). (2019). The heavy burden of obesity: The economics of prevention. In *OECD health policy studies*. Paris: OECD Publishing.

Páez, D., & Pérez, J. A. (2020). Social representations of COVID-19 (Representations sociales del COVID-19). *International Journal of Social Psychology, 35*(3), 600–610.

Palaiodimos, L., Kokkinidis, D. G., Li, W., Karamanis, D., Ognibene, J., Arora, S., Southern, W. N., & Mantzoros, C. S. (2020). Severe obesity is associated with higher in-hospital mortality in a cohort of patients with COVID-19 in the Bronx, New York. *Metabolism Clinical and Experimental, 108*, 154262.

Papadokostaki, E., Tentolouris, N., & Liberopoulos, E. (2020). COVID-19 and diabetes: What does the clinician need to know? *Primary Care Diabetes, 14*(5), 558–563.

Pitocco, D., Viti, L., Tartaglione, L., Di Leo, M., Rizzo, G. E., Manto, A., Rizzi, A., Caputo, S., & Pontecorvi, A. (2020). Diabetes and severity of COVID-19: What is the link? *Medical Hypotheses, 143*, 109923.

Popkin, B., Du, S., Green, W. D., Beck, M., Algaith, T., Herbst, C., Alsukait, R., Alluhidan, M., Alazemi, N., & Shekar, M. (2020). Individuals with obesity and COVID-19: A global perspective on the epidemiology and biological relationships. *Obesity Reviews, 21*(11), e13128.

Rao, H., Shi, X., & Zhang, X. (2017). Using the Kulldorff's scan statistical analysis to detect spatio-temporal clusters of tuberculosis in Qinghai Province, China, 2009-2016. *BMC Infectious Diseases, 17*(578), 1–11.

Richardson, A. S., Meyer, K. A., Howard, A. G., Boone-Heinonen, J., Popkin, B. M., Evenson, K. R., Shikany, J. M., Lewis, C. E., & Gordon-Larsen, P. (2015). Multiple pathways from the neighborhood food environment to increased body mass index through dietary behaviors: A structural equation-based analysis in the cardia study. *Health & Place, 36*, 74–87.

Ripley, B., Venables, B., Bates, D. M., Hornik, K., Gebhardt, A., Firth, D., & Ripley, M. B. (2013). *Package 'mass', cran R*.

Rosario, D. K., Mutz, Y. S., Bernardes, P. C., & Conte-Junior, C. A. (2020). Relationship between COVID-19 and weather: Case study in a tropical country. *International Journal of Hygiene and Environmental Health, 229*, 113587.

Sanchis-Gomar, F., Lavie, C. J., Mehra, M. R., Henry, B. M., & Lippi, G. (2020). Obesity and outcomes in COVID-19: When an epidemic and pandemic collide. *Mayo Clinic Proceedings, 95*(7), 1445–1453.

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*. https://doi.org/10.1016/j.scs.2020.102418

Seixas, R., Nunes, T., Machado, J., Tavares, L., Owen, S. P., Bernardo, F., & Oliveira, M. (2018). Demographic characterization and spatial cluster analysis of human Salmonella 1, 4,[5], 12: i:-infections in Portugal: A 10 year study. *Journal of Infection and Public Health, 11*(2), 178–182.

Snyder, B. F., & Parks, V. (2020). Spatial variation in socio-ecological vulnerability to Covid-19 in the contiguous United States. *Health & Place, 66*, 102471.

Sun, Y., Hu, X., & Xie, J. (2020). Spatial inequalities of COVID-19 mortality rate in relation to socioeconomic and environmental factors across England. *The Science of the Total Environment*. https://doi.org/10.1016/j.scitotenv.2020.143595

Sun, F., Matthews, S., Yang, T. C., & Hu, M. (2020). A spatial analysis of COVID-19 period prevalence in US counties through June 28, 2020: Where geography matters? *Annals of Epidemiology, 52*, 54–59.

Sun, Z., Zhang, H., Yang, Y., Wan, H., & Wang, Y. (2020). Impacts of geographic factors and population density on the COVID-19 spreading under the lockdown policies of China. *The Science of the Total Environment, 746*, 141347.

Tanne, J. H., Hayasaki, E., Zastrow, M., Pulla, P., Smith, P., & Rada, A. G. (2020). Covid-19: How doctors and healthcare systems are tackling coronavirus worldwide. *BMJ*, 368.

Tosepu, R., Gunawan, J., Effendy, D. S., Lestari, H., Bahar, H., & Asfian, P. (2020). Correlation between weather and covid-19 pandemic in jakarta, Indonesia. *The Science of the Total Environment, 725*, 138436.

Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica*, 307–333.

Walton, M., Pearce, J., & Day, P. (2009). Examining the interaction between food outlets and outdoor food advertisements with primary school food environments. *Health & Place, 15*(3), 841–848.

Wu, X., Nethery, R. C., Sabath, M. B., Braun, D., & Dominici, F. (2020). Air pollution and COVID-19 mortality in the United States: Strengths and limitations of an ecological regression analysis. *Science Advances, 6*(45). eabd4049.

Xie, Z., Qin, Y., Li, Y., Shen, W., Zheng, Z., & Liu, S. (2020). Spatial and temporal differentiation of COVID-19 epidemic spread in mainland China and its influencing factors. *The Science of the Total Environment, 744*, 140929.

Yi, H., Ng, S. T., Farwin, A., Pei Ting Low, A., Chang, C. M., & Lim, J. (2021). Health equity considerations in COVID-19: Geospatial network analysis of the COVID-19 outbreak in the migrant population in Singapore. *Journal of Travel Medicine, 28*(2). taaa159.

Yin, C., Zhao, W., Cherubin, F., & Pereira, P. (2021). Integrate ecosystem services into socio-economic development to enhance achievement of sustainable development goals in the post-pandemic era, Geography and Sustainability. *Geography and Sustainability., 2*(1), 68–73.

Zhou, C., Su, F., Pei, T., Zhang, A., Du, Y., Luo, B., Cao, Z., Wang, J., Yuan, W., Zhu, Y., Song, C., Chen, J., Xu, J., Li, F., Ma, T., Jiang, L., Yan, F., Yi, J., Hu, Y., Lia, Y., & Xiao, H. (2020). COVID-19: Challenges to GIS with Big data. *Geography and Sustainability., 1*(1), 77–87.