




Spatial analysis of socio-economic factors and their relationship with the cases of COVID-19 in Pernambuco, Brazil

Carlos Fabricio Assunção da Silva¹  | Mayara Costa Silva¹  | Alex Mota dos Santos²  |
Anderson Paulo Rudke^{3,4}  | Cristine Vieira do Bonfim^{5,6}  | Gabriela Tobias Portis⁷  |
Pedro Monteiro de Almeida Junior⁸  | Maria Beatriz de Santana Coutinho⁸ 

¹Department of Cartographic and Survey Engineering, Federal University of Pernambuco, Recife, Brazil

²Center of Agroforestry Sciences and Technologies, Federal University of Southern Bahia, Itabuna, Brazil

³Department of Sanitary and Environmental Engineering, Federal University of Minas Gerais, Belo Horizonte, Brazil

⁴Federal University of Technology - Paraná, Londrina, Brazil

⁵Social Research Department, Joaquim Nabuco Foundation, Recife, Brazil

⁶Postgraduate Program in Collective Health, Federal University of Pernambuco, Recife, Brazil

⁷Federal Institute of Goiás, Goiânia, Brazil

⁸Department of Statistics, Center of Nature and Exact Sciences, Federal University of Pernambuco, Recife, Brazil

Correspondence

Carlos Fabricio Assunção da Silva,
Department in Cartographic and Surveying
Engineering, Center of Technologies
and Geosciences, Federal University of
Pernambuco - UFPE, Avenida Acadêmico
Hélio Ramos, s/n, Cidade Universitária, Recife
50740-530, Brazil.
Email: carlos.assuncao@ufpe.br

Abstract

Objectives: To analyse the spatial distribution of rates of COVID-19 cases and its association with socio-economic conditions in the state of Pernambuco, Brazil.

Methods: Autocorrelation (Moran index) and spatial association (Geographically weighted regression) models were used to explain the interrelationships between municipalities and the possible effects of socio-economic factors on rates.

Results: Two isolated clusters were revealed in the inner part of the state in sparsely inhabited municipalities. The spatial model (Geographically Weighted Regression) was able to explain 50% of the variations in COVID-19 cases. The variables *proportion of people with low income, percentage of rented homes, percentage of families in social programs, Gini index* and *running water* had the greatest explanatory power for the increase in infection by COVID-19.

Conclusions: Our results provide important information on socio-economic factors related to the spread of COVID-19 and can serve as a basis for decision-making in similar circumstances.

KEY WORDS

geographical dimensions, geographically weighted regression, health inequalities, infectious disease, pandemic

INTRODUCTION

Humanity is undergoing one of its biggest global health crises in its recent history, the social and economic damage caused by Coronavirus Disease 2019 (COVID-19). Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) showed pandemic potential in 2019, when several cases of pneumonia emerged simultaneously in Wuhan (China) [1,2]. WHO declared the COVID-19 outbreak a pandemic on 11 March 2020 and the world has been on alert ever since due to COVID-19's high transmissive potential, which has burdened health systems [3,4]. From China, the virus quickly

reached other countries [5], and the scientific community has been trying to understand the factors related to the advance of COVID-19 in a given region and pursued measures in order to contain its spread [6,7].

In Brazil, COVID-19 spread quickly despite well-established health and social protection systems [8]. In the state of Pernambuco in the Northeast region, the first case was confirmed on 12 March 2020. Despite measures taken to reduce the spread of COVID-19, the number of cases grew, but so did the research that was undertaken in response. Research has played an important role in shaping policies to contain epidemics, as analysis of spatial patterns of any

disease incidence or vector occurrence can provide information on the processes that determine the risk of transmission and identify priority areas for health actions [9].

Spatiotemporal analysis has been used in epidemiological studies over the last two decades [10-14]. Its tools can be employed to describe epidemics, to help generate hypotheses about the cause of diseases and to investigate control techniques [15]. In the case of COVID-19, spatiotemporal analysis served to identify areas with a greater likelihood of increase in cases, to direct contingency measures and optimise the efficiency of control programs [16]. For example, Ramírez-Aldana et al. [16] developed a statistical approach to describe how COVID-19 cases were spatially distributed, and how the socio-economic and climatic characteristics of Iranian provinces could predict the number of cases.

To ensure statistical models with robust explanatory power, it is necessary to choose explanatory variables that are adequate for the phenomenon under study. The selection process for these variables is one of the most complex steps in the development of a model [17], as the drivers of spatial patterns may be directly or indirectly related to the studied event [18]. The spatial pattern drivers can be exogenous, generated by factors or processes not directly related to the variable under study (for example, landscape heterogeneity), or endogenous, generated by factors or processes inherent to the variable or event under study (for example, vector dispersion) or a combination of both [18]. In case of COVID-19, as it is transmitted through the air, socio-economic variables related to population behaviour, given the need for mobility restrictions, have been successful in spatially explaining the number of cases of the disease [19-21]. One of the applied methods for assessing spatial relationships of epidemiological phenomena was Geographically weighted regression (GWR). GWR regression builds separate equations while incorporating the response variables, the confirmed cases of COVID-19 in this case, and the explanatory variables [22]. Thus spatial regression techniques including the GWR model [23] can be used to analyse the spatial causes for the increase in COVID-19 cases [24].

The aim of our study was to analyse the spatial distribution from the rates of COVID-19 cases and its association with socio-economic conditions in the state of Pernambuco/Brazil. Such an analysis not only investigates socio-economic factors that facilitate the spread of the disease in Pernambuco's municipalities but also variables that are spatially related to it, taking into account the spatial effect to obtain adequate inferences.

MATERIALS AND METHODS

Study area

The state of Pernambuco in the Northeast region of Brazil has a territorial extension of 98,075.90 km² and consists of 185 municipalities. The estimated population for the year 2020 was 9.6 million [25]. According to the Instituto Brasileiro

de Geografia e Estatística (IBGE), the body responsible for the national demographic census, the municipalities of Pernambuco can be grouped into five mesoregions: Região Metropolitana do Recife, Mata Pernambucana, Agreste Pernambucano, Sertão Pernambucano and São Francisco Pernambucano (Figure 1). The geographic space of the state of Pernambuco is fairly heterogeneous, with dynamic areas that are connected to the global economy and its technological innovations.

The Human Development Index (HDI) of Pernambuco in 2017 was 0.673, which is below Brazil's HDI (0.778) [26]. In 2017 19% (36) of Pernambuco's municipalities did not have a general sewage system, and more than 30% (57) of those suffer edfrom intermittent water supply for more than 6 months, caused mainly by drought and water source shortages [27].

Data source

Data on the accumulated confirmed cases of COVID-19 (Table 1) were obtained from the Centro de Informações Estratégicas de Vigilância em Saúde de Pernambuco (CIEVS-PE, <https://www.cievspe.com/>), for the period March 2020 to March 2021 to calculate the number of COVID-19 cases per 100,000 inhabitants (incidence rate) to ensure an adequate comparison between Pernambuco's municipalities. The explanatory variables used are presented in Table 1.

Vector files of Pernambuco's municipality boundaries were obtained through IBGE. This base, along with the data of the analysed variables, was processed in Qgis 3.10.5 software and the statistical analyses in R-4.1.0 software. In R software, the following libraries were used as follows: ggplot2, readxl, maptools, spdep, leaflet, RColorBrewer, tmap, tmaptools, spatialreg and spgwr.

Methods

This study consisted of the investigation of the variables capable of spatially explaining the distribution of rates of COVID-19 cases in Pernambuco. Based on this, three statistical approaches were used: the Moran Index, the Ordinary Least Square (OLS) and the Geographically Weighted Regression (GWR), due to its recurrent application in spatial analysis studies [28]. The Moran index estimates the magnitude of the spatial autocorrelation between areas, with significant values considered above 95% (p -value <0.05). In OLS, an analysis was performed to choose the model with the best coefficient of determination (R^2) and the Akaike information criteria [28]. The GWR model considers the spatial dependence and influence of neighbouring areas in independent local models [29] and allows for detection of heterogeneity and spatial variation of variables. The GWR model was initially proposed by Brunson et al. [29] to explore spatial nonstationarity, but it cannot adequately explain

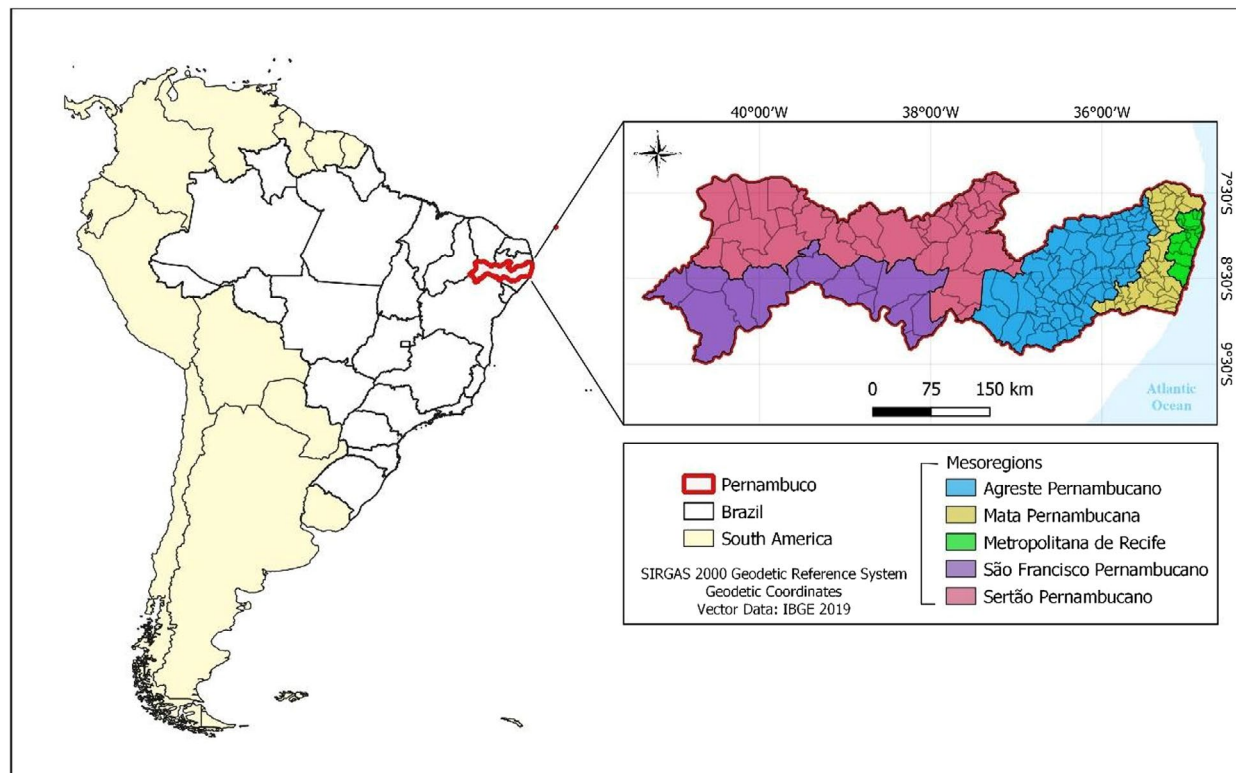


FIGURE 1 Study area

TABLE 1 Description of the variables with the respective categories, descriptions, measurement units and information sources

Variable	Category	Description	Unit	Source
Incidence rate of COVID-19	Response variable	Incidence rate of COVID-19.	Count/unit	[47]
Proportion of people with low income	Explanatory variable	Percentage of the population residing in the municipality with monthly per capita household income of up to half the minimum wage.	Count/unit	[25,48]
Percentage of rented households	Explanatory variable	Percentage of permanent private households rented in the municipality.	Count/unit	[25,48]
Percentage of families with the social program	Explanatory variable	Percentage of inhabitants of the municipality assisted by the Bolsa Família Program, which is a conditional cash transfer program for low-income families in Brazil.	Count/unit	[25,48]
Gini index	Explanatory variable	Degree of concentration of income in a given group.	Count/unit	[25,48]
Running water	Explanatory variable	Percentage of population in households with piped water.	Count/unit	[25,48]

the relationships between some sets of variables in a geographic region. Therefore, according to Carvalho et al. [30], it fits a regression model to each observed point, weighting all other observations as a function of the distance to that point. Regarding the parameters used in the GWR model, the bandwidth for the GWR model was obtained through cross-validation, which permits selection of the bandwidth that presents the smallest mean square errors. The adaptive method was used to assign the weights. Lastly, the Gaussian kernel was used in this application.

RESULTS

Overall, spatial autocorrelation of rates of COVID-19 cases could be observed in some parts of Pernambuco (Global Moran's $I = 0.318$, p -value < 0.01). Figure 2 reveals the grouping of municipalities with accumulated COVID-19 case rates between March 2020 and March 2021. On the one hand, two HH (high-high) clusters were observed, adding up to 20 municipalities, the largest of them with 11, in the North-central portion. On the other hand, two LL (low-low)

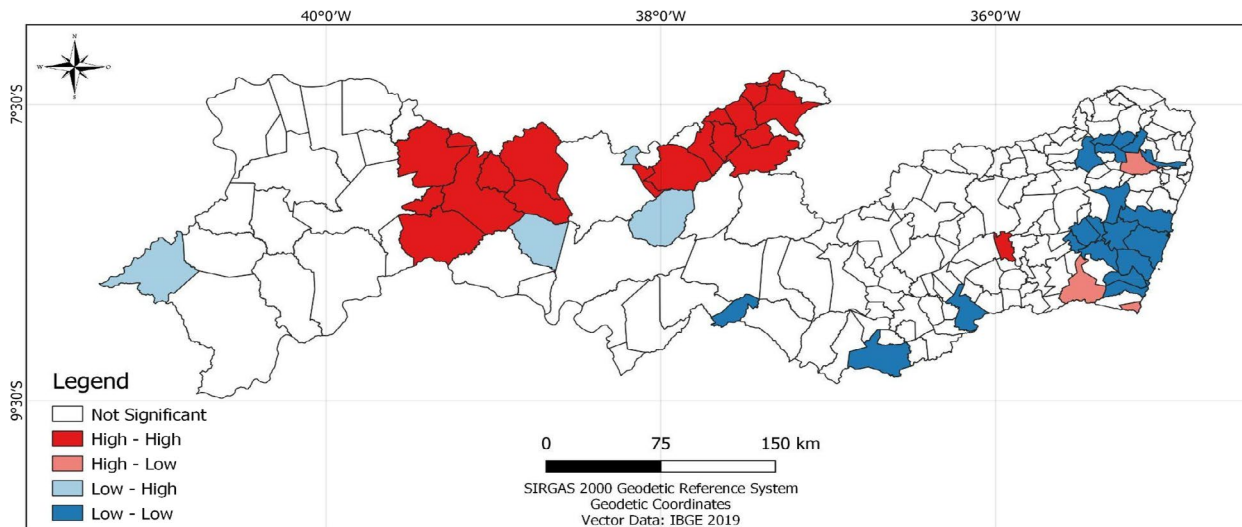


FIGURE 2 Map of the Local Indicators of Spatial Association (LISA) of the COVID-19 case rate

clusters were identified, adding up to 20 municipalities, the largest of them in the far East, also with 11 municipalities.

Table 2 shows the correlation matrix of the six covariates used in the OLS regression model with the incidence rate. The covariate *percentage of rented homes* has the highest correlation with *running water* (0.60). Overall the covariates did not show strong correlations with each other, which avoids possible multicollinearity problems.

Table 3 presents the covariates were used in the OLS regression model. The covariates *proportion of people with low income*, *percentage of rented homes* and *Gini index* are positively associated with the increase in the rate of COVID-19 cases in Pernambuco, while *percentage of eligible families in social programs* and *percentage of the population in households with running water* are negatively associated.

All variables had a VIF value <1.7, which indicates that no multicollinearity was observed in the OLS regression model. The adjusted R^2 indicates that the model was able to explain about 18% of the total variation in the rate of COVID-19

cases in Pernambuco. However, when performing the spatial autocorrelation test on the residuals, a significant positive value (Global Moran's $I = 0.287$, p -value = 0.000) was obtained. Thus, the OLS regression assumption that the residuals are independent was violated, consequently, for this case, the use of other techniques that can incorporate the spatial relationship in the model is required. Therefore, the GWR model was used to characterise the relationship between the rates of COVID-19 cases with the covariates. In Table 4 the GWR model coefficients for the COVID-19 case rate are presented in terms of statistical measures.

Based on the selection of the best model through the AIC method, the adjusted R^2 of the GWR model was 0.50. In other words, the spatial model is able to explain 50% of the variations in the rate of COVID-19 cases in Pernambuco. Thus, the GWR model has an important advantage over the OLS model. One of the benefits of using the GWR model is that, with the local coefficient of determination R^2 , it is possible to highlight the areas in which the model performs better. The best values for

TABLE 2 Correlation matrix between the variables used to build the statistical models

	Incidence rate of COVID-19	Proportion of people with low income	Percentage of rented homes	Percentage of families in a social program	Gini index	Running water
Incidence rate of COVID-19	1		0.25	-0.24	0.27	0.04
Proportion of people with low income		1	-0.35	0.09	0.17	
Percentage of rented households	0.25	-0.35	1	-0.07	-0.04	0.6
Percentage of families in a social program	-0.24	0.09	-0.07	1	-0.14	-0.06
Gini index	0.27	0.17	-0.04	-0.14	1	
Running water	0.04	-0.14	0.6	-0.06		1

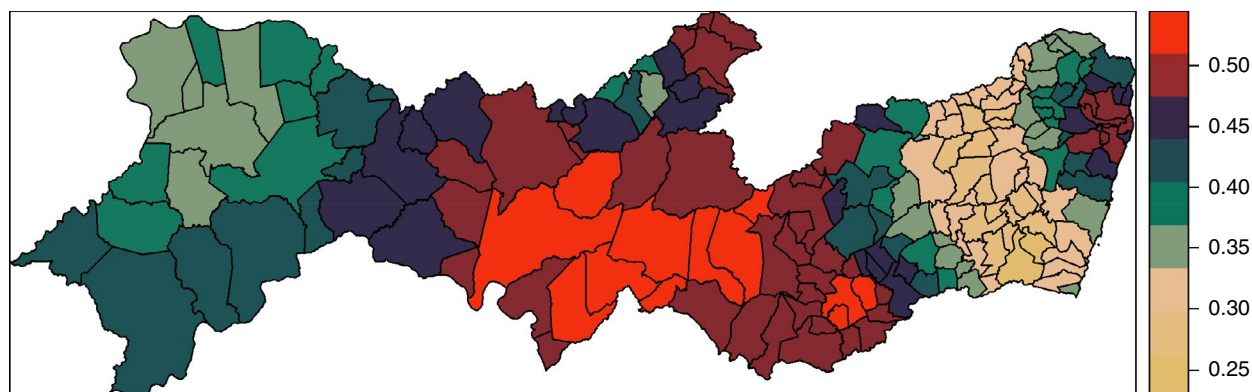
The gray color in the table has no meaning. It just highlights the diagonal of the correlation matrix.

TABLE 3 Adjustment of the OLS regression model

Variables	Estimate	Standard error	t-value	p-value	VIF
(Intercept)	1010.086	2038.163	0.496	0.620797	–
Proportion of people with low income	14.664	12.925	1.135	0.258066	1.195
Percentage of rented homes	108.155	24.243	4.461	1.44e–05	1.748
Percentage of families in a social program	–34.003	11.473	–2.964	0.003455	1.034
Gini index	8861.813	2463.902	3.597	0.000417	1.056
Running water	–17.686	7.598	–2.328	0.021060	1.564
R_2	0.18				
AIC	2671.72				

TABLE 4 Adjustment of the GWR model

Variables	Minimum	1st quartile	Median	3rd quartile	Maximum	Global
(Intercept)	–3.8685e+03	–1.7923e+03	2.3987e+03	3.9413e+03	7.2764e+03	1010.086
Proportion of people with low income	–2.7809e+01	–3.5117e–01	6.8567e+00	1.0700e+01	5.3753e+01	14.664
Percentage of rented homes	8.1351e+00	7.2446e+01	9.3828e+01	1.3628e+02	2.8156e+02	108.155
Percentage of families in a social program	–7.0915e+01	–3.4956e+01	–1.3829e+01	–6.6533e–01	1.7429e+01	–34.003
Gini index	–1.4174e+04	8.8172e+02	5.6597e+03	9.4660e+03	1.3967e+04	8861.813
Running water	–3.3938e+01	–1.9638e+01	–1.4184e+01	–4.8296e+00	4.6326e+01	–17.686
R^2	0.50					
AIC	3192.322					

FIGURE 3 R^2 of the GWR model

R^2 in the state of Pernambuco are the municipalities that constitute the regions: The Central portion of the Metropolitan region of Recife mesoregion, the central area of São Francisco and Sertão Pernambucano mesoregions and the Northwestern portion of the Pernambuco countryside. In these mesoregions, R^2 values are mostly >0.40 (Figure 3).

Figure 4a shows the estimated coefficients for the GWR model considering the variable *proportion of people with low income*. As noted, the highest coefficients were observed to be located in the Western portion of the Sertão Pernambucano and Sertão de São Francisco mesoregions (Figure 4a). The coefficients were also significant in these regions (p -value <0.05 , see Figure 4b).

Regarding the variable *percentage of rented homes*, the highest values of the coefficient are located in the western portion of the Sertão de São Francisco and Sertão Pernambucano mesoregions (Figure 5a). For these regions, the coefficients were statistically significant (p -value <0.05 , Figure 5b).

As for the variable *percentage of eligible families in a social program* (Figure 6a), the highest coefficients were observed in the Western portion of the Sertão Pernambucano mesoregion, in the Northern portion of the Zona da Mata Pernambucana mesoregion and the Eastern portion of the Agreste Pernambuco. Some regions had significant coefficients (Figure 6b).

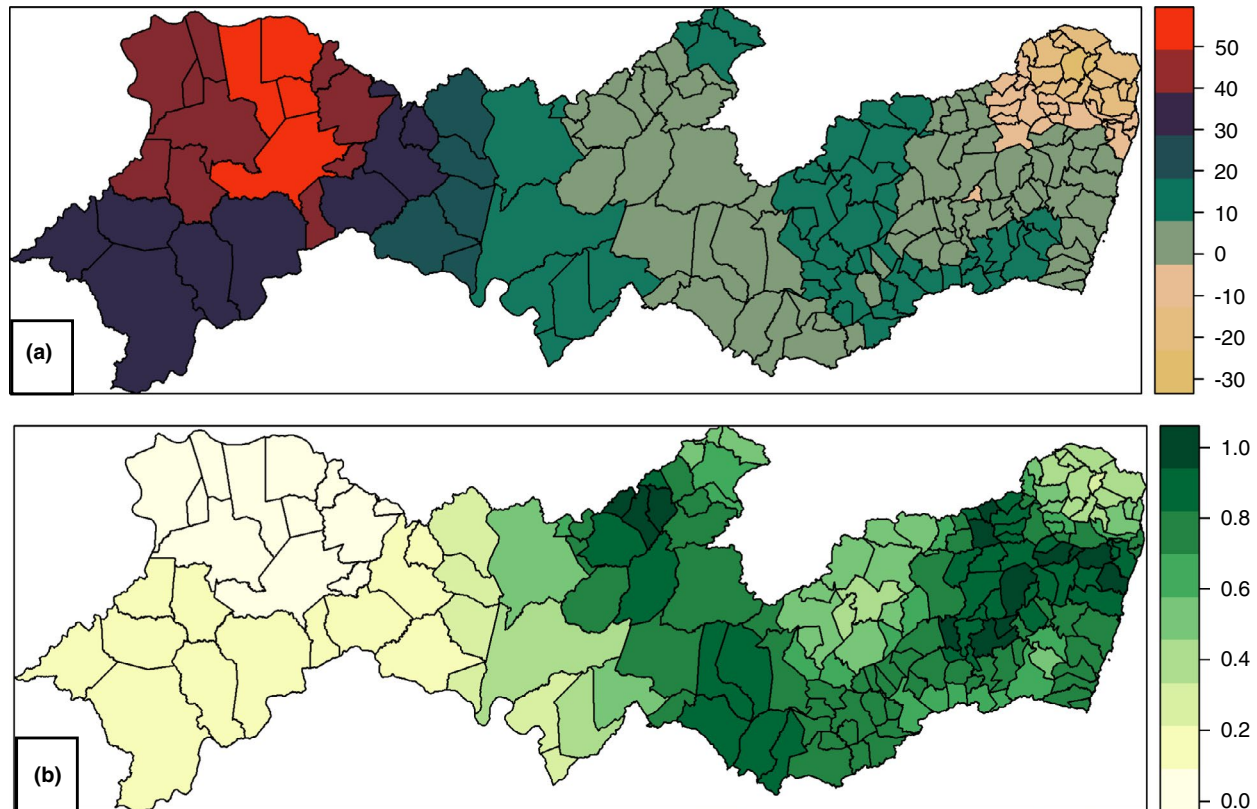


FIGURE 4 (a) Coefficients and (b) p -values for the Proportion of people with low-income variable

Figure 7a shows that the highest coefficient values for the Gini Index variable were concentrated in the Northeastern portion of the Zona da Mata Pernambucana and the Northeastern portion of the Agreste Pernambucana. Figure 7b reveals that these regions have significant coefficients.

The highest values for the *running water* variable were seen in the Northeastern portion of the central mesoregion of the Sertão Pernambucano. In Figure 8b shows that these regions have significant coefficients.

DISCUSSION

We found great spatial variability in the registered cases of COVID-19, and the robustness of the spatial models allows for the verification of the local R^2 and thus identification of areas where the model performs better. Epidemiological studies that involve spatial analyses can help identify hot spots with a higher risk of transmission, which is essential for the prevention and control of the epidemic, and can help explain geographic disparities in the prevalence [31].

The pandemic advanced beyond the big cities to small municipalities, places where historically there are gaps in care and where small hospitals predominate [32]. According to Noronha et al. [32], small hospitals lack the specialised resources to treat patients with severe symptoms of COVID-19, a factor that may be related to the high transmission rates. Furthermore, according to Quinino et al. [33] in the first

wave (from March 3rd to June 30th, 2021) cases were recorded in densely populated areas, in the second wave mobility between municipalities favoured the spread of COVID-19 cases throughout the state.

Thus, cities in the interior, especially the Sertão Pernambucano and Sertão do São Francisco Mesoregions, have historically presented the lowest Human Development Index (HDI), which is based on a series of factors that can have a direct impact to the well-being of the population [28]. This scenario may accentuate social differences in this region of the country through the contraction of the economy Josephson et al. [34]. According to Karaye and Horney [22], this is in part due to endemic inequalities within these populations such as income, education, nutrition, transportation, housing, jobs, the environment, psychosocial stress and medical care, all of which contribute to poor health.

Thus, as the percentage of the low-income population increases, so does the rate of cases. In general, according to Rocha et al. [8], in Brazil existing socio-economic inequalities have affected the course of the epidemic more than conventional risk factors (of age, health status and others). In Pernambuco, this holds true as the highest values of the coefficient (percentage of the low-income population) were observed in parts of the poorest mesoregions of the state, Sertão de São Francisco and Sertão Pernambucano. Our results are corroborated by the work of Sung [35], who found a positive relationship between the percentage of people with low income and cases of COVID-19 in the USA. Finch & Finch [36] revealed that during the first weeks of the pandemic the

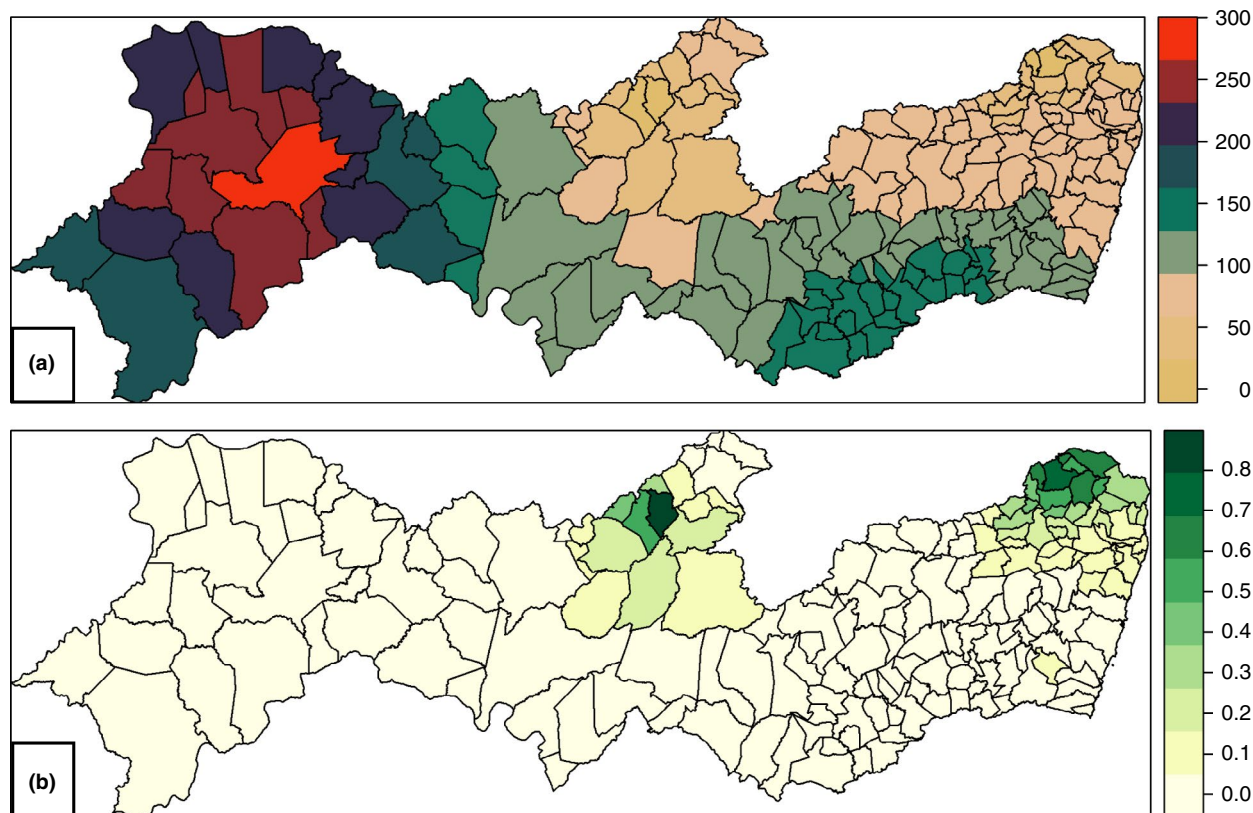


FIGURE 5 (a) Coefficients and (b) *p*-values for the *percentage of rented homes* variable

most disadvantaged counties in the United States had the largest number of confirmed cases of COVID-19.

Regarding the percentage of rented homes, our study demonstrates a positive effect on the rates of COVID-19 cases. This socio-economic aspect is not only associated with the inhabitants' income, but also to human mobility, as inhabitants do not settle in the same place, resulting in a large flow of residents in the same area, which reduces social distancing. Our results corroborate those of Dutta et al. [37] in India, where the urbanisation rate had a direct relationship with the increase in COVID-19 cases.

The percentage of families in a social program had a negative effect on the rates of COVID-19 cases. Thus, the smaller the number of families assisted by the Bolsa Família social program, the higher the rate of cases. Similarly, Rocha et al. [8] observed a negative correlation between social assistance coverage and COVID-19 cases. This finding demonstrates that targeted policies and actions are needed in order to protect those with greater socio-economic vulnerability. Despite this, there was inertia in central government [8]. Two factors blocked the success of the actions of the central government in Brazil: (1 frequent changes in health leadership, which generated administrative instability), and (2 the political context, which posed a major challenge to the response [8]).

The Gini Index, which reflects income inequality, had a positive relationship with the rates of the COVID-19 cases. Therefore, the greater the income inequality, the higher the rate of COVID-19 cases. Elgar et al. [38] also observed that

the increase in income inequality was associated with the increase in cases and mortality rates due to COVID-19 in 84 countries, including Brazil. For many decades, Brazil has been showing a high level of income inequality [39] that sets the country the challenge of two epidemics: COVID-19 and social inequality itself, which puts millions of Brazilians at risk of returning to extreme poverty [40].

Running water had a negative effect on the rate of COVID-19 cases, that is, the less access to piped water, the higher the rate of cases. This result is alarming, as it is well known that personal and environmental hygiene is a factor in the spread of COVID-19. Hand hygiene is known for being the best way to reduce the transmission of viruses and bacteria [41]. This result corroborates those of Ekumah et al. [42] who evaluated 25 countries in sub-Saharan Africa and found that individuals in households without water and access to sanitation are more likely to spread the COVID-19 virus. Thus, according to Casazza [43], the lack of access to water is accompanied by other conditions that exclude certain parts of society from COVID-19 prevention and render them more susceptible. According to WHO [44], hand hygiene is one of the fundamental pillars to interrupt the transmission of COVID-19. Water scarcity is recognised in part of the study area, where arid and semi-arid ecosystems predominate [45], especially in the Sertão Pernambucano and Sertão do São Francisco mesoregions. Thus water scarcity not only influences hygiene but indirectly also the economy, poverty, freedom of movement and inequality [46].

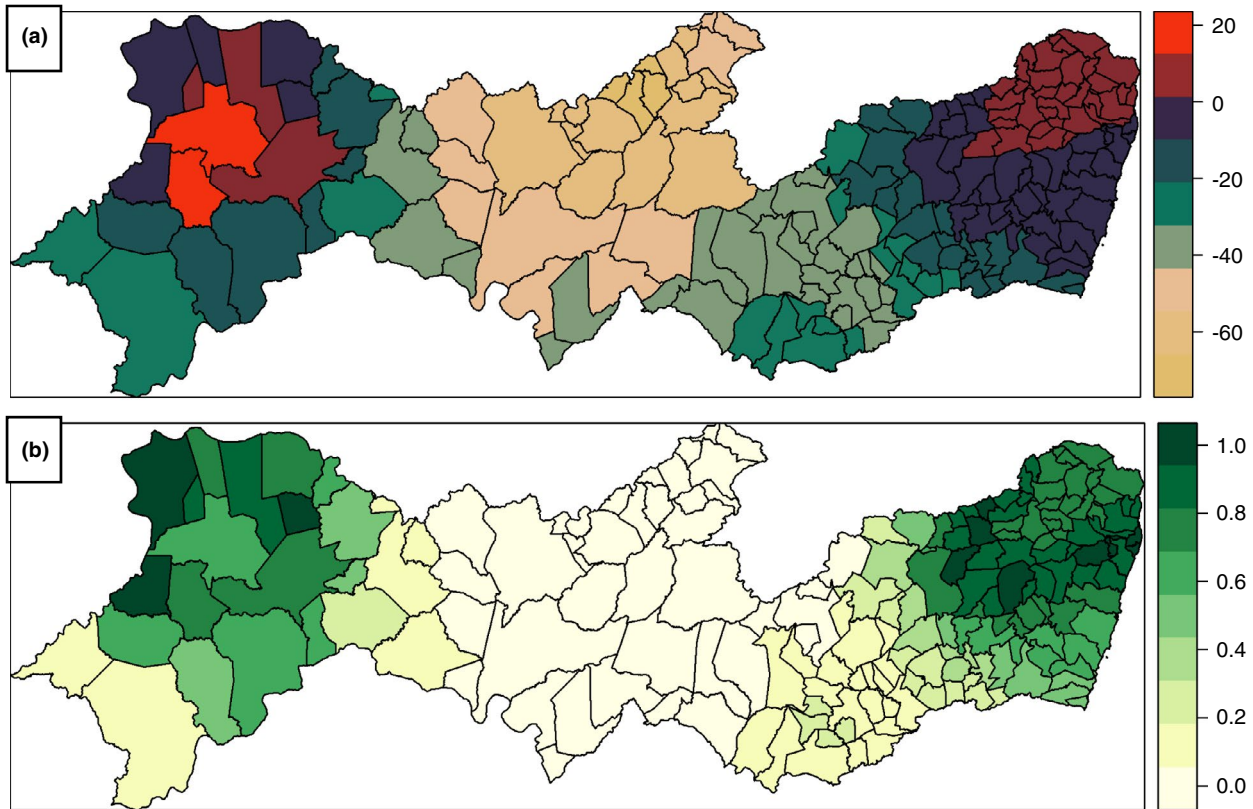


FIGURE 6 (a) Coefficients and (b) *p*-values for the variable *percentage of eligible families in a social program*

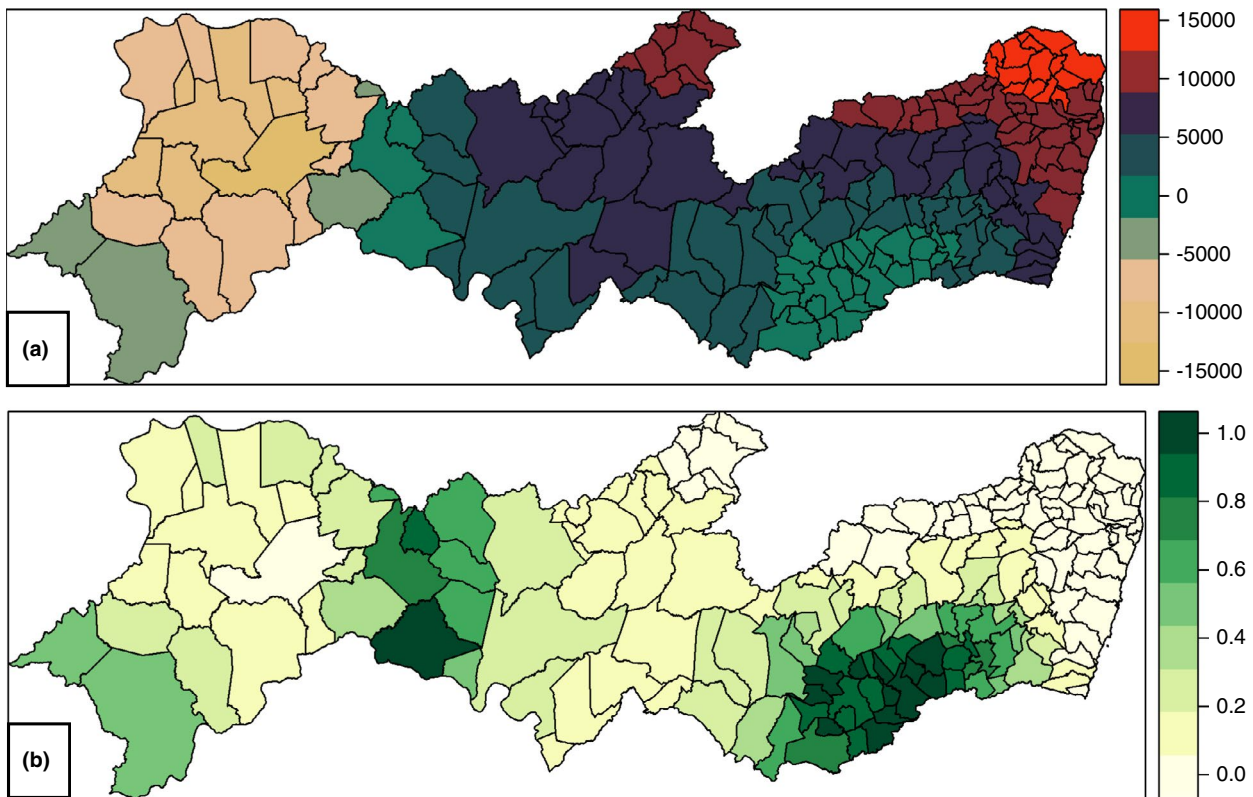


FIGURE 7 (a) Coefficients and (b) *p*-values for the Gini Index variable

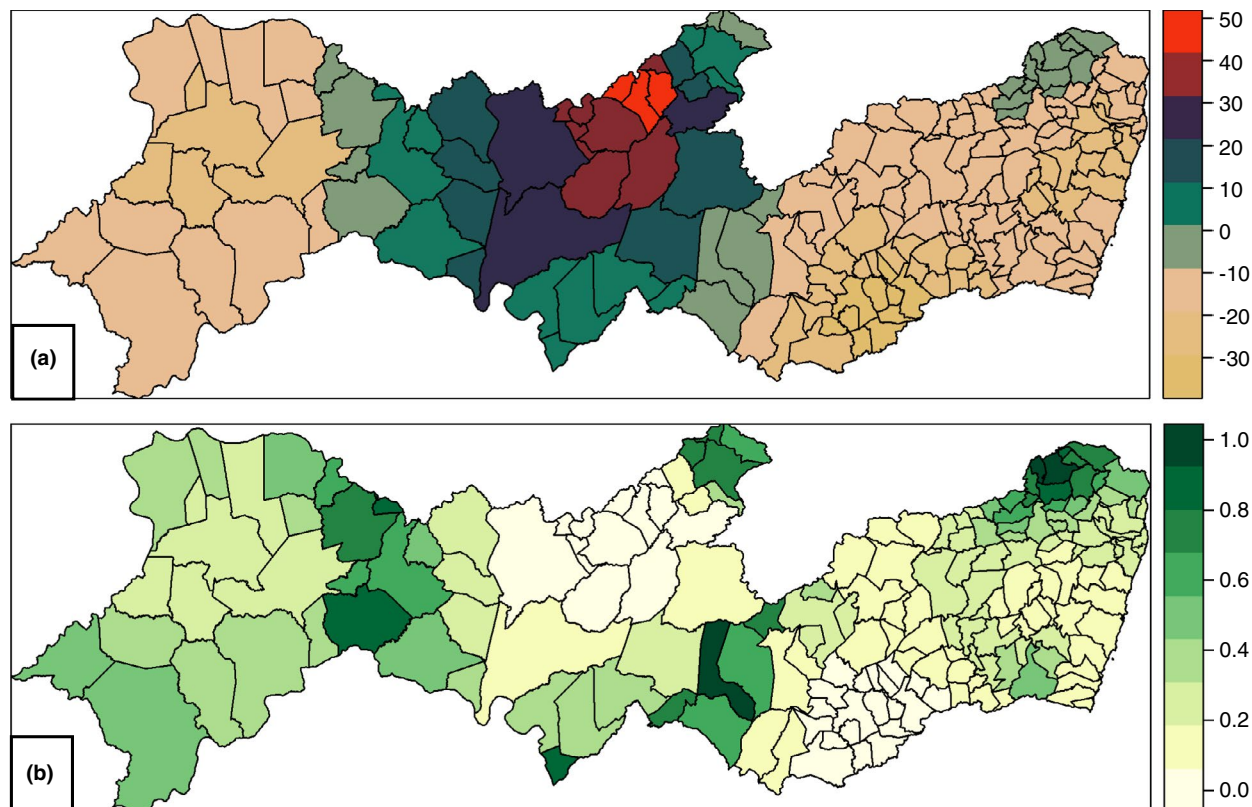


FIGURE 8 (a) Coefficients and (b) *p*-values for the Piped Water variable

This study has limitations in terms of the underlying demographic information. Up-to-date census data should already be available, but the central government of Brazil suspended the census in 2020 due to a lack of resources for IBGE and in 2021 due to the COVID-19 pandemic, which impacts the analyses of socio-economic issues. Another limitation is that it was not possible to determine the accuracy of the COVID-19 case data, specifically for the early stages of the pandemic, when data collection protocols were still being defined.

CONCLUSIONS AND RECOMMENDATIONS


We evaluated the socio-economic factors that are spatially associated with the advance of COVID-19 in Pernambuco, a Brazilian state with a high rate of cases. We conclude that the rates of COVID-19 cases were predominant in the interior part of the state, in sparsely inhabited municipalities, in two isolated clusters (Central and Northeastern portion of the Sertão Pernambucano mesoregion and in the Central portion of the São Francisco Pernambucano mesoregion); in general, the spatial model was able to explain 50% of the variations in the rates of COVID-19 cases in the state of Pernambuco; the variables *proportion of people with low income, percentage of rented homes, percentage of families in a social program, Gini Index* and *running water* could best explain the spread of COVID-19.

For future studies, the interactions and support given to the local health system should be emphasised: number of Emergency Care Units available, number of Intensive Care Units for the treatment of COVID-19, protocols adopted for the use of masks, etc. Assessing the thoroughness of containment measures adopted by government officials, measures adopted for social distancing, incentives to reduce mobility, social subsidies, etc. may shed light on the environmental determinants that may have an impact on the cases of COVID-19.


ACKNOWLEDGEMENTS


We thank the Department of Cartographic Engineering of the Federal University of Pernambuco.

ORCID

Carlos Fabricio Assunção da Silva  <https://orcid.org/0000-0001-7009-8996>

Mayara Costa Silva  <https://orcid.org/0000-0001-9086-1960>

Alex Mota dos Santos  <https://orcid.org/0000-0002-5156-3968>

Anderson Paulo Rudke  <https://orcid.org/0000-0003-2970-5453>

Cristine Vieira do Bonfim  <https://orcid.org/0000-0002-4495-9673>

Gabriela Tobias Portis  <https://orcid.org/0000-0001-8458-2359>

Pedro Monteiro de Almeida Junior  <https://orcid.org/0000-0003-0887-1639>

Maria Beatriz de Santana Coutinho  <https://orcid.org/0000-0001-7669-8562>

REFERENCES

1. Wu F, Zhao S, Yu B, Chen YM, Wang W, Song Z-G, et al. A new coronavirus associated with human respiratory disease in China. *Nature*. 2020;579:265–9. <https://doi.org/10.1038/s41586-020-2008-3>
2. Zhou P, Yang X-L, Wang X-G, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*. 2020;579:270–3. <https://doi.org/10.1038/s41586-020-2012-7>
3. Andrade RO. Covid-19 is causing the collapse of Brazil's National Health Service. *BMJ*. 2020;368:m3032. <https://doi.org/10.1136/bmj.m3032>
4. Blumenthal D, Fowler EJ, Abrams M, Collins SR. Covid-19 — Implications for the Health Care System. *N Engl J Med*. 2020;383:1483–8. <https://doi.org/10.1056/NEJMs2021088>
5. Zhang X-A, Fan H, Qi R-Z, Zheng W, Zheng K, Gong J-H, et al. Importing coronavirus disease 2019 (COVID-19) into China after international air travel. *Travel Med Infect Dis*. 2020;35:101620. <https://doi.org/10.1016/j.tmaid.2020.101620>
6. Dutta A, Fischer HW. The local governance of COVID-19: disease prevention and social security in rural India. *World Dev*. 2021;138:105234. <https://doi.org/10.1016/j.worlddev.2020.105234>
7. Rudke AP, Martins JA, de Almeida DS, Martins LD, Beal A, Hallak R, et al. How mobility restrictions policy and atmospheric conditions impacted air quality in the State of São Paulo during the COVID-19 outbreak. *Environ Res*. 2021;198:111255. <https://doi.org/10.1016/j.envres.2021.111255>
8. Rocha R, Atun R, Massuda A, Rache B, Spinola P, Nunes L, et al. Effect of socioeconomic inequalities and vulnerabilities on health-system preparedness and response to COVID-19 in Brazil: a comprehensive analysis. *Lancet Glob Health*. 2021;9(6):e782–92. [https://doi.org/10.1016/S2214-109X\(21\)00081-4](https://doi.org/10.1016/S2214-109X(21)00081-4)
9. Reisen WK. Landscape epidemiology of vector-borne diseases. *Annu Rev Entomol*. 2010;55:461–83. <https://doi.org/10.1146/annurev-ento-112408-085419>
10. Farnsworth ML, Ward MP. Identifying spatio-temporal patterns of transboundary disease spread: examples using avian influenza H5N1 outbreaks. *Vet Res*. 2009;40(3):1–14. <https://doi.org/10.1051/vetre/2009003>
11. Fauver JR, Petrone ME, Hodcroft EB, Shioda K, Ehrlich HY, Watts AG, et al. Coast-to-coast spread of SARS-CoV-2 during the early epidemic in the United States. *Cell Press*. 2020;181:990–6. <https://doi.org/10.1016/j.cell.2020.04.021>
12. Kim B, Rundle AG, Goodwin ATS, Morrison CN, Branas CC, El-Sadr W, et al. COVID-19 testing, case, and death rates and spatial socio-demographics in New York City: an ecological analysis as of June 2020. *Health Place*. 2021;68:102539. <https://doi.org/10.1016/j.healthplace.2021.102539>
13. Lipsitt J, Chan-Golston AM, Liu J, Su J, Zhu Y, Jerrett M. Spatial analysis of COVID-19 and traffic-related air pollution in Los Angeles. *Environ Int*. 2021;153:106531. <https://doi.org/10.1016/j.envint.2021.106531>
14. Stevenson M, Robinson TP, Stevens KB, Rogers DJ, Clements AC. *Spatial analysis in epidemiology*. Oxford: Oxford University Press; 2008.
15. Ward MP, Carpenter TE. Techniques for analysis of disease clustering in space and in time in veterinary epidemiology. *Prev Vet Med*. 2000;45(3–4):257–84. [https://doi.org/10.1016/S0167-5877\(00\)00133-1](https://doi.org/10.1016/S0167-5877(00)00133-1)
16. Ramírez-Aldana R, Gomez-Verjan JC, Bello-Chavolla OY. Spatial analysis of COVID-19 spread in Iran: insights into geographical and structural transmission determinants at a province level. *PLoS Negl Trop Dis*. 2020;14(11):e0008875. <https://doi.org/10.1371/journal.pntd.0008875>
17. Ratner B. Variable selection methods in regression: ignorable problem, outing notable solution. *J Targeting Meas Anal Mark*. 2010;18:65–75. <https://doi.org/10.1057/jt.2009.26>
18. Dale MR, Fortin MJ. *Spatial analysis: a guide for ecologists*. Cambridge: Cambridge University Press; 2014.
19. Birenbaum-Carmeli D, Chassida J. Covid-19 in Israel: socio-demographic characteristics of first wave morbidity in Jewish and Arab communities. *Int J Equity Health*. 2020;19:153. <https://doi.org/10.1186/s12939-020-01269-2>
20. Ehlert A. The socio-economic determinants of COVID-19: a spatial analysis of German county level data. *Socioecon Plann Sci*. 2021;78:101083. <https://doi.org/10.1016/j.seps.2021.101083>
21. Varkey RS, Joy J, Sarmah G, Panda PK. Socioeconomic determinants of COVID-19 in Asian countries: an empirical analysis. *J Public Aff*. 2020;21:1–10. <https://doi.org/10.1002/pa.2532>
22. Karaye IM, Horney JA. The impact of social vulnerability on COVID-19 in the US: an analysis of spatially varying relationships. *Am J Prev Med*. 2020;59(3):317–25. <https://doi.org/10.1016/j.amepre.2020.06.006>
23. Fotheringham AS, Brunsdon C, Charlton M. *Geographically weighted regression: the analysis of spatially varying relationships*. John Wiley & Sons. 2002.
24. Wu X, Zhang J. Exploration of spatial-temporal varying impacts on COVID-19 cumulative case in Texas using geographically weighted regression (GWR). *Environ Sci Pollut Res*. 2021;1–15. <https://doi.org/10.1007/s11356-021-13653-8>
25. IBGE Instituto Brasileiro de Geografia e Estatística. 2021. Available from <https://www.ibge.gov.br/>. Accessed July 16, 2021.
26. Pesquisa Nacional por Amostra de Domicílios. 2020. Available from <https://www.ibge.gov.br/estatisticas/sociais/trabalho/9171-pesquisa-nacional-por-amostra-de-domicilios-continua-mensal.html>. Accessed July 16, 2021.
27. Pesquisa Nacional de Saneamento Básico. 2020. Available from <https://www.ibge.gov.br/estatisticas/multidominio/meio-ambiente/9073-pesquisa-nacional-de-saneamento-basico.html>. Accessed July 16, 2021.
28. Silva C, Melo S, Santos A, Junior PA, Sato S, Santiago K, et al. Spatial modeling for homicide rates estimation in Pernambuco State-Brazil. *ISPRS Int J Geo-Inf*. 2020;9(12):740–58. <https://doi.org/10.3390/ijgi9120740>
29. Brunsdon C, Fotheringham AS, Charlton ME. Geographically weighted regression: a method for exploring spatial nonstationarity. *Geogr Anal*. 1996;28(4):281–98. <https://doi.org/10.1111/j.1538-4632.1996.tb00936.x>
30. Carvalho LEX, Silva HN, Loureiro CFG, Meneses HB. Regressão linear geograficamente ponderada em ambiente SIG. *TRANSPORTES*. 2006;14(2):18–26. <https://doi.org/10.14295/transportes.v14i2.65>
31. Sun F, Matthews SA, Yang TC, Hu MH. A spatial analysis of the COVID-19 period prevalence in US counties through June 28, 2020: where geography matters? *Ann Epidemiol*. 2020;52:54–9. <https://doi.org/10.1016/j.annepidem.2020.07.014>
32. Noronha KVMSD, Guedes GR, Turra CM, Andrade MV, Botega L, Nogueira D, et al. The COVID-19 pandemic in Brazil: analysis of supply and demand of hospital and ICU beds and mechanical ventilators under different scenarios. *Cad Saude Publica*. 2020;36:e00115320. <https://doi.org/10.1590/0102-311X00115320>
33. Quinino LRDM, Vasconcellos FHM, Diniz IS, Aguiar LR, Ramos YTD, Bastiani FD. Spatial and temporal aspects and factors associated with the spread of Covid-19 to the interior of the State of Pernambuco, Brazil. *Ciênc Saude Colet*. 2021;26:2171–82. <https://doi.org/10.1590/1413-81232021266.00642021>
34. Josephson A, Kilic T, Michler JD. Socioeconomic impacts of COVID-19 in low-income countries. *Nat Hum Behav*. 2021;5(5):557–65. <https://doi.org/10.1038/s41562-021-01096-7>
35. Sung B. A spatial analysis of the effect of neighborhood contexts on cumulative number of confirmed cases of COVID-19 in US Counties

- through October 20 2020. *Prev Med.* 2021;147:106457. <https://doi.org/10.1016/j.ypmed.2021.106457>
36. Finch WH, Finch MEH. Poverty and Covid-19: rates of incidence and deaths in the United States during the first 10 weeks of the pandemic. *Front Sociol.* 2020;5:47. <https://doi.org/10.3389/fsoc.2020.00047>
 37. Dutta I, Basu T, Das A. Spatial analysis of COVID-19 incidence and its determinants using spatial modeling: a study on India. *Environ Challenges.* 2021;4:100096. <https://doi.org/10.1016/j.envc.2021.100096>
 38. Elgar FJ, Stefaniak A, Wohl MJ. The trouble with trust: time-series analysis of social capital, income inequality, and COVID-19 deaths in 84 countries. *Soc Sci Med.* 2020;263:113365. <https://doi.org/10.1016/j.socscimed.2020.113365>
 39. Salata A. Inequalities and the Brazilian new democracy: income distribution between classes in recent decades. *Sociol Antropol.* 2016;6:181–208. <https://doi.org/10.1590/2238-38752016v6i8>
 40. Oriá F. Between two epidemics: Coronavirus and inequality in Brazil. *Global Policy, Brazil*; 2020. Available from <https://www.globalpolicyjournal.com/blog/21/12/2020/between-two-epidemics-coronavirus-and-inequality-brazil-conversation-ggf-2035>. Accessed July 16, 2021.
 41. Donde OO, Atoni E, Muia AW, Yillia PT. COVID-19 pandemic: water, sanitation and hygiene (WASH) as a critical control measure remains a major challenge in low-income countries. *Water Res.* 2020;191:116793. <https://doi.org/10.1016/j.watres.2020.116793>
 42. Ekumah B, Armah FA, Yawson DO, Quansah R, Nyieku FE, Owusu SA, et al. Disparate on-site access to water, sanitation, and food storage heighten the risk of COVID-19 spread in Sub-Saharan Africa. *Environ Res.* 2020;189:109936. <https://doi.org/10.1016/j.envres.2020.109936>
 43. Casazza IF. O acesso à água e os excluídos da prevenção à Covid-19. *Fiocruz.* 2020. Available from <http://coc.fiocruz.br/index.php/pt/todas-as-noticias/1789-o-acesso-a-agua-e-os-excluidos-da-prevencao-ao-covid-19.html#YPG2hz2SmUk>. Accessed July 16, 2021.
 44. WHO. Water, sanitation, hygiene and waste management for the COVID-19 virus. 2020. Available from <https://www.who.int/publications/i/item/WHO-2019-nCoV-IPC-WASH-2020.4>. Accessed July 16, 2021.
 45. Schulz K, Voigt K, Beusch C, Almeida-Cortez JS, Kowarik I, Walz A, et al. Grazing deteriorates the soil carbon stocks of Caatinga forest ecosystems in Brazil. *For Ecol Manage.* 2016;367:62–70. <https://doi.org/10.1016/j.foreco.2016.02.011>
 46. Boretti A. Covid19 pandemic as a further driver of water scarcity in Africa. *GeoJournal.* 2020;21(1):1–28. <https://doi.org/10.1007/s10708-020-10280-7>
 47. CIEVS-PE Centro de Informações Estratégicas de Vigilância em Saúde de Pernambuco. 2021. Available from <https://www.cievspe.com/>. Accessed July 16, 2021.
 48. SIDRA Instituto Brasileiro de Geografia e Estatística. 2020. Available from <https://sidra.ibge.gov.br/tabela/5457>. Accessed July 16, 2021.

How to cite this article: da Silva CFA, Silva MC, dos Santos AM, Rudke AP, do Bonfim CV, Portis GT, et al. Spatial analysis of socio-economic factors and their relationship with the cases of COVID-19 in Pernambuco, Brazil. *Trop Med Int Health.* 2022;27:397–407. <https://doi.org/10.1111/tmi.13731>