Socioeconomic determinants of pandemics: a spatial methodological approach with evidence from COVID-19 in Nice, France

Laurent Bailly,[1] Rania Belgaied,[2] Thomas Jobert,[3] Benjamin Montmartin[4]

[1]Department of Public Health, Centre Hospitalier Universitaire de Nice, Recherche en Santé des Populations Evaluation Clinique et Thérapeutique, (RESPECT) Unité de Recherche Clinique Côte d'Azur (UR2CA), Université Côte d'Azur, Nice; [2]Departement of Economics, Université Côte d'Azur, Groupe de Recherche en Droit, Economie, Gestion (GREDEG), Unité de Recherche Clinique Côte d'Azur (UR2CA), Nice; [3]Université Côte d'Azur, Centre National de la Recherche Scientifique (CNRS), Groupe de Recherche en Droit, Economie, Gestion (GREDEG), France; [4]SKEMA Business School, Université Côte d'Azur ,Groupe de Recherche en Droit, Economie, Gestion (GREDEG), Observatoire Français des Conjonctures Économiques, Nice, France

Abstract

During the period 4 January 4 – 14 February 2021 the spread of the COVID-19 epidemic peaked in the city of Nice, France with a worrying number of infected cases. This article focuses on analyzing the explicit, spatial pattern of virus spread and assessing the geographical factors influencing this distribution. Spatial modelling was carried out to examine geographical disparities in terms of distribution, incidence and prevalence of the virus, while taking socio-economic factors into account. A multiple linear regression model was used to identify the key socio-economic variables. Global and local spatial autocorrelation were measured using Moran and LISA indices, followed by spatial autocorrelation analysis of the residuals. Similarly, we used the Geographically Weighted Regression (GWR) model and the Multiscale Geographically Weighted Regression (MGWR) model to assess the influence of socio-economic factors that vary on a global and local scale. Our results reveal a marked geographical polarization, with affluent areas in the Southeast of the city contrasting sharply with disadvantaged neighbourhoods in the Northwest. Neighbourhoods with low Localized Human Development Index (LHDI), low levels of education, social housing and immigrant populations all pointed to worrying values. On the other hand, people who use public transport were significantly more likely to be contaminated by the virus. These results underline the importance of geographically predicting COVID-19 distribution patterns to guide targeted interventions and health policies. Understanding these spatial patterns using models such as MGWR can help guide public health interventions and inform future health policies, particularly in the context of pandemics.

Correspondence: Rania Belgaied, Departement of Economics, Université Côte d'Azur, Groupe de Recherche en Droit, Economie, Gestion (GREDEG), Unité de Recherche Clinique Côte d'Azur (UR2CA), Nice, France
E-mail address: rania.belgaied@etu.unice.fr

Key words: COVID-19, spatial analysis, spatial autocorrelation, public health, Geographic Information System (GIS); France.

Conflict of interest: the authors declare no potential conflict of interest, and all authors confirm accuracy.

Availability of data and materials: all data generated or analyzed during this study are included in this published article.

Acknowledgments: the authors would like to thank Dr. Christian Pradier and Dr. Eugènia Mariné-Barjoan for their assistance in providing data from SI-DEP, as well as the Public Health Department's team.

Received: 28 February 2025.
Accepted: 23 June 2025.

Geospatial Health 2025; 20:1383
doi:10.4081/gh.2025.1383

Introduction

The COVID-19 health crisis revealed the critical influence of environmental and social factors on population health. According to Pinter-Wollman *et al.* (2018), the built environment plays a key role in both the prevention of chronic diseases and the spread of infectious diseases. Measures imposed during the lockdown to mitigate viral transmission elevated housing conditions to a central concern in public health discourse. This is not surprising as housing is a multidimensional concept that extends beyond the physical dwelling; it encompasses the household and the wider neighbourhood environment, all of which can affect health outcomes. The effects of housing on health can be analysed along three dimensions: the dwelling itself, the household it accommodates, and the broader neighbourhood setting. This multidimensional framework underpins our analysis of COVID-19 incidence patterns.

Stiglitz's observation that "COVID is not an equal opportunity killer" is true in terms of both mortality rates (Schellekens & Sourrouille, 2020; Shahbazi & Khazaei, 2020; Revollo-Fernández *et al.*, 2022) and incidence rates (Ruthberg *et al.*, 2020; Etowa & Hyman, 2021). Early research on the pandemic focused on epidemiological and clinical (Ruthberg *et al.*, 2020; Xu *et al.*, 2020), environmental (Luo *et al.*, 2021), demographic (Patterson *et al.*, 2022) and ecological characteristics of infected patients. However, as the crisis evolved, other potential determinants, such as socio-economic inequalities emerged as key explanatory factors in the spread and impact of COVID-19 (Benita *et al.*, 2022; Dowrick *et*

al., 2022; Godefroy & Lewis, 2022). Schellekens and Sourrouille (2020) described COVID-19 as a missile targeting the most vulnerable segments of society, while Shahbazi and Khazaj (2020), using the Human Development Index (HDI) as an indicator of the incidence and mortality rate of COVID-19, found that developed countries were the most affected. Similarly, Etowa *et al.* (2021) have shown that the risk and burden of COVID-19 infection are not evenly distributed between population subgroups. Bambra *et al.* (2020) interpreted the emergence of these inequalities through a syndemic lens, highlighting the synergistic interaction between COVID-19 incidence rate, the socio-environmental and the socio-economic factors that exacerbate vulnerability. In this framework, socio-economic deprivation emerges as a key driver of COVID-19transmission (Rohleder *et al.*, 2022).

Early research on the COVID-19 pandemic largely relied on statistical and biomedical approaches, focusing on individual risk factors and national-level indicators. However, as Chossegros (2020) argues, an alternative reading of the pandemic is possible. This interpretation highlights the territorial aspect of health inequalities and underscores the influence of structural disparities in determining exposure and vulnerability. It highlights the need for spatially informed analyses that go beyond aggregate figures and consider the localized dynamics of the disease.

The importance given to spatial analysis, considering location, spatial interaction, spatial structure and spatial processes is well established across disciplines such as epidemiology, econometrics and environmental science. In the field of economics, Patel *et al.* (2020) and Fontán-Vela *et al.* (2023) emphasized that during the COVID-19 pandemic, exposure risk and disease severity varied significantly according to factors such as housing overcrowding, working conditions, living or housing conditions, education level and income. These findings underscore the necessity of considering contextual and socio-spatial factors when analysing health outcomes. Household-level characteristics, such as overcrowding, household composition, and migrant status (Siljander *et al.*, 2022), play a pivotal role in determining vulnerability. At the regional level, poverty (Talavera & Perez, 2009) and the quality of the healthcare system (Scott & Coote, 2007) also help explain differences between regions. In addition, demography is a factor in spatial differentiation in health (Hu & Goldman, 1990), whether national or regional (Frohlich & Mustard, 1996). Education levels and the organization of the healthcare system further compound these inequalities, reinforcing the need for a geographically nuanced approach to pandemic analysis. The spatial distribution of COVID-19 cases has revealed significant clustering patterns, particularly in urban environments. These clusters not only reflect the transmission dynamics of the virus but also reveal underlying socio-economic and environmental inequalities. To better understand these localized disparities, a geographically detailed approach is required; one that captures spatial variations in incidence rates and the structural factors influencing them. To do so, we relied on the level of 'Ilots Regroupés pour l'Information Statistique'(IRIS), which is the smallest statistical unit used by the French National Institute of Statistics and Economic Studies (INSEE) (https://www.insee.fr/fr/metadonnees/definition/c1523 - accessed25/3/2024).

Working at the IRIS level addresses a frequent limitation in the literature: the use of spatial units that are too large or internally heterogeneous, such as municipalities or regions. These larger units often mask important intra-urban disparities. The IRIS mesh offers geographic granularity and demographic consistency, enabling more accurate assessments of localized health inequalities. In addition, testing access was not a barrier in this context: the city of Nice had 198 testing centres, ensuring equitable access to screening during the study period. This minimizes potential biases related to underdetection or unequal access to health services. Nice provided a particularly relevant case, having experienced a significant outbreak between January and February 2021. This time frame allowed for the analysis of spatial disparities under consistent socio-sanitary conditions, before major policy interventions were introduced.

The objective of this study was to explore how socio-economic conditions relate to the spatial distribution of COVID-19 incidence across the 144 IRIS units of Nice. To achieve this, we integrated spatial interactions and accounted for local heterogeneity through a combination of global and local spatial econometric models.

Materials and Methods

Study area and timeframe

According to the 2019 census (https://www.insee.fr/fr/information/6444222 - accessed 26/8/2024), Nice had a population of 341,003 inhabitants, spread across a surface area of 71.92 km², resulting in a population density of 4,741 inhabitants per km². Socio-economic data from INSEE statistics (https://www.insee.fr/fr/statistiques/2011101?geo=COM-06088#chiffre-cle-5 - accessed 14/11/2023) indicate that over 30% of tenant households live below the poverty line, while the unemployment rate among young people aged 15 to 24 years exceeds 25%. In this study, we used IRIS as the spatial units of analysis, each typically encompassing around 2,000 inhabitants. From the 146 IRIS units in Nice, we excluded two non-residential units (central train station and airport), yielding a final sample of 144.

The study period covered the six-week period of 4 January – 14 February 2021, during which Nice experienced a sharp increase in COVID-19 incidence, reaching an average rate of 463.5 cases per 100,000 inhabitants, exceeding the national alert threshold of 400 cases per 100,000 as defined by public health authorities (Ministère du travail, 2024). This period corresponds to the early phase of the third COVID-19 wave in France, marked by the emergence of more transmissible variants and the time before mass vaccination campaigns were deployed or major restrictions implemented. Indeed, the first containment measures in the Alpes-Maritimes (such as partial weekend lockdowns) began as late as on February 27, 2021 (Vie Publique, 2021), This timeframe provided a temporally and territorially homogeneous context. No major disparities in access to testing were observed, as the city offered 198 testing locations covering its full territory. In total, 10,078 cases were included in the study.

Type of data and information sources

Our database combined epidemiological, socio-economic, socio-demographic and socio-environmental data. Epidemiological data on the COVID-19 incidence rate were provided by the "Système National d'Information sur le dépistage de la COVID-19" (SIDEP) (https://www.santepubliquefrance.fr/presse/2020/systeme-de-surveillance-du-covid-19-pour-une-analyse-territoriale-contextualisee-de-la-dynamique-epidemique - Accessed 24/8/2023) and referred to residents of the city of Nice

with a first positive result based on the Polymerase Chain Reaction (PCR). Only sporadic cases were included, excluding individuals living in institutions for dependent, elderly people. The incidence rate was used as the primary health indicator to monitor the epidemic, based on data published by "Santé Publique France" (SPF), the national public health agency responsible for epidemic surveillance. The remaining variables were retrieved from INSEE (https://www.insee.fr/fr/information/6444222 - accessed 26/8/2024) databases and covered 343 geolocated indicators across multiple dimensions.

Selection of variables

In a first step, we applied the 'Least Absolute Shrinkage and Selection Operator'(LASSO) regression method with penalization based on Akaike Information Criterion (AIC) to reduce this initial pool. This led to 37 pre-selected 37 candidate variables that were categorized into six groups: income/inequality; housing conditions; population density; contamination through work; understanding of health rules &cultural factors; and contamination through the school environment. Five socio-economic variables were ultimately selected and presented (Figure 1). The selection process was guided by a review of the literature and the availability of data, To further validate the selection and reduce multicollinearity, we conducted multiple linear regressions using Ordinary Least Squares (OLS). This two-step approach ensured the identification of statistically robust and interpretable variables while avoiding overfitting. The final set of variables included the Local Human Development Index (LHDI) that varies from 0 to 1 and is based on the combination of three factors: life expectancy at birth, level of education and gross national income per inhabitant (Shahbazi & Khazaei, 2020; Schellekens & Sourrouille, 2020; Etowa *et al.*, 2021; Revollo-Fernández *et al.*, 2022); the proportion of the population living in low-income housing (Brun & Simon, 2020; Schellekens & Sourrouille, 2020; Etowa & Hyman, 2021; Luo *et al.*, 2021; Nazia *et al.*, 2022); the proportion of employed people using public transport (Pinter-Wollman *et al.*, 2018; Brun & Simon, 2020; Cordes & Castro, 2020); the immigrant population (Brun & Simon, 2020; Nazia *et al.*, 2022; Rohleder *et al.*, 2022; Scarpone *et al.*, 2020; Vang & Ng, 2023; Vilinová & Petrikovičová, 2023); and the proportion of people aged between 18 and 24 years attending school (Cordes & Castro, 2020; Kim & Castro, 2020; Patterson *et al.*, 2022).

Approach

We adopted an exploratory two-phase approach, the first of which focused on the identification of spatial dependence of the data (Le Gallo, 2002) by analyzing spatial autocorrelation patterns. The second phase explored the spatially varying relationships between selected socio-economic variables and COVID-19 incidence rates using Geographically Weighted Regression (GWR) and its Multiscale Extension (MGWR). In order to ensure comparability across models and control for differences in variable scales, all variables (dependent and explanatory) were standardized using z-score transformation prior to estimation. Each variable was centred and scaled (mean of 0, Standard Deviation - SD of 1), including the COVID-19 incidence rate (measured as the number of cases per 100,000 inhabitants). This approach allowed all model coefficients to be interpreted as the expected change (in SDs) of the dependent variable resulting from one SD change in the explanatory variable. Standardization was also necessary to ensure consistency with the MGWR 2.2 software, which automatically applies z-score transformation to all variables. While this choice departed from models estimated using natural units, it enables a coherent interpretation of effect sizes across variables and models. It was also justified by the fact that most explanatory variables are proportions or indices ranging from 0 to 1, which would otherwise be difficult to compare directly.

Statistics

To quantify spatial autocorrelation between the IRIS units under study (Le Gallo, 2002), we used Moran's *I* statistic (Oliveau, 2010), which ranges from -1 to +1 and measures both the direction and strength of the spatial autocorrelation. A positive value indicates a positive spatial correlation, while a negative value indicates a negative correlation (Cliff & Ord, 1972). The higher the index, the stronger the spatial correlation (Kim & Castro, 2020). Conversely, an index close to 0 indicates the absence of autocorrelation, suggesting a random distribution of observations. The Moran's *I* formula is calculated as follows:

$$\text{Moran's } I = \frac{\sum_i \sum_j w_{ij}(z_i-\bar{z})(z_j-\bar{z})}{\sum_i (z_i-\bar{z})^2} \quad \text{Eq. 1}$$

where z_i and z_j represent the values of the variable in spatial units; *i* individual units; *j* neighbours; $\bar{z}$ the mean of the variable; and w_{ij} the weighted (neighbourhood) matrix.

Spatial weighted matrices are essential for capturing the structure of spatial relationships between regions through their relative positions. There are three types of weighted matrix (Oliveau, 2010) with the dimensions (*n**n), where *n* here is the number of IRIS units studied: i) the distance matrix (two IRIS are considered neighbours if they are situated within the defined neighbourhood distance. The default distance is the shortest distance where each IRIS has at least one additional neighbour); ii) the contiguity matrix (two regions *i* and *j* are contiguous of the *k* order if *k* is the minimum number of boundaries needed to cross to get from *i* to *j*.); and iii) the K-nearest neighbour matrix (all IRIS units have the same fixed number of neighbours).

Spatial autocorrelation can have several sources. It may come from spatially autocorrelated omitted variables or from measurement errors: the effect not captured by the explanatory variables can appear in the errors in the form of spatial autocorrelation (Le Gallo, 2002). This will then be considered as a tool for diagnosing spatial dependence, *i.e.* residual spatial autocorrelation serves as an essential diagnostic for the correct specification of models. It can therefore be used to verify the existence of spatial dependence between residuals. Significant residual autocorrelation can indicate that important explanatory variables have been omitted or misspecified (Cliff & Ord, 1972). Therefore, by calculating the global Moran's *I* for the regression residuals, we can detect these errors and adjust the model accordingly, ensuring a more accurate representation of the underlying spatial dynamics. Not using residuals leads to three major statistical problems in modelling: underestimation of standard errors, bias in parameter estimates and model specification errors (Gaspard *et al.*, 2019).

To this end, the global Moran's *I* is calculated in relation to the residuals of the regression estimated by the OLS model with the weight matrix and takes the following matrix form:

$$I = \frac{N}{S_0}\left(\frac{\varepsilon' W \varepsilon}{\varepsilon' \varepsilon}\right) \quad \text{Eq. 2}$$

where $\varepsilon = y - X\hat{\beta}$ is (the vector of residuals from the OLS regression); W = the spatial weights matrix; $S_0 = S_i S_j w_{ij}$ a standardization factor equal to the sum of all elements of W and N = the number of spatial units (IRIS in our case).

Although global indicesof spatial autocorrelation provide an overview of the spatial structure of the COVID-19 distribution, they lack precision when it comes to highly localized phenomena (Guillain & Le Gallo, 2008). The Local Indicator of Spatial Association (LISA), developed by Anselin (1995), was applied to assess the local level of spatial autocorrelation investigating the situation regarding IRIS units in proximity to each other. This was done to detect the potential emergence of possible and potential clusters based on COVID-19 incidence rates in the 144 IRIS units of Nice. The formula for Moran's local index is calculated as follows:

$$I_i = x_i \sum_j w_{ij} x_j \quad \text{Eq. 3}$$

where x_i and x_j represent COVID-19 incidence rates in IRIS units i and j, respectively; and w_{ij} is the spatial weighting matrix.

LISA divides neighbouring IRIS units into four cluster categories: High-High (HH) = areas with high incidence rates surrounded by neighbours with high numbers of cases; High-Low (HL) = areas with high incidence rates surrounded by neighbours with low numbers of cases; Low-Low (LL) = areas with low incidence rates surrounded by low numbers of cases; and Low-High (LH) = areas with a low incidence rate surrounded by high numbers of cases. To account for the multiple comparison issue inherent to local spatial statistics, we used pseudo p-values generated through 999 Monte Carlo permutations, as recommended by Anselin (1995) to ensure statistical reliability.

We used three spatial regression models to explore the relationship between COVID-19 incidence rate and socio-economic indicators: OLS, GWR and MGWR, where the OLS method assumes constant relationship across space. In our case, it was based on two key assumptions: firstly that observations were independent and constant within the study area, and secondly that there was no correlation between the error terms (Anselin & Rey, 1991; Nazia *et al.*, 2022). The exploratory analyses revealed polarization and spatial heterogeneity in the distribution of COVID-19, which justified the use of statistical tools sensitive to intra-urban variations, so GWR and MGWR (Han *et al.*, 2021; Fotheringham *et al.*, 2002) were applied. Lagrange Multiplier (LM) and robustness tests applied to the data did not reveal any significance problems. Unlike global models, the regression coefficients in the GWR model are not fixed but vary according to the geographical coordinates of observations. As a result, local regression parameter estimates were obtained at each observation point (Lu *et al.*, 2017). Thus, the coefficients of the explanatory parameters formed continuous surfaces estimated at certain points in space (Fotheringham *et al.*, 2003):

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^{p} \beta_k(u_i, v_i) x_{ik} + \epsilon_i \quad \text{Eq. 4}$$

where y_i is the COVID-19 incidence rate at location i; u_i, v_i the spatial coordinates; x_{ik} the explanatory variable; and e_i the residuals.

The GWR model uses a single optimal bandwidth for all explanatory variables, which assumes that all factors affect the COVID-19 rate at the same spatial scale. In this study, a Gaussian kernel was used to weight observations according to their spatial proximity (Brunsdont *et al.*, 1998). This choice reduced the influence of distant observations as distance increases progressively giving more weight to nearby observations (Fotheringham *et al.*, 2003). Model estimation is based on a matrix of weights W_i, whose values decrease as a function of the distance separating units i and j:

$$w(d_{ij}) = \exp\left(-\frac{1}{2}\left(\frac{d_{ij}}{b}\right)^2\right) \text{ if } d_{ij} < b, \text{ 0 otherwise.} \quad \text{Eq. 5}$$

The bandwidth b was estimated using a cross-validation approach, with the aim of minimizing the mean-square error (MSE):

$$MSE = \sum_{i=1}^{n}\left(y_i - \widehat{y_{i\neq i}}(b)\right)^2 \quad \text{Eq. 6}$$

The Gaussian kernel function used to weight the observations is essential for determining the spatial extent of the influence of neighbouring observations. The coefficients (β) can be obtained by minimizing the sum of the weighted squares:

$$\sum_{j=1}^{n} w_j(i)\left(y_j - \beta_0(u_i, v_i) - \sum_{k=1}^{p} \beta_k(u_i, v_i) x_{jk}\right)^2 \quad \text{Eq. 7}$$

with the weighted least squares estimator given by Fotheringham *et al.* (2003):

$$\hat{\beta}(u_i, v_i) = (X^T W(u_i, v_i) X)^{-1} X^T W(u_i, v_i) Y \quad \text{Eq. 8}$$

The choice of a single bandwidth can be penalizing, particularly when the explanatory variables influence the dependent variable at different spatial scales. This limitation can reduce the reliability of statistical inferences and bias the results (Wheeler, 2009; Lu *et al.*, 2017).To counter these problems, Fotheringham *et al.* (2017) extended GWR by developing the MGWR (Oshan *et al.*, 2019; Yu *et al.*, 2020), which allows the use of different optimal bandwidths specific to each explanatory variable since it eliminates the assumption of variations within the same scale (Yu *et al.*, 2020). By allowing each factor to influence COVID-19 levels at a different spatial scale, resulting in an optimal number of neighbours considered for each parameter estimate, the spatial relationships can be modelled more accurately, which favours the prediction of explanatory variables (Shabrina *et al.*, 2021):

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^{p} \beta_k(u_i, v_i, b_k) x_{ik} + \epsilon_i \quad \text{Eq. 9}$$

where thenon-fixed bandwidth b is specific to the variable k, *i.e.* b_k.

Model performance was assessed using the corrected Akaike Information Criterion (AICc) and adjusted R^2 values. These metrics confirmed that the MGWR model provides a better fit than

both global OLS and standard GWR models. The improvement in model fit underscores the relevance of accounting for spatial scale variability in the relationship between socio-economic factors and COVID-19 incidence.

The global and local spatial autocorrelation indices were computed using GeoDa software. The local models were estimated using MGWR software, version 2.2 (Spatial Analysis Research Center (SPARC), Tempe, AZ, USA), developed by Fotheringham *et al.* (2017). Output and mapping were produced using R studio.

Results

The mapping of the different socio-economic variables in Nice reflects social, economic and service access inequalities. Each map reveals a pattern of geographical polarization, with sharp contrasts in their spatial distribution (Figure 1). Polarization was found to be particularly strong for the LHDI variable, which showed a marked disparity between the city's advantaged neighbourhoods (Centre/Southeast) and the more disadvantaged ones in the Northwest. Similarly, the spatial pattern of low-income housing (HLM) revealed a polarization between the western part where social housing is focused and the Centre-East where it is virtually

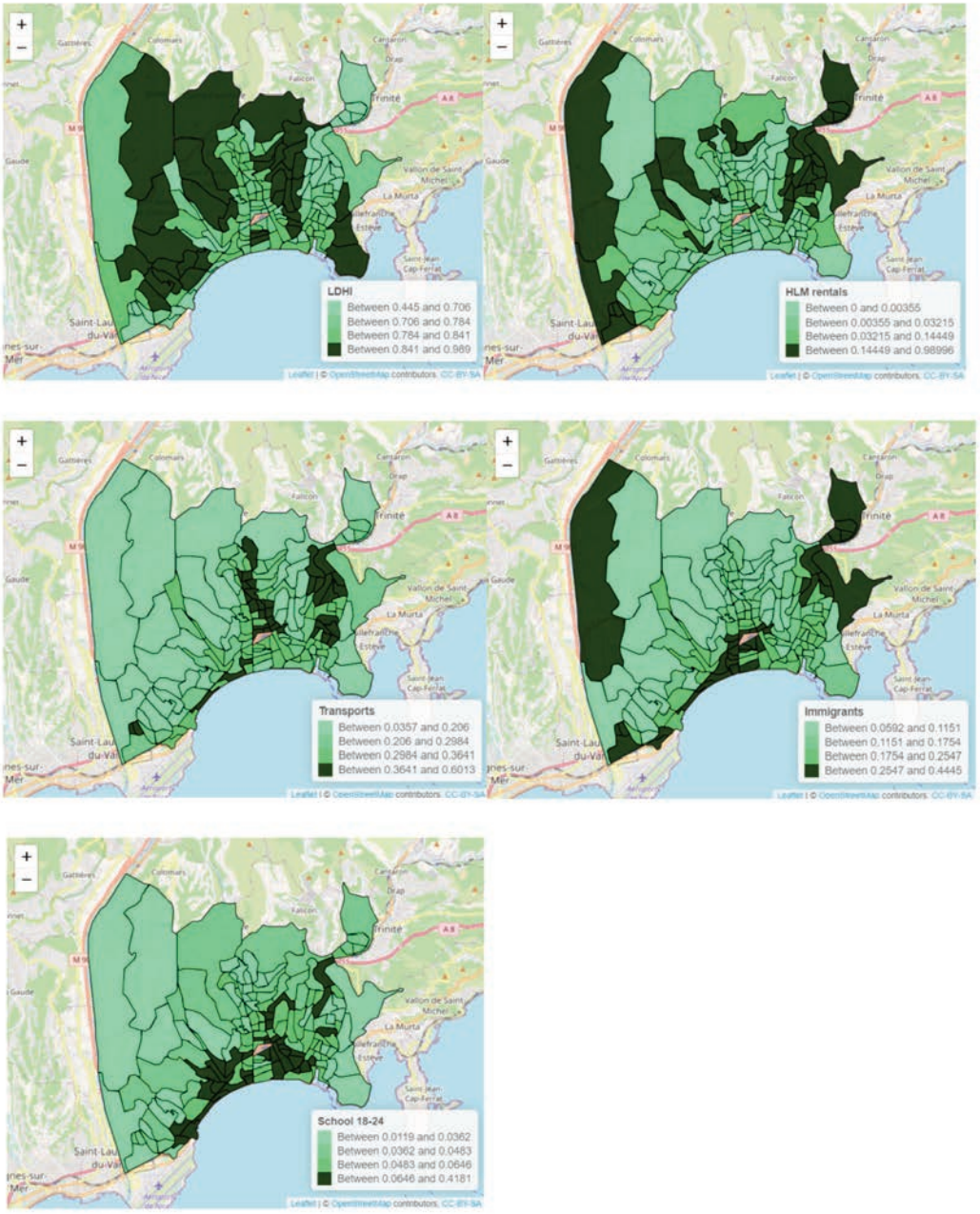

Figure 1. Cartographic representation of the independent variables.

absent. This reflects a concentration of social housing in specific areas, which could indicate residential segregation - some areas are clearly more favoured with respect to social housing than others.

Figure 2 maps the spatial distribution of the COVID-19 incidence rate across the IRIS units. It shows that no area in Nice was entirely spared, although the pandemic appears concentrated in working-class neighbourhoods in the far western and north-eastern parts of the city. The study period reported here, which heralds the start of the third wave of the pandemic, shows the extent of the epidemiological situation. The OLS method, used to identify significant predictors of the prevalence rate of COVID-19 incidence in Nice, explained 41.0% of the variation highlighting that the proportion of the population being migrants and those living in Low-Income Housing (HLM) were positively correlated with the dependent variable. The LHDI, on the other hand, revealed a negative correlation between the COVID-19 incidence rate and the proportion of the working population using public transport and those aged 18 to 24 years attending school (Table 1).

Spatial dependence was measured by defining a neighbourhood structure for the 144 IRIS units in Nice. To ensure the robustness of spatial autocorrelation results, we tested four alternative spatial weighting matrices: first-order contiguity, second-order contiguity, inverse-distance and exponential distance matrices. The results are presented in Table 2. Among these, the first-order contiguity matrix was retained as the main specification, as it produced the strongest and most significant spatial clustering of COVID-19 incidence, with a similar pattern of spatial dependence observed across all specifications. Although the 2nd-order contiguity matrix did not reach statistical significance, both inverse-distance and exponential distance matrices were strongly significant, thus confirming that the observed clustering patterns were not an artefact of the spatial structure but indeed a reflection of robust spatial processes (Table 2). These results indicate that the distribution of the COVID-19 incidence rate had a significant positive correlation with the incidence rates of neighbouring IRIS units during the study period in Nice. Spatial autocorrelation of residuals allowed us to check whether this dependency was captured by the model or whether 'unmodelled' spatial structures still existed (Table 3). To this end, Moran's *I* was calculated with the weight matrix used for the dependent variable on the residuals of the chosen classical linear regression. As seen in the table, Moran's *I* on the residuals was not statistically significant, which made us conclude that there was no significant spatial autocorrelation in the residuals of our model.

Still, diagnostic tests, such as LM tests, are required to assess the robustness of the model. The regression output allowed us to diagnose the spatial dependence of the residuals using a number of tests (supplementary data). Firstly, Moran's error test (-0, 0508), with a very high *p*-value for the residuals of Moran's *I* (p = 0.13798), made us believe that a global spatial relationship could be ruled out. The LM test for the spatial lag model also had a high *p*-value (p= 0.10633), suggesting that application of a Spatial Autoregressive (SAR) would not be appropriate. This finding is reflected in the decision rule recommended by Anselin and Florax (1995), which is based on the significance of the tests. Furthermore, the LM (error) (p-value=0.07964) and the robust LM (error) (p=0.38707) tests indicated that we also could not use the Standard Error (SEM) model. These findings, consistent with the LM diagnostics, confirmed the absence of global spatial dependence in the data. To complement this, we also estimated the outcome of the SAR and the SEM using the first-order contiguity matrix. The results, reported in Table 4, showed that the spatial autoregressive coefficient (ρ) in the SAR model was not statistically significant (ρ = 0.0328, p = 0.783) and the spatial error coefficient (λ) in the SEM model was also non-significant (λ = -0.1353, p = 0.349). None of the models could improve the OLS model fit. These results confirmed that there was no significant global spatial dependence, which justified the use of local regression approaches such as GWR and MGWR to capture spatial heterogeneity in the relationships between COVID-19 incidence and socio-economic

Table 1. Odd Least Squares (OLS) regression.

Variable	Model	T-value	p
Intercept	0.000	0.000	1.000
LHDI	-0.319	-1.958	0.050
School (18-24 years olds)	-0.123	-1.824	0.068
Immigrant	0.218	2.061	0.039*
Commuter	-0.325	-3.635	0.001***
HLM	0.257	2.380	0.017*

LHDI, local human development index; HLM, social housing; *significance at the ≤0.05 level; ***significance at the ≤0.001 level.

Table 2. Global autocorrelation of the dependant variable.

COVID incidence indicator (rate & matrix)	Moran's *I*	Z-value	p
First-order contiguity	0,315	6.981	0.001
2nd-order contiguity	0,031	1.105	0.134
Exponential distance	0,025	6.674	0.001
Inverse distance weighted	0,137	5.832	0.001

Table 3. Global autocorrelation of the residuals.

Regression indicator (residual & matrix)	Moran's *I*	Z-value	p
First-order contiguity matrix	-0.030	-0.4056	0.356
2nd-order contiguity matrix	-0.018	-0.3341	0.367
Inverse-distance weighted matrix	-0.051	-1.4627	0.053

Table 4. Comparison of global spatial regression models.

Model	Spatial parameter	Estimate	p	AIC	R^2	LLH
SAR	ρ	0.0328	0,783	1,885.06	0.39	-935.53
SEM	λ	-0.1353	0,349	1,882.49	0.40	-935.24

LLH, log- likelihood; SAR, spatial autoregressive regression; SEM, spatial error model; ρ, spatial autoregressive coefficient; λ, spatial error coefficient.

factors. We could therefore conclude that there was local spatial dependence in our data.

Among the four spatial weighting matrices tested (first-order contiguity, second-order contiguity, inverse-distance and exponential distance), the exponential distance spatial weighting matrix was selected as the reference specification for the LISA analysis of the COVID-19 incidence rates in Nice as it provided the most spatially coherent and statistically robust clustering patterns (Figure 3). To address the issue of multiple comparisons, pseudo *p*-values were computed using 999 Monte Carlo permutations following the procedure recommended by Anselin (1995). This ensured the statistical validity of the local spatial association results as illustrated in the accompanying significance map (Figure 3) that revealed significant spatial clustering. We identified two HH clusters in the extreme western and north-eastern parts of Nice (together including 8 IRIS units). A large LL cluster was also visible in central Nice (35 IRIS units) suggesting spatial concentration of low incidence rates. An addition 25 IRIS units fell into the HL or LH categories, while 78 IRIS units were not statistically significant. These results confirmed the presence of positive local spatial dependence and a polarized spatial distribution of COVID-19 incidence across the city during the time the study was carried out.

We then compared the performance of global and local models. Table 5 summarizes key metrics (bandwidths, effective number of parameters, critical T-values, R^2 and AICc) for our local models. The multi-scale MGWR model provided the best performance,

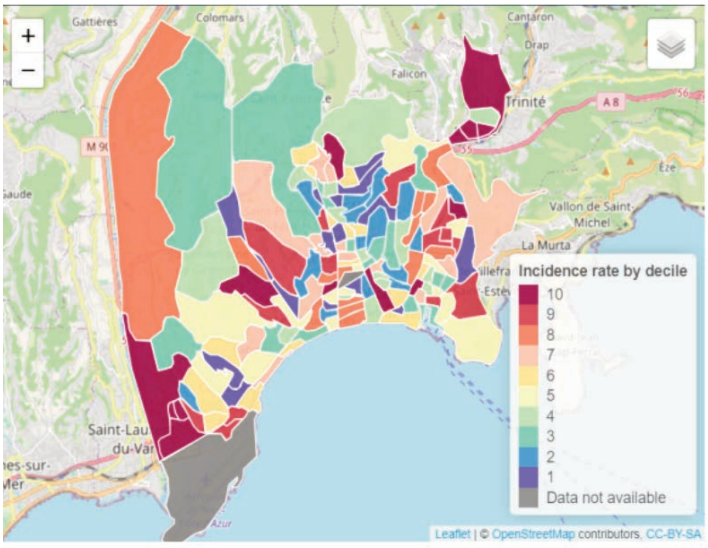

Figure 2. Spatial distribution of the cumulative incidence rate of COVID-19 by decile. Timeframe: 4 January to 14 February 2021.

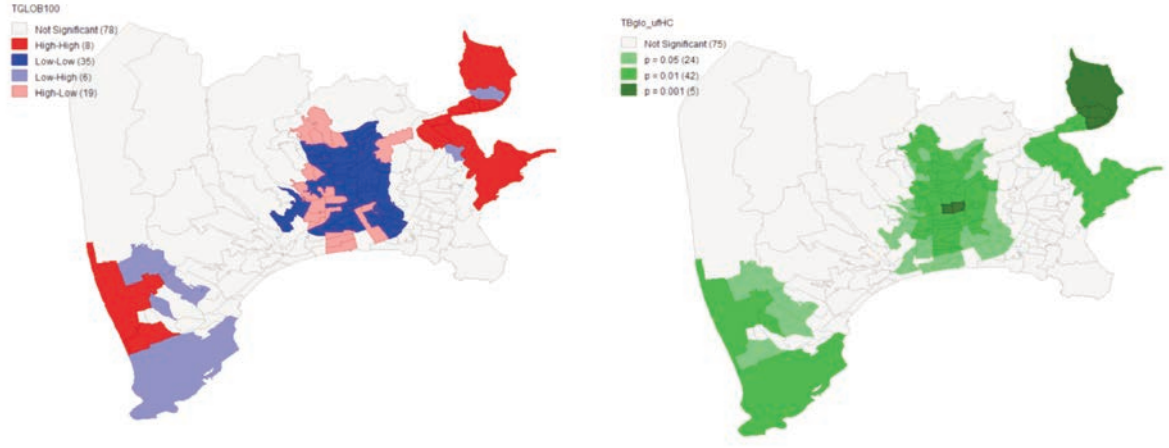

Figure 3. Local Indicators of Spatial Association (LISA) applied to the COVID-19 incidence rate in Nice using a distance-based spatial weighting matrix. Left: LISA cluster typology; Right: Significance levels of local spatial associations (p-values from 999 Monte Carlo permutations).

explaining 52% of the variance in COVID-19 incidence, outperforming both OLS (adjusted R^2=0.389) and standard GWR (adjusted R^2=0.429).

The MGWR model revealed spatial variation in the effects of each variable (Table 6). Figures 4 to 8 display the spatial distribution of coefficients for the five socio-economic predictors, highlighting how their effects vary across the city. The maps in blue shade show the significant estimates of local parameters based on the significance of the *p*-value at the 5% threshold.

The maps depicted in Figures 4-8 show that all the socio-eco-

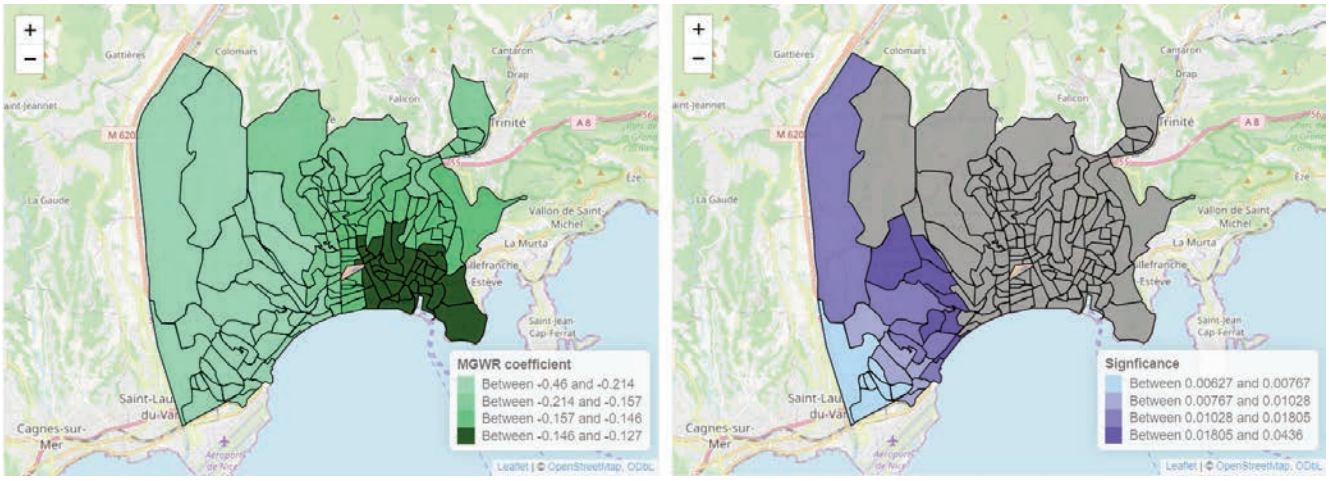

Figure 4. Coefficient and significance of the local human development index variable in the MGWR model.

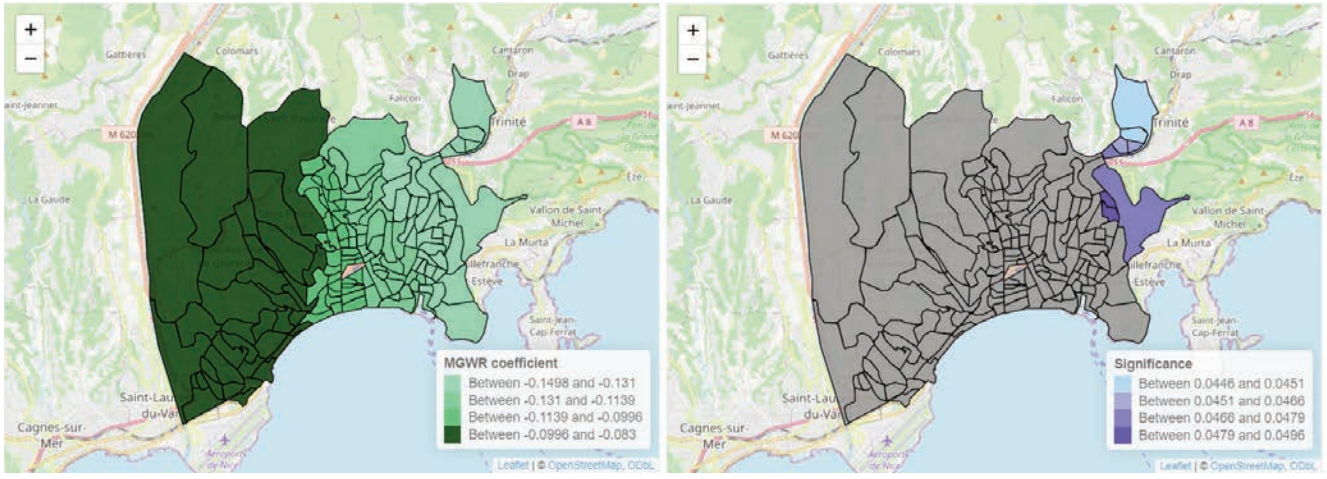

Figure 5. Coefficient and significance of the variable 'School-enrolled - aged between 18 and 24 years' in the MGWR model.

Table 5. Comparaison between the geographically weighted models.

Diagnostic	GWR Model	MGWR Model	Intercept	LDHI	School	Migrant	Commuter	HLM
Bandwidth	140		143	135	143	143	60	137
Degree of dependency	0.889	0.827	0.939	0.887	0.938	0.920	0.630	0.857
Critical T (95 %)	2.209		2.105	2.212	2.109	2.145	2.687	2.271
AICc	342.074	337.163						
R^2	0.471	0.520						
N=144								

GWR, geographically weighted regression; MGWR, multiscale geographically weighted regression; LHDI, localized human development index; HLM, social housing; AICc, corrected Akaike information criterion; N, number of IRIS units investiaged.

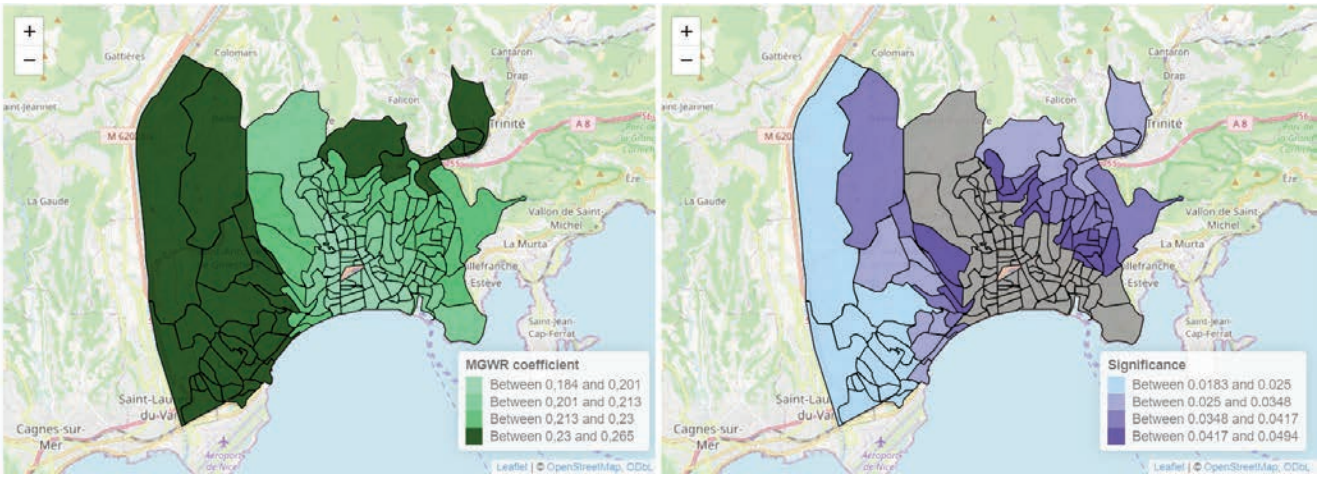

Figure 6. Coefficient and significance of the variable 'immigrant population' in the MGWR model.

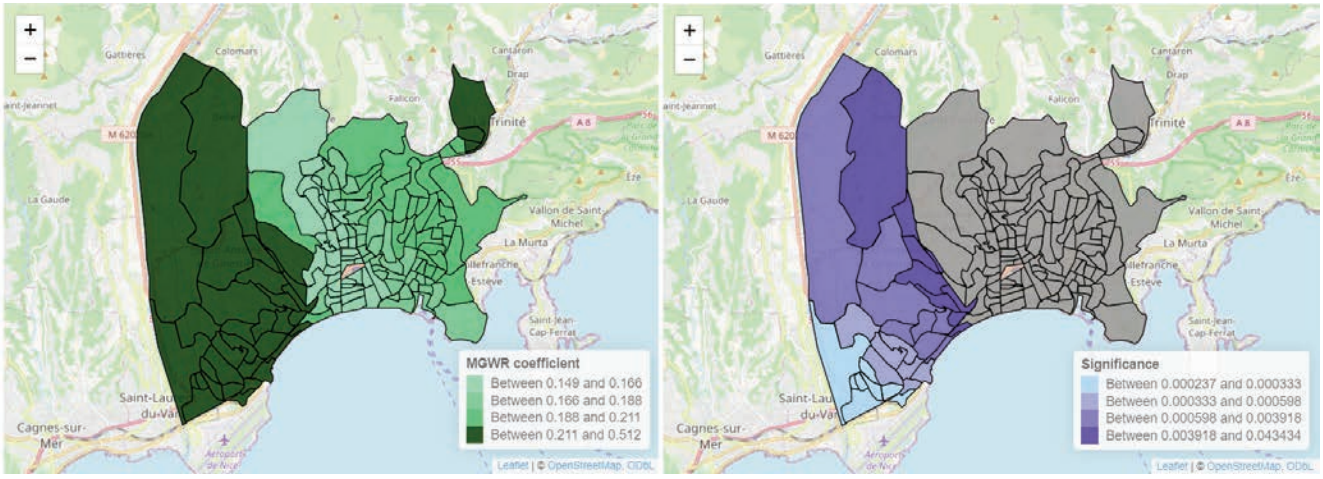

Figure 7. Coefficient and significance of the variable 'Population living in social housing (HLM)'.

Table 6. Comparison of global and local parameter estimates across regression models.

Variable	OLS			GWR				MGWR					
	Coefficient	Standard error	T-stat.	Standard error	Min	Median	Max	Mean	Standard error	Min	Median	Max	Mean
Intercept	0.000	0.065	0.000	0.041	-0.129	-0.013	0.050	-0.019 *	0.023	-0.142	-0.077	-0.041	-0.076 **
LDHI	- 0.319 .	0.163	-1.958	0.028	-0.360	-0.286	-0.268	-0.299*	0.105	-0.460	-0.157	-0.127	-0.210 **
School (18-24-year olds)	-0.123 .	0.067	-1.824	0.022	-0.170	-0.120	-0.086	-0.123*	0.017	-0.150	-0.114	-0.083	-0.116 **
Immigrant	0.218 *	0.106	2.061	0.041	0.105	0.188	0.265	0.181*	0.021	0.184	0.213	0.265	0.217 **
Cummuter	-0.325 ***	0.089	-3.635	0.032	-0.387	-0.317	-0.252	-0.316***	0.165	-0.713	-0.222	-0.034	-0.280 ***
HLM	0.257 *	0.108	2.380	0.134	0.026	0.152	0.513	0.201**	0.111	0.149	0.188	0.512	0.234 **

Adjusted R^2 = 0.389 (OLS); 0.429 (GWR); 0.476 (MGWR). AIC: 501.999 (OLS); 339.918 (GWR); 333.323 (MGWR). GWR, geographically weighted regression; MGWR, multiscale geographically weighted regression; LHDI, localized human development index; HLM, social housing; AIC, Akaike Information Criterion; **significance at the ≤0.01 level; ***significance at the ≤0.001 level.

nomic factors had a significant impact. The LHDI variable, the 18-24-year-olds enrolled in school, and employed people using public transport all displayed a negative relationship with the COVID-19 incidence rate. According to the MGWR model, one SD increase in the LHDI rate was associated with a corresponding decrease at the 0.32 level in COVID-19 cases per 100,000 inhabitants, reflecting a moderate protective effect. As seen in Figure 4, the strongest local reductions was seen in the western parts of Nice (–0.460 to –0.214), while weaker effects were observed in the Southeast (–0.146 to –0.127). One SD increase in the proportion of 18–24-year-olds enrolled in schools was associated with SD decrease of 0.12 in incidence, indicating a limited protective effect. As shown in Figure 5, the spatially strongest, negative associations appeared in the western IRIS units (–0.099 to –0.083) and the weakest in the East (–0.149 to –0.131). For those employed using public transport, the average MGWR estimate indicated that one SD increase was associated with a corresponding decrease of 0.28 in incidence. Overall, that the highest coefficients were observed in the city centre, ranging from -0.141 to -0.034 (Figure 8).

As for the people in low-income housing (HLM) and immigrants, these variables positively influence the COVID-19 incidence rate. According to the MGWR model, one SD increase in the proportion of people living HLM was associated with an increase of 0,25 in COVID-19 incidence. Locally, the spatial distribution showed the highest estimates in the West, as well as in a few IRIS units in the extreme eastern part of the city, with coefficient variations ranging from 0.211 to 0.512. As seen in Figure 7, notably affected neighbourhoods included Ariane in the Northwest and Pasteur in the Centre. Similarly, increases by one SD in the migrants led to an increase of 0.22 in COVID-19 incidence, which corresponded to a moderately positive effect (range: 0.23 to 0.265) in the extreme north-eastern and western parts of the city, while the lower effects had lower increases ranging from 0.184 to 0.201 (Figure 6).

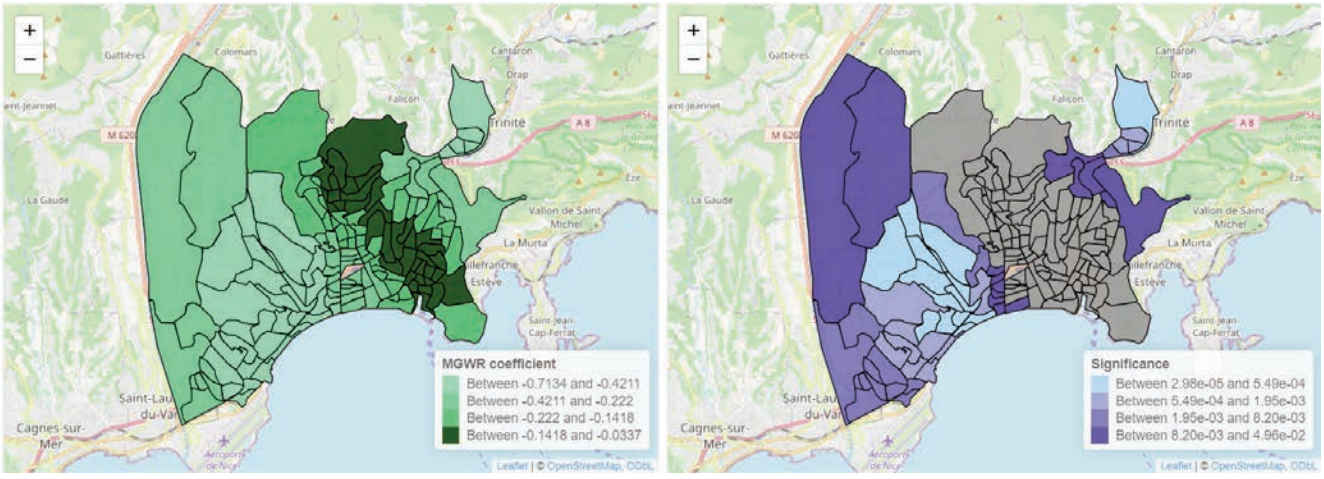

Figure 8. Coefficient and significance of the variable 'Commuters - individuals using public transportation' in the MGWR model.

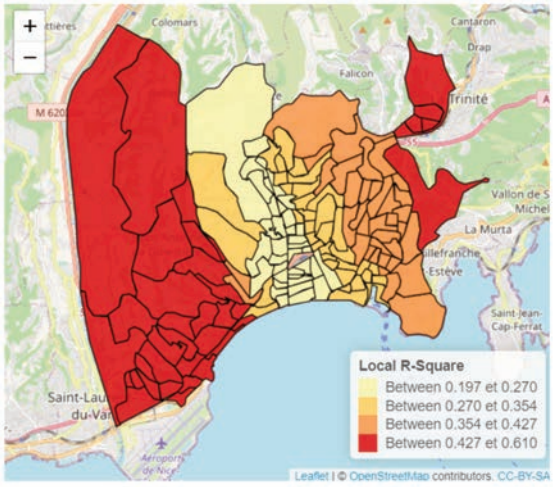

Figure 9. Local R^2 variations over the study area.

The R^2 value indicates that the MGWR model explains 52% of the variations in the COVID-19 incidence rate in Nice, which is higher than that shown by the OLS model. To illustrate this, Figure 9 shows the spatial variations in local R^2 values in the study area of the COVID-19 incidence rate associated with the socio-economic factors of COVID-19 disease distribution for each IRIS unit according to the MGWR model. The values with the highest R^2 ($0.427 \leq R^2 \leq 0.61$) were seen in the extreme eastern and western parts of the city. This indicates a strong prediction of the concentration of infected cases in these areas. The values with the lowest R^2 were found in the Centre, which indicates good performance of the MGWR model in this study area, since we found the same results as the clusters previously given by LISA.

Discussion

Since the beginning of the twentieth century and throughout the history of pandemics, scientists have regarded human contact as a critical vector in the spatial spread of disease-causing viruses. Variations in this spread are often associated with socio-economic factors. In other words, different socio-economic groups may be vulnerable in different ways, depending on lifestyle and social status. To explore this point, we highlighted five socio-economic factors to explain and visualize the spatial distribution of the incidence rate of COVID-19. A central question was if it is completely random or if the incidence rate in one particular place (IRIS in our case) influenced that of neighbouring IRIS units. To answer this question, we compared GWR and its extension MGWR, which have been widely applied in epidemiology and socio-economic research. The former has proven effective in exploring spatial heterogeneity (Apparicio *et al.*, 2007; Wheeler, 2009; Dziauddin & Idris, 2017; Han *et al.*, 2021; Lotfata, 2022), while the latter allows each explanatory variable to operate at its own spatial scale (Maiti *et al.*, 2021; Shabrina *et al.*, 2021; Lotfata, 2022; Ma *et al.*, 2022; Nazia *et al.,* 2022). For example, Nazia et al. have shown that there is heterogeneity between the distribution of COVID-19-infected cases and risk factors by illustrating the spatial variation between the incidence rate and socio-economic factors using GWR.

Our results show that the MGWR model outperforms global OLS and standard GWR in terms of model fit, highlighting the importance of accounting for spatial scale heterogeneity for socio-economic predictors of COVID-19 incidence. Local Moran's *I*, which measures spatial dependency, revealed the existence of three clusters in our study area: hotspots in the West and Northeast and a coldspot in the Centre. The analysis revealed a significant negative association between the LHDI and the incidence rate of COVID-19, an inverse relationship that can be explained by the characteristics of this indicator that assesses the standard of living in each area; not only based on economic data, but also on the well-being of its inhabitants. This indicator captures the intrinsic vulnerability of populations, particularly those most affected by the virus (Scarpone *et al.*, 2020). Indeed, as shown in Figure 1, these areas are known for their concentration of social housing and our model supports this finding by revealing the very strong positive association between areas characterized by social housing and migrants on the one hand and the highest incidence rates of COVID-19 on the other. With the majority of migrants in low-income housing, this positive association can be explained by the more precarious living conditions (overcrowding) and working conditions (menial jobs), which makes this population particularly exposed to the risk of contamination. Difficulties in accessing healthcare and information are also contributing factors, notably due to language barriers (Vang & Ng, 2023) or socio-economic status (Kirksey *et al.*, 2021). Despite the various confinements and restrictions put in place, people of migrant background face an increased risk of exposure to the virus, reflecting social inequalities in health and ethno-racial discrimination. As mentioned above, this "virus of inequalities" affected several populations, particularly those most at risk from respiratory or chronic diseases. Migrant and 'racialized' communities were among the workers mobilized to survive the crisis when out of work due to the closure of businesses, such as restaurants, hotels and domestic work designated as non-essential during the pandemic. Professional vulnerability thus increased during the pandemic, particularly among immigrants, which not only increased the risk of contracting COVID-19, but also led to loss of income, deterioration in material conditions, financial stress and food insecurity. Thus, the high concentration of COVID-19 among migrants in Nice was largely due to socio-economic conditions, not necessarily physiological risks.

As mentioned earlier, people using public transport to get to work were less likely to have COVID-19 than those using their own vehicles. To address this counterintuitive finding, we have some tentative explanations. Firstly, the public transport variable we used comes from census data and reflects usual commuting modes rather than actual mobility behaviour during the pandemic. It specifically captured the proportion of workers using public transport for home to work and return in normal times, not during COVID-19 restrictions. Secondly, we noted that in Nice, more than three-quarters of these commuting flows are intra-communal. As such, IRIS units with high public transport usage may reflect central areas where better access to healthcare, stricter enforcement of sanitary measures or higher compliance with wearing facemasks could have contributed to lower incidence rates. Given these uncertainties, it is clear that the interpretation of this variable is complex. Indeed, it may be capturing features of central IRIS units rather than actual exposure risk, and should therefore be treated with caution.

Employment category is a concept that can explain the LHDI variable as well as the transport variable. A person's occupational category may expose him or her to a major risk of contamination, and work may, in most cases, involve interaction with others caused by frequent contact or the mode of transport used to get to the workplace (Rule *et al.*, 2018). Work that cannot be conducted from home but necessitates human contact increases the risk of infection. Occupation is therefore a direct determinant of infection, as well as being an indirect determinant of the extent at which the disease spreads. Thus, work is correlated with the level of education, a variable used to calculate the LHDI indicator. We can therefore conclude that a low level of education may be an indirect factor in the development of severe forms of COVID-19. A low level of education can also lead to low income, which can affect living conditions, such as housing in deprived areas, which can increase the risk of COVID-19 and other pathogen infections. However, it is important to note that the development of severe forms of COVID-19 is more likely related to pre-existing medical conditions and social determinants of health rather than increased exposure.

Strengths and limitations

The methodology used has several strengths; firstly, it applied a variety of spatial econometric techniques. Secondly, it focused

on an urban setting, where testing accessibility was not a barrier and no major policy interventions (lockdowns or vaccination schemes) were in effect during the study period. This temporal window was particularly informative as it could capture the onset of the third wave and the emergence of new variants, but preceded mass vaccination and major policy interventions. Its short duration also limited potential seasonal biases, offering a temporally homogeneous framework. Finally, the use of IRIS units added a fine-scale approach, enabling the detection of intra-urban disparities that broader spatial units tend to obscure. Together, these elements provided a robust empirical and methodological framework to assess local drivers of COVID-19 spread. This study thus contributes by combining disaggregated epidemiological data, socio-economic indicators, and multiscale spatial models to reveal inequalities with regard to disease incidence. Even if the reliance on aggregated data may limit generalizability, the robustness of the methodology and consistency of the results with broader literature support its validity.

The work is, however, not without limitations; firstly, the analysis focused on a short six-week period immediately following the end-of-year holidays. During this time, no restrictive policies were in place, allowing relatively unfiltered social interactions. The brevity of the period may have restricted the generalizability of the results by reducing the number of positive cases per IRIS, potentially affecting the robustness of incidence estimates; secondly, COVID-19 incidence was measured using voluntarily reported test results, which may introduce a detection bias. Although Nice had wide testing availability, disparities in the test-seeking behaviour could have influenced case counts, especially across different socio-economic groups; thirdly, this analysis relied on ecological data aggregated at the IRIS level. While IRIS units are relatively homogeneous compared to larger administrative divisions, this scale still presents a risk of the ecological fallacy, meaning that inferences made at the group level may not reflect individual-level relationships. Nonetheless, the IRIS scale offered access to a rich set of contextual variables (343 initially), far more than would be available at the individual level; fourthly, the interpretation of the "public transport" variable presents a conceptual limitation, specifically, the main mode of travel between home and work. Importantly, as most workers live and work within the city over three-quarters of commuting flows are intra-municipal. Therefore, this variable measured the share of workers who regularly use public transportation in each IRIS under normal conditions. The surprising negative correlation with COVID-19 incidence could either reflect a higher compliance with preventive measures or be due to structural characteristics of central IRIS, where public transit is more common but where the incidence remained lower. It is difficult to disentangle whether we captured behavioural effects or underlying neighbourhood traits. This ambiguous interpretation of the transport variable must therefore be acknowledged as a limitation. In addition, our analysis was based on the cumulative incidence rate. Daily or weekly data might have been more appropriate for a finer spatial analysis. Future research should extend our approach to other urban areas or explore the temporal dynamics of incidence using longitudinal data.

Finally, contrary to what might have been expected, the proportion of people aged between 18- and 24-years and attending schools was not found to be a strong factor in transmission. One possible explanation for this is that our study period did not cover the entire duration of the pandemic; these IRIS units might have been massively contaminated during the first waves of the epidemic leading to the development of 'heard immunity' in the populations there by the time of our analysis.

Conclusions

This study provides new insights into the spatial determinants of COVID-19 incidence during a six-week window, a critical period in early 2021marked by high transmission rates in Nice. By relying on fine-scale IRIS units and a rich set of geolocated, socio-economic indicators, our analysis identified key contextual factors shaping the spatial heterogeneity of the pandemic. Methodologically, the multi-scale approach revealed significant spatial variations in the strength and direction of associations between COVID-19 incidence and local determinants. The most disadvantaged neighbourhoods, characterized by limited development, high concentrations of low-income housing and immigrant population experienced the highest infection rates. Conversely, the spatial relationship between public transport use and lower incidence rates Unlike traditional global models, MGWR enables geographically targeted predictions and interventions tailored to intra-urban disparities. Overall, these findings demonstrate the value of spatially disaggregated analysis for informing local health strategies and confirm that covid-19 is strongly spatially correlated, and that spatial analysis is an essential step in implementing effective preventive measures. This article emphasizes that spatially informed public health strategies are essential in mitigating the unequal impacts of epidemics across neighbourhoods and in building more resilient, equitable urban health systems.

References

Anselin L, 1995. Local Indicators of Spatial Association—LISA. Geogr Anal 27:93–115.

Anselin L, Florax RJGM, 1995. Small Sample Properties of Tests for Spatial Dependence in Regression Models: Some Further Results. In New Directions in Spatial Econometrics. Anselin L, Florax, RJGM (Eds.), Springer, Berlin, Heidelberg, pp. 21–74.

Anselin L, Rey S, 1991. Properties of Tests for Spatial Dependence in Linear Regression Models. GeogrAnal 23:112–31.

Apparicio P, Séguin AM, Leloup X, 2007. Modélisation spatiale de la pauvreté à Montréal : apport méthodologique de la régression géographiquement pondérée. Can Geogr 51:412–27.

Bambra C, Riordan R, Ford J, Matthews F, 2020. The COVID-19 pandemic and health inequalities. J Epidemiol Community Health 74:964–8.

Benita F, Rebollar-Ruelas L, Gaytán-Alfaro ED, 2022. What have we learned about socioeconomic inequalities in the spread of COVID-19? A systematic review. Sustain Cities Soc 86:104158.

Brun S, Simon P, 2020. L'invisibilité des minorités dans les chiffres du Coronavirus : le détour par la Seine-Saint-Denis. De Facto – Institut Convergences Migrations, Dossier : Inégalités ethno-raciales et pandémie de coronavirus 68–78.

Brunsdon C, Fotheringham S, Charlton M, 1998. Geographically weighted regression – Modelling spatial non-stationarity. J Royal Statist Soc Series D 47:431–43.

Chossegros P, 2020. L'épidémie de COVID-19: une autre histoire pourrait être racontée. La Presse Médicale Formation 1:447–

50.
Cliff A, Ord K, 1972. Testing for spatial autocorrelation among regression residuals. Geographical Analysis 4:267–84.
Cordes J, Castro MC, 2020. Spatial analysis of COVID-19 clusters and contextual factors in New York City. Spatial and Spatio-temporal Epidemiology 34:100355.
Dowrick A, Rai T, Hinton L, Eacott B, Baker S, Askew M, Ziebland S, Locock L, 2022. Health inequalities, ethnic minorities and COVID-19: interactive theatre workshop drawing on a qualitative interview study. Lancet 400 Suppl1:S9.
Dziauddin MF, Idris Z, 2017. Use of Geographically Weighted Regression (GWR) Method to estimate the effects of location attributes on the residential property values. Indonesian J Geogr 49:97–110.
Etowa J, Hyman I, 2021. Unpacking the health and social consequences of COVID-19 through a race, migration and gender lens. Can J Public Health 112:8–11.
Etowa J, Hyman I, Dabone C, Mbagwu I, Ghose B, Sano Y, Osman M, Mohamoud H, 2021. Strengthening the collection and use of disaggregated data to understand and monitor the risk and burden of COVID-19 among racialized populations. Can Stud Popul 48:201–16.
Fontán-Vela M, Gullón P, Bilal U, Franco M, 2023. Social and ideological determinants of COVID-19 vaccination status in Spain. Public Health 219:139–45.
Fotheringham A, Brunsdon C, Charlton M, 2002. Geographically weighted regression: the analysis of spatially varying relationships. John Wiley& Sons.
Fotheringham AS, Brunsdon C, Charlton M, 2003. Geographically Weighted Regression: The Analysis of Spatially Varying Relationships. John Wiley & Sons.
Fotheringham AS, Yang W, Kang W, 2017. Multiscale Geographically Weighted Regression (MGWR). Annals of the American Association of Geographers 107:1247–65.
Frohlich N, Mustard C, 1996. A regional comparison of socioeconomic and health indices in a Canadian province. Soc Sci Med 42:1273–81.
Gaspard G, Kim D, Chun Y, 2019. Residual spatial autocorrelation in macroecological and biogeographical modeling: a review. Journal of Ecology and Environment 43:19.
Godefroy R, Lewis J, 2022. What explains the socioeconomic status-health gradient? Evidence from workplace COVID-19 infections. SSM Popul Health 18:101124.
Guillain R, Le Gallo J, 2008. Identifier la localisation des activités économiques : une approche par les outils de l'analyse exploratoire des données spatiales. Économie appliquée 61:5–34.
Han Y, Yang L, Jia K, Li J, Feng S, Chen W, Zhao W, Pereira P, 2021. Spatial distribution characteristics of the COVID-19 pandemic in Beijing and its relationship with environmental factors. Science of The Total Environment 761:144257.
Hu Y, Goldman N, 1990. Mortality differentials by marital status: an international comparison. Demography 27:233–50.
Kim S, Castro MC, 2020. Spatiotemporal pattern of COVID-19 and government response in South Korea (as of May 31, 2020). Int J Infect Dis 98:328–33.
Kirksey L, Tucker DL, Taylor E, White Solaru KT, Modlin CS, 2021. Pandemic superimposed on epidemic: COVID-19 disparities in black Americans. J Natl Med Assoc 113:39–42.
Le Gallo J, 2002. Économétrie spatiale : l'autocorrélation spatiale dans les modèles de régression linéaire. Économie & Prévision 155:139–57.
Lotfata A, 2022. Using geographically weighted models to explore obesity prevalence association with air temperature, socioeconomic factors, and unhealthy behavior in the USA. J Geovis Spat Anal 6:14.
Lu B, Brunsdon C, Charlton M, Harris P, 2017. Geographically weighted regression with parameter-specific distance metrics. Int J Geogr Inf Sci 31:982–98.
Luo Y, Yan J, McClure S, 2021. Distribution of the environmental and socioeconomic risk factors on COVID-19 death rate across continental USA: a spatial nonlinear analysis. Environ SciPollutRes Int 28:6587–99.
Ma J, Zhu H, Li P, Liu C, Li F, Luo Z, Zhang M, Li L, 2022. Spatial patterns of the spread of COVID-19 in Singapore and the influencing factors. ISPRS Int J Geo-Inf11:152.
Maiti A, Zhang Q, Sannigrahi S, Pramanik S, Chakraborti S, Pilla F, 2021. Spatiotemporal effects of the causal factors on COVID-19 incidences in the contiguous United States. Sustain Cities Soc 68:102784.
Ministère du Travail, de la Santé et des Solidarités, 2024. Professionnels de santé [WWW Document]. Ministère du Travail, de la Santé et des Solidarités.https://sante.gouv.fr/soins-et-maladies/maladies/ maladies-infectieuses/coronavirus/professionnels-de-sante/ Accessed 3.18.24.
Nazia N, Law J, Butt ZA, 2022. Spatiotemporal clusters and the socioeconomic determinants of COVID-19 in Toronto neighbourhoods, Canada. Spatial Spatio-temporal Epidemiol 43:100534.
Oliveau S, 2010. Autocorrélation spatiale: leçons du changement d'échelle. L'Espace géographique 39:51–64.
Oshan TM, Li Z, Kang W, Wolf LJ, Stewart Fotheringham A, 2019. MGWR: A python implementation of multiscale geographically weighted regression for investigating process spatial heterogeneity and scale. ISPRS Int J Geo-Inf 8:060269.
Patel JA, Nielsen FBH, Badiani AA, Assi S, Unadkat VA, Patel B, Ravindrane R, Wardle H, 2020. Poverty, inequality and COVID-19: the forgotten vulnerable. Public Health 183:110–11.
Patterson K, Chalifoux M, Gad R, Leblanc S, Paulsen P, Boudreau L, Mazerolle T, Pâquet M, 2022. Tendances démographiques de l'exposition à une éclosion de COVID-19 et de la transmission de cas dans une communauté rurale canadienne, 2020. RMTC 48:512–19.
Pinter-Wollman N, Jelić A, Wells NM, 2018. The impact of the built environment on health behaviours and disease transmission in social systems. Philos Trans R Soc Lond B Biol Sci 373:20170245.
Revollo-Fernández D, Rodríguez-Tapia L, Medina-Rivas C, Morales-Novelo JA, 2022. Socio-economic determinants of COVID-19 in Mexico. Public Health 207:28–30.
Rohleder S, Costa DD, Bozorgmehr PK, 2022. Area-level socioeconomic deprivation, non-national residency, and COVID-19 incidence: a longitudinal spatiotemporal analysis in Germany. eClinicalMedicine 49:101485.
Rule AM, Apau O, Ahrenholz SH, Brueck SE, Lindsley WG, de Perio MA, Noti JD, Shaffer RE, Rothman R, Grigorovitch A, Noorbakhsh B, Beezhold DH, Yorio PL, Perl TM, Fisher EM, 2018. Healthcare personnel exposure in an emergency department during influenza season. PLoS One 13:e0203223.
Ruthberg J, Quereshy H, Jella T, Kocharyan A, D'Anza B, Maronian N, Otteson T, 2020. Geospatial analysis of COVID-

19 and otolaryngologists above age 60. Am J Otolaryngol 41:102514.

Scarpone C, Brinkmann ST, Große T, Sonnenwald D, Fuchs M, Walker BB, 2020. A multimethod approach for county-scale geospatial analysis of emerging infectious diseases: a cross-sectional case study of COVID-19 incidence in Germany. Int J Health Geogr 19:32.

Schellekens P, Sourrouille D, 2020. COVID-19 Mortality in rich and poor countries: a tale of two pandemics? World Bank, Washington, DC. https://doi.org/10.1596/1813-9450-9260

Scott A, Coote W, 2007. Incentives and the Quality of Primary Care in Australia: Do Regional Primary Care Organisations Make a Difference? . iHEA 2007 6th World Congress: Explorations in Health Economics Paper, Available at SSRN: https://ssrn.com/abstract=995067

Shabrina Z, Buyuklieva B, Ng MKM, 2021. Short-term rental platform in the urban tourism context: a geographically weighted regression (GWR) and a multiscale GWR (MGWR) approaches. Geogr Anal 53:686–707.

Shahbazi F, Khazaei S, 2020. Socio-economic inequality in global incidence and mortality rates from coronavirus disease 2019: an ecological study. New Microbes New Infect 38:100762.

Siljander M, Uusitalo R, Pellikka P, Isosomppi S, Vapalahti O, 2022. Spatiotemporal clustering patterns and sociodemographic determinants of COVID-19 (SARS-CoV-2) infections in Helsinki, Finland. Spatial Spatio-temporal Epidemiol 41: 100493.

Talavera A, Perez EM, 2009. Is cholera disease associated with poverty? J Infect Dev Ctries 3:408–11.

Vang ZM, Ng E, 2023. The impacts of COVID-19 on immigrants and the healthy immigrant effect: Reflections from Canada. Prev Med 171:107501.

Vie publique. Covid-19: confinement du littoral des Alpes-Maritimes, vie-publique.fr, 2021. https://www.vie-publique.fr/en-bref/278682-covid-19-confinement-du-littoral-des-alpes-maritimes Accessed 5.18.25.

Vilinová K, Petrikovičová L, 2023. Spatial autocorrelation of COVID-19 in Slovakia. Trop Med Infect Dis 8:298.

Wheeler D, 2009. Simultaneous coefficient penalization and model selection in geographically weighted regression: the geographically weighted lasso. Environ Plan A 41:722–742.

Xu XW, Wu XX, Jiang XG, Xu KJ, Ying LJ, Ma CL, Li SB, Wang HY, Zhang S, Gao HN, Sheng JF, Cai HL, Qiu YQ, Li LJ, 2020. Clinical findings in a group of patients infected with the 2019 novel coronavirus (SARS-CoV-2) outside of Wuhan, China: retrospective case series. BMJ 368:m606.

Yu H, Fotheringham AS, Li Z, Oshan T, Wolf LJ, 2020. On the measurement of bias in geographically weighted regression models. Spat Stat 38:100453.

Online supplementary materials

First-order contiguity matrix.

Second-order contiguity matrix.

Inverse distance matrix.

Moran residuals with first-order contiguity matrix.

Moran residuals with second-order contiguity matrix.

Regression with Geoda, weight matrix: first-order contiguity.