

Explaining COVID-19 contagion in Portuguese municipalities using spatial autocorrelation models

Paulo Mourao*^a/ Ricardo Bento^{b1}

^a Universidade do Minho-Escola de Economia e Gestão, Departamento de Economía, Campus de Gualtar, 4710-057 Braga, Portugal

^b Universidade de Trás-os-Montes e Alto Douro (UTAD), Centro de Estudos Transdisciplinares para o Desenvolvimento (CETRAD), Departamento de Engenharias, Quinta de Prados, 5000-801 Vila Real, Portugal

Received: 30 June 2020 / Accepted: 10 April 2021

Abstract

This paper investigates the pattern of COVID-19 contagion in Portuguese municipalities from March 23rd to April 5th (the exponential phase). We have resorted to spatial autocorrelation models to discuss how the conglomeration of highly infectious spaces has also contributed to infecting neighbouring municipalities. We have used several indicators for the contagion of COVID-19 from the number of infectious individuals to rates of infectious. As explicative variables, additionally to spatial proximity, we also considered population density, the share of the elderly population as well as the length of municipal perimeter/border. Our results show that highly dense municipalities tended to contaminate close areas. Lengthier perimeters also showed a positive effect on the contagious indicators for a given municipality.

Keywords

COVID-19 / Portugal / Spatial contagion / Spatial autocorrelation.

Explicando o contaxio da COVID-19 nos concellos portugueses usando modelos de autocorrelación espacial

Resumo

Este artigo investiga o patrón de contaxio da COVID-19 nos municipios portugueses entre o 23 de marzo e o 5 de abril (a fase exponencial). Recorremos a modelos de autocorrelación espacial para analizar como a veciñanza de espazos altamente infecciosos tamén contribuíu a infectar os municipios próximos. Utilizamos varios indicadores para o contaxio da COVID-19, desde o número de individuos infecciosos ata as taxas de infección. Como variables explicativas, ademais da proximidade espacial, tamén consideramos a densidade de poboación, a proporción da poboación de persoas maiores e a lonxitude do perímetro/fronreira municipal. Os nosos resultados indican que os municipios altamente densos tenden a contaminar áreas próximas. Os perímetros máis longos tamén mostraron un efecto positivo nos indicadores contaxiosos para un municipio determinado.

Palabras clave

COVID-19 / Portugal / Contaxio espacial / Autocorrelación espacial.

JEL Codes: G01, H12.

* P. Mourao: paulom@eeg.uminho.pt (corresponding author), R. Bento: rbento@utad.pt

¹ P. Mourao acknowledges financial support from Fundação para a Ciência e a Tecnologia- grant number UIDB/03182/2020. R. Bento acknowledges financial support from Fundação para a Ciência e a Tecnologia- grant number UIDB/04011/2020. The authors acknowledge the suggestions provided by two anonymous referees. Remaining limitations are authors' exclusive ones.

1. Introduction and motivation

The first case of a Portuguese citizen infected with COVID-19 was identified on February 22nd, 2020; this was the case of a citizen employed on a cruise docked off Japan. However, it was only 9 days before the first two infected people were registered in Portuguese territory –two cases of 'import' of the disease insofar as these cases were associated with visits by these citizens to Valencia and northern Italy.

If on 10/03/2020 the Daily Bulletins of the General Directorate of Health –DGS– (official body for making data available in Portugal on the phenomenon) registered 41 positive cases, the evolution of cases of the Portuguese infected in the following weeks showed an evolution close to one sigmoid function that has been identified in other examples of European countries: 448 (03/17/2020), 2362 (03/24/2020), 7443 (03/31/2020), and 11730 (04/06/2020).

Until 03/23/2020, DGS Daily Bulletins disaggregated positive cases (as well as recovered cases and fatalities by Basic regions for the application of regional policies (NUTS 2) in the country. As of 03/24/2020, the referred bulletins started to reveal the aforementioned distribution by municipality. Since then, it has been possible to observe the spatial distribution of cases of citizens whose tests for the existence of infection by COVID-19 were positive.

There has been no scientific publication that applied spatial autocorrelation analysis techniques to the evolution observed in the Portuguese territory. We consider that such an analysis is necessary for three reasons. First, the urgency of the fight against COVID-19. The current pandemic has shown the weakness of the various health policy action systems around the globe, as long as an effective vaccine has been distributed to the population by December 2020, the most appropriate prophylactic reaction has been related to the reduction of contagion networks, by limiting the mobility of the main carriers of the virus-humans. Understanding how the dynamics of spatial contagion are associated with the dynamics of contagion groups is one of the objectives of this investigation. Second, one of the causes of contagion from the earliest researches was associated with social proximity between human beings. Therefore, understanding how human densities –both demographic densities (inhabitants per square kilometer) and differentiation of age cohorts– requires an analysis of the spatial distribution (Oliva Denis & Aldrey Vázquez, 2018). Third, contributing to the delimitation of future epidemic outbreaks, as understanding how such outbreaks are concentrated in certain spaces and due to certain variables of the groups located in those spaces is an essential step to avoid pandemic shocks like the current one.

The proliferation of the COVID-19 pandemic found spatial analysis techniques to be an important resource to support the design of geographic-administrative policies. On the one hand, there is the perception that neighboring spaces (even neighboring municipalities in different countries) tend to exchange significant demographic flows. Several studies have proved this in a clear way –from tourist flows to students, investors, or for consumption purposes. Thus, with the COVID-19 pandemic associated with the fluidity of demographic movements, spatial analysis resources made it possible to observe the extent to which dimensions such as contiguity help to explain contagion and in the case of validation to what extent measures such as border control or sanitary fences are effective in reducing contagion rates. Then, studies like Gulyiev (2020) that focused on the initial cases in Wuhan (China) or Sun, Matthews, Yang & Hu (2020) in the United States of America, quickly explored some spatial regression techniques that validated the hypothesis that, in these countries, spatial contiguity effects were significant. Finally, works such as those by Krisztin, Piribauer & Wögerer (2020) have even proved the influence of spatial contiguity enhanced by international connections, namely flights, for the proliferation of the virus.

Given the strategic relevance of spatial analysis for pursuing the contagion across Portuguese regions, we will consider some explicative dimensions which follow previous research using these methods for other countries (Mollalo, Vahedi & Rivera, 2020; Páez, López, Menezes, Cavalcanti & Pitta, 2020; or Sannigrahi, Pilla, Basu, Sarkar Basu & Molter, 2020). So, we will consider municipal

perimeter, the ageing index and the population density. As we will detail in the respective section, the municipal perimeter suggests that larger municipalities (in terms of area) tend to see more people's movements within the surrounding spaces. Additionally, a higher population density also enhances these opportunities. Finally, a higher value of the ageing index is associated with a higher proportion of aged people which has been found to correlate with higher exposures to the risk of COVID-19 contagion.

Mollalo et al. (2020) investigated county-level variations of disease incidence across the continental United States, compiling environmental, socioeconomic, topographic, and demographic variables that could explain the spatial variability of disease incidence, including the percentage of older adults. Similar studies were carried by Mansour, Al Kindi, Al-Said, Al-Said & Atkinson (2021) in Oman, where they also concluded that an increase in the number of the elderly is associated with an increased rate of disease incidence.

Páez et al. (2020) investigated the influence of environmental (meteorological) factors, coupled with economic and demographic control variables, on the progression of the incidence of COVID-19 in the coterminous provinces in Spain, demonstrating that population density and percentage of older adults displayed strong associations with incidence of COVID-19. Singh & Adhikari (2020) compared the age and contact structures of the populations of India, China and Italy, emphasizing the effect these have on the spread of an infectious disease, particularly where the prevalence of three-generations households is higher as is the case of Portugal Albuquerque (2011).

Ramírez-Aldana, Gómez-Verjan & Bello-Chavolla (2020) also developed a spatial statistical approach to describe how COVID-19 cases are spatially distributed in Iran and tried to socioeconomic and climatic predictors for the number of cases, and concluded that the COVID-19 spread was spatially correlated, and that provinces with older population structures were the most susceptible to present a higher number of COVID-19 cases.

Cont, Kotlicki & Xu (2020) also observe significant differences in epidemic dynamics across regions in England, with higher fatality and contagion levels between regions, pointing to the importance of integrating demographic and geographical heterogeneity when modelling the impact of COVID-19, noting the particular relevance of elderly populations, where the percentage of symptomatic carriers is higher, which causes higher rates of infection, as supported by previous epidemiological evidence (Sayampanathan et al., 2020).

Considering the presented studies, we will consider municipal perimeter, the ageing index and the population density as explicative dimensions. As we will detail in the respective section, the municipal perimeter suggests that larger municipalities (in terms of area) tend to have more chances of exchanging people's movements with the surrounding spaces. Additionally, a higher population density also enhances these opportunities. Finally, a higher value of the ageing index is associated with a greater proportion of old people which has been found as correlated with higher exposures to the risk of COVID-19 contagion (Ramírez-Aldana et al., 2020; You, Wu, & Guo, 2020; Zhang & Schwartz, 2020).

Although data regarding positive cases used in this study refers only to the first of the three waves that already occurred in Portugal, the succeeding waves, replicated the spatial dynamics pattern marked by contagion of contiguity municipalities of the initial outbreaks.

The remaining sections of this work are as follows. Section 3 discusses the materials, data and methods used for this research. It also discusses the achieved results. Section 4 concludes the work.

2. Materials and methods

2.1. Study area

The study area is Portugal mainland, a coastal nation in southwestern Europe, located at the western end of the Iberian Peninsula, bordering Spain on its northern and eastern frontiers, with a

total of 92,225 km², ranging from 42°09'N to 36°57'N latitude and 6°11'W to 9°30'W longitude. It enjoys a Mediterranean temperate climate, with an annual average temperature ranging from 7,5°C in January to 22°C in August and an annual average precipitation ranging between 2 mm in the driest month (July) and 109 mm in the wettest month (December), the south being warmer and drier than the north.

It is comprised of 278 municipalities, with an estimated total population of almost 9.78 million in 2018, with a settlement pattern heavily marked by litoralization and metropolization processes (Cavaco, 2016) in a bipolar territorial structure anchored in the two metropolitan areas which together account for almost 47% of the total population, namely Lisbon with 2.85 million and the national capital located in the south, and Porto with 1,72 million in the north.

2.2. Dataset

The dataset for this study is composed of epidemic data and auxiliary data. The epidemic data mainly includes the total number of daily confirmed cases collected from the situation reports published by the Directorate-General of Health (www.covid19.min-saude.pt) from March 3, 2020, but containing information on confirmed cases broken down by municipality only from March 23, 2020.

The auxiliary data for each municipality includes the total population, population density and the ageing index (the ratio of the number of persons aged 65 and over to the number of young people aged 14 or younger), collected from Portugal annual estimates of the resident population in 2018 (www.ine.pt), total municipal area as well as length of municipal perimeter/border which are given by the Directorate-General of Territorial Development (<http://www.dgterritorio.pt>).

2.3. Methods

This research aims to discuss the spatial-temporal patterns of the COVID-19 epidemic from spatial patterns of the point prevalence (number of confirmed cases divided by resident population) per municipality and per day. Epidemic hotspots can be clustered (spatial clusters) or exist individually (spatial outliers) and they are identified using spatial association measures. The most appropriate of these measures is Local Moran's I index (Anselin, 1995; Le Gallo & Ertur, 2003). Local Moran's I is a global index of spatial association measures that assess the clustering of spatial data at the local level (using local clusters) or globally (using all the available data), used to explore the spatial structure of infectious diseases. For example, spatial analysis has been conducted for spatial epidemic dynamics of the COVID-19 outbreak in China (Kang, Choi, Kim & Choi, 2020). Given the heterogeneity in spatial and epidemic data, autocorrelation analysis tends to show significant local variations, so local Moran's I (Anselin, 1995) is more suitable to explore local autocorrelation or dependency of spatial epidemic data. Local Moran's I at spatial location i can be expressed as:

$$I_i = \frac{z_i - \bar{z}}{\sigma^2} \sum_{j=1, j \neq i}^n [\omega_{ij} (z_j - \bar{z})] \quad (1)$$

where z_i is the value of the point prevalence z at location i ; \bar{z} is the average value of z with the sample number of n ; z_j is the value of the prevalence rate z at all the other locations (where $j \neq i$); σ^2 is the variance of point prevalence; and w_{ij} is the spatial weight matrix, which defined spatial interaction throughout the study regions. In general, $\omega_{ij} = 1$ if location i and location j were neighbouring (shared a common boundary); otherwise, $\omega_{ij} = 0$. In this study, spatial contiguity was assessed as the first order queen's contiguity that defines spatial neighbours as those areas with shared borders and vertexes. Significance was tested by comparison to a reference distribution obtained by random permutations (Anselin, 1995). This analysis used 999 permutations to determine the differences among the spatial units.

A high positive Local Moran's I value implies that the location under study has similarly high or low values as its neighbours, thus the locations are spatial clusters. Spatial clusters include high-high clusters (high values in a high value neighbourhood) and low-low clusters (low values in a low value neighbourhood). In this context, low-low clusters are “low epidemic spots”, while high-high spatial clusters can be regarded as “regional epidemic hotspots”.

A high negative Local Moran's I value means that the location under study is a spatial outlier. Spatial outliers are those values that are obviously different from the values of their surrounding locations. Spatial outliers include high-low (a high value in a low value neighbourhood) and low-high (a low value in a high value neighbourhood) outliers. In epidemics, high-low spatial outliers can be regarded as isolated “individual hotspots”.

Spatial autocorrelation analysis was performed in GeoDa V.1.14.0.4 software (Anselin & Rey, 2014) and parametric regressions by StataIC 16.1.

Figures 1-3 show the different clusters observed throughout the period (Figure 1) or in the first date (Figure 2) and in the last recorded day (Figure 3). Figure 2 panels are identified as Moran's I scatterplot. The Moran scatterplots show the relationship between a specific variable (z) and its spatial lag (Wz) and classifies this statistically significant relationship for each municipality into one of four possible quadrants. Let us clarify this statistical significance regarding to the statistical significance of the estimated z and Wz computed for each municipality.

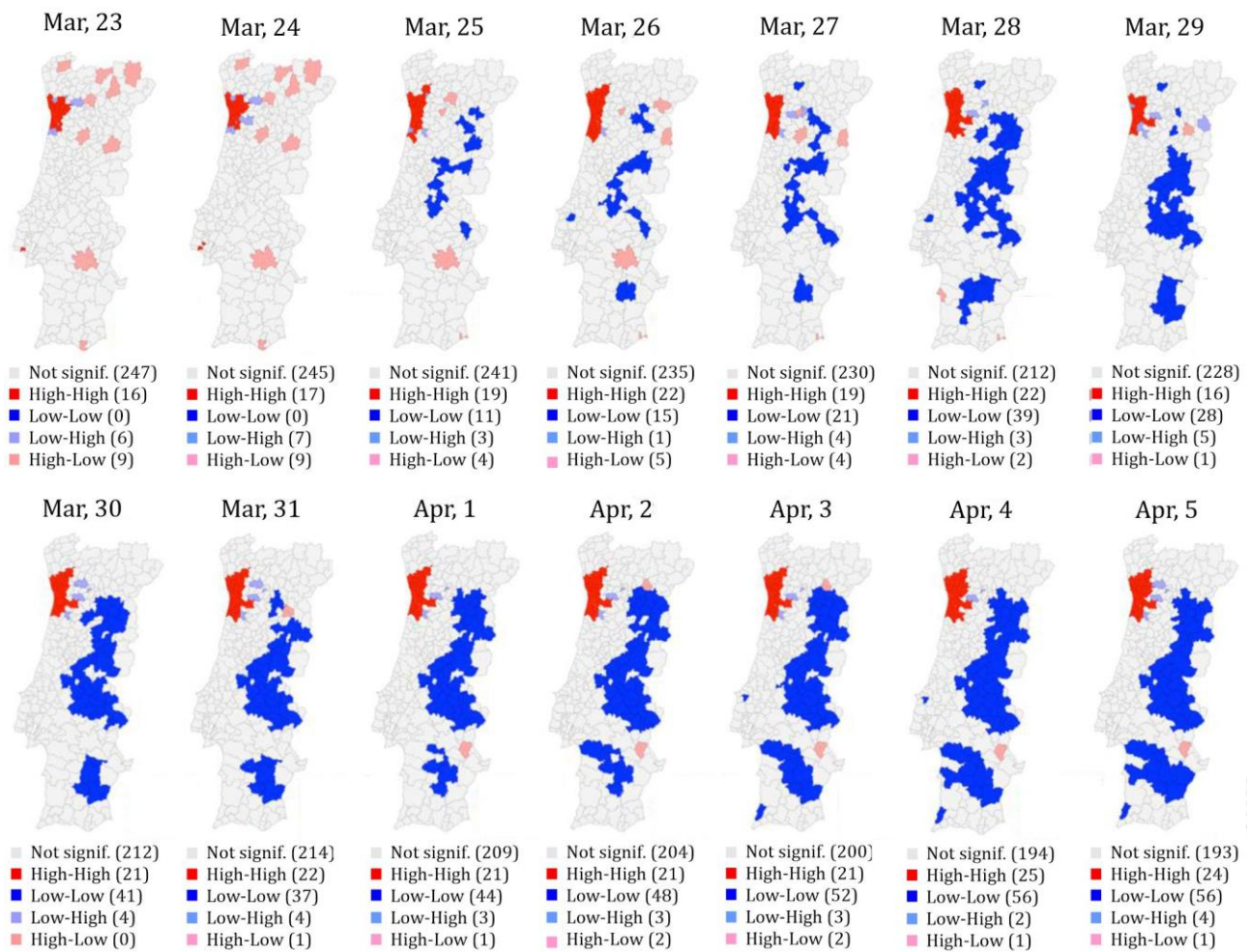


Figure 1. Clusters in Portugal following Moran's I analysis (March, 23 - April, 05). Source: own elaboration.

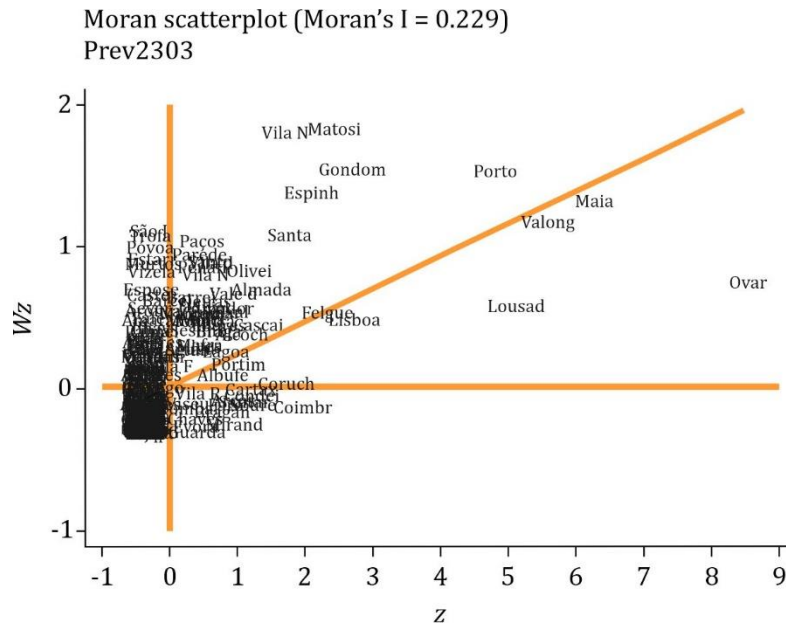


Figure 2. Moran scatterplot (March 23) – Distribution of municipalities considering point prevalence (z) and their spatial lag (Wz). Source: own elaboration.

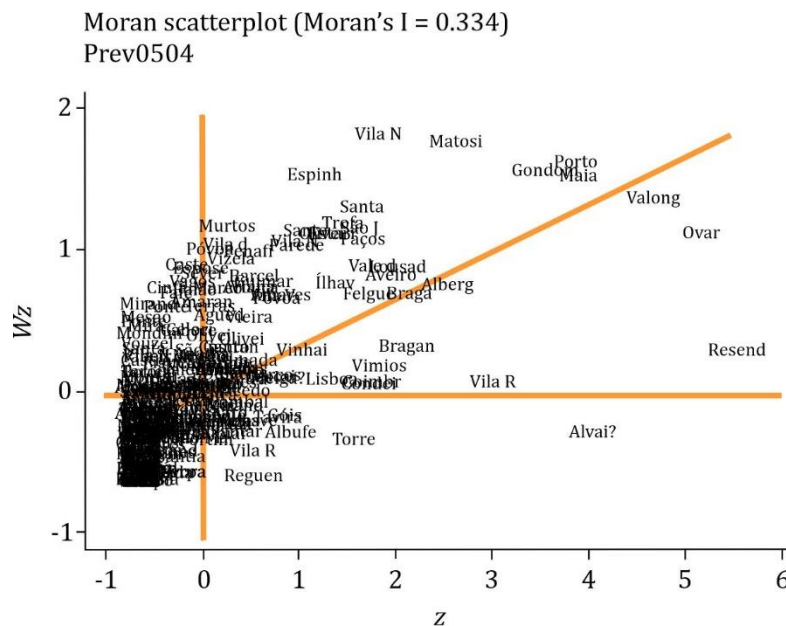


Figure 3. Moran scatterplot (April 05) – Distribution of municipalities considering point prevalence (z) and their spatial lag (Wz). Source: own elaboration.

Therefore, following Figure 1 (at each reported day), our z-variable is point prevalence (observed for each municipality) and Wz is a constructed variable giving the point prevalence of the surrounding/contiguous neighboring municipalities. The concentration of municipalities in the low-low cluster is evident (negative values for both z and Wz in Figures 2 and 3). The cases of high-low

(positive values for z and negative for Wz) are important because they relate to municipalities with social or public institutions which generated local epidemic foci. The cases of low-high are also worth analysing because these are municipalities with smaller values for prevalence ratios although these places are surrounded by spaces with significantly higher prevalence ratios. Figure 1 thus shows a concentration of cases in municipalities with higher incidence rate values; these cases have been concentrated in the north of the country. This situation led to the concentration of cases marked in red (high-high) in the north of Portugal in the observed period. These cases were identified between the urban areas around Braga (more to the north) and Aveiro (more to the south), which are generically referred to as areas of the north coast of Portugal, characterized by higher population densities. Reasons were put forward in reports in the daily press or by the National Statistics Institute itself (Instituto Nacional de Estatística [INE], 2020; Neves, 2020). One of the reasons for this distribution was the industrial concentration in this area that limited the telework decreed for the country (or work at home) on 19th March and driving many industrial employees to maintain the rhythms of work in places where there was no adequate prophylactic distance. In contrast, Figure 1 also shows the presence of municipalities identified in blue associated with the least affected areas, being concentrated in the space identified as 'interior of Portugal' (east of 8° W); these municipalities were also characterized by being surrounded by other municipalities with low incidence rates.

2.4. Parametric regressions

After the previous steps, which are the first when analysing spatial data, we intend to include exogenous dimensions for explaining the level of contagion observed throughout the Portuguese regions. For this purpose, several authors (Anselin, 1992; Anselin & Hudak, 1992) suggest the use of parametric regressions. Therefore, we explored spatial error models and spatial lag models. Starting from the standard ordinary least squares (OLS) model

$$Y_i = X_i \beta + e_i$$

the spatial error model uses the following identification for the error term

$$e_i = \lambda W_i e_i + u_i$$

while the spatial lag model recurs to a different identification

$$e_i = \rho W_i w_j + u_i$$

In simple terms, a spatial error model highlights the relevance of the observed residuals in the OLS estimation from neighbouring spaces while a spatial lag model highlights the relevance of the observations of the dependent variable in the neighbouring spaces. For our cases, the spatial lag model will try to explain the distribution of the prevalence variable by adding the prevalence observed in the surrounding area; the spatial error model will try to explain the distribution of the prevalence variable throughout Portuguese municipalities by adding the statistical residuals observed in the surrounding area (usually related to omitted explicative variables). As Anselin (1992) claims, spatial errors indicate a violation of the traditional OLS assumption of uncorrelated error terms. In the case of spatial errors, this suggests the omission of relevant and spatially correlated covariates/exogenous variables. In the case of spatial lags models, the statistical significance of spatial lags indicates the existence of diffusion processes, meaning that events in one place increase the likelihood of similar events in neighboring places.

As Anselin (1992) observed, a further step is to try to test hypotheses able to explain the found spatial autocorrelation. In our case, as we want to test some hypotheses regarding the spatial distribution of COVID-19 infectious cases in Portuguese municipalities between 23/03/2020 and 05/04/2020, we need to recur to models like the spatial error model or to the spatial lag model (Anselin & Hudak, 1992). We will include the population density as an explicable variable, following Li, Richmond & Roehner (2018), a relevant proxy for the contagious propensity across the spaces (Chandra, Kassens-Noor, Kuljanin & Vertalka, 2013; Holko, Mędreń, Pastuszek & Phusavat, 2016; Hu, Nigmatulina & Erckhoff, 2013; Kang et al., 2020), the ageing index (because older people have been identified as the most infectable cohort in this first wave (Davies et al., 2020; Gao et al., 2020), and the perimeter of each municipality (as a proxy for the likelihood of importing infected cases from neighbouring areas, which follows Lima et al. (2012)). The descriptive stats for these variables are available on request. Considering other hypotheses tested for other countries, we also suggest for future works the inclusion of other explicative variables such as the industrial concentration, the share of industrial production on local income, the average number of people per household, the share of students in the population, the number of hospitals and/or care homes in the municipality, the number of people daily transported by public vehicles, or the amount of CO₂ emissions recorded at each place (Gao et al., 2020; Holko et al., 2016).

For each day in the observed period, our diagnoses observed three tests for spatial error dependence (I-Moran, simple Lagrange multiplier, LM, of λ , and robust Lagrange multiplier of λ) and two tests for spatial lag dependence (simple Lagrange multiplier of ρ and robust Lagrange multiplier of ρ). Independent of the situation observed on the various days, all these tests provided significant values (p -value lower than 0.100; full results available on request). YLM tests are relevant for choosing an appropriate spatial regression model. Significance of LM(λ) points to a spatial error model, while significance of LM(ρ) points to a spatial lag model. The Lagrange multiplier test for spatial dependence (LM error test) is based on the estimation of the regression model with spatially autocorrelated errors versus an Ordinary Least Squares estimation without spatially autocorrelated errors. In the case of statistical significance of the respective parameter (λ), OLS estimation should be abandoned in favour of a spatial error model. Spatial dependence in regression models may not be properly described in the error. Instead, it may be accounted for by entering a spatial lag Wy in the endogenous variable Y . For assessing the relevance of a spatial lag (Wy) for explaining Y (against the null hypothesis of an OLS estimation) we shall use a LM test for the spatial lag (ρ). Following the empirical literature on the topic (Anselin, 1995; Anselin & Hudak, 1992), it is relevant to use these tests in their robust specification for avoiding endogenous influences between the spatial parameters ρ and λ .

Given the stability of our explicative variables throughout the days (actually, these variables are computed annually by the official National Statistics Institute, we only exhibit the spatial error model and the spatial lag model for our initial and last dates observed (the full range of estimations are available on request).

The results presented in Table 1 prove that both estimated λ and ρ are significant and positive, confirming the relevance of our models to explain the spatial dynamics of COVID-19 prevalence in Portuguese municipalities. Providing a clear interpretation, we can state that increasing values of the prevalence in neighbouring Portuguese municipalities had led to increases in the prevalence observed in a given Portuguese municipality. Additionally, the residuals observed for a certain municipality by our error model (the spatial error model, preferable if considering the Wald test, the Lratio test and the Lmultiplier test) were also spatially correlated with its neighbouring municipalities' residuals, proving the spatial interlinkages in the observed process. We also observed how relevant certain exogenous dimensions are important for explaining the observed dynamics. We conclude that the statistical significance of the population density and of the municipal perimeter explain higher prevalence.

Following the Elhorst taxonomy upon Sarrias (2020) –Figure 4–, we have also tested the interest on running spatial Durbin models, in which spatially lagged explanatory variables also exert significant effects on the dependent variable. However, considering the current set of explanatory variables and

through the robust version of the LR tests, we were not able to reject the null hypothesis of a non-significant γ ; therefore, we opted for not exhibiting the estimation of the spatial Durbin models here, given the expected statistically non-significance of the estimated coefficients for the spatially lagged explanatory variables (p -value Day1 = 0.434; p -value Day14 = 0.368).

Table 1. Spatial error model and spatial lag model results

Variables	Spatial error model		Spatial lag model	
	Day 1 (March 23, 2020)	Day 14 (April 5, 2020)	Day 1 (March 23, 2020)	Day 14 (April 5, 2020)
Municipal perimeter	0.0023 (0.0015)	0.0084 (0.0085)	0.0021* (0.0013)	0.0113 (0.0076)
Ageing index	-0.0008 (0.0008)	-0.0082* (0.0044)	-0.0002 (0.0006)	-0.0034 (0.0034)
Population density	0.0003*** (0.0001)	0.0017*** (0.0006)	0.0003*** (0.0001)	0.0013*** (0.0004)
Constant	0.3519 (0.3277)	5.7251* (2.9945)	-0.3371867 (0.2240)	-1.4229 (1.3552)
λ / ρ	1.21714*** (0.1419)	1.1180*** (0.0948)	1.1381*** (0.1535)	1.1087*** (0.1007)
Log-Likelihood	-379.67	-867.96	-379.06	-868.03
Observations	278	278	278	278
Diagnostics for spatial dependence				
Wald test	73.523	139.207	54.981	121.180
Lratio test	48.551	99.988	49.770	99.864
Lmultiplier test	44.334	192.213	47.776	175.550

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

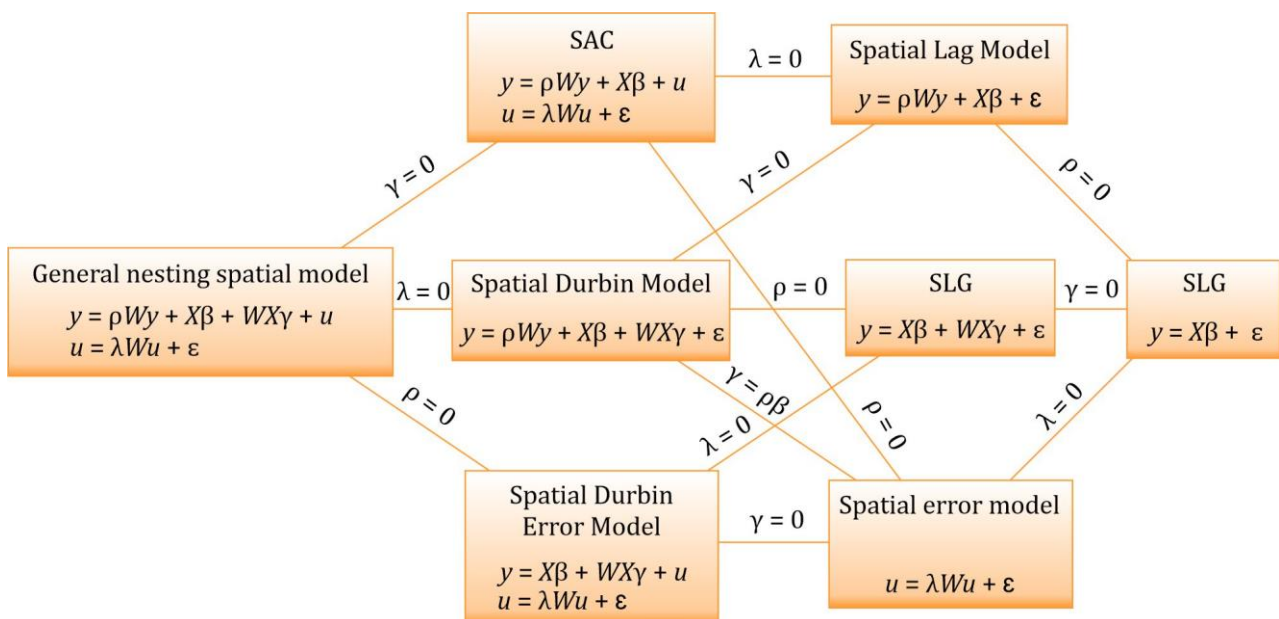


Figure 4. Elhorst taxonomy upon Sarrias (2020).

3. Conclusions and further work

The pandemic known as COVID-19 has so far experienced two waves in most countries and in some the presence of a third wave is under observation. The first wave started, for most countries, between January 2020 and the end of March 2020. This work is the first study on spatial autocorrelation patterns, at the municipal level, observed in the first wave, in Portugal.

The main conclusions point to municipal contiguity as a determining factor for contagion. Thus, municipalities surrounded by other municipalities with high rates of infection ended up being infected with high incidence values. This exercise demonstrates the need for mobility limitations in the future in order to reduce the rates of contagion of pandemics spread by humans.

These pioneering results enlarge the current discussion with a need to generate heterogeneous national health policies considering the significant variations within each country. Therefore, it is important for a more effective direction of the current fight against COVID-19 to consider the spatial heterogeneity here highlighted.

Further models –namely spatial panel data models– shall be considered in more extended empirical efforts in this domain. We also intend to enlarge the set of observed variables as soon as available data allow it. Namely, we consider it relevant to recur to the number of tests done at each municipality, which would allow an extension of these foci, generating a highly correlated control variable with the number of cases, but would also drive the analysis to the heterogeneous pattern of test distribution in the different Portuguese municipalities throughout the pandemic phases. We also consider the insertion of additional explicative variables relevant –as soon as these data become available: we are specifically referring to data like commuting or tourist interest which would account for the flow of people across the municipal borders to capture the probability of importing the virus.

As mentioned, we intend, in future work to extend the range of explanatory variables that we use here. We have already referred to variables like the industrial concentration, the share of industrial production on local income, the average number of people per household, the share of students in the population, the number of hospitals and or care homes in the municipality, the number of people transported daily by public vehicles, or the amount of CO₂ emissions recorded at each place (Gao et al., 2020; Holko et al., 2016).

We also intend to test this extended set of variables during the weeks of the different pandemic waves observed in Portugal. We intend to do so to the extent that, depending on the various pandemic control instruments used, it is unlikely that there will be a limited number of the same variables always responsible for the spatial contagion observed between neighboring municipalities in the different waves. Thus, we postulate that different variables will have a different capacity for explaining spatial autocorrelation in different epidemic waves.

References

- Anselin, L. (1992). Spatial data analysis with GIS: An introduction to application in the social sciences. *Technical Report 92-10*. Santa Barbara, CA: National Center for Geographic Information and Analysis. Retrieved from: <http://ncgia.ucsb.edu/technical-reports/PDF/92-10.pdf>
- Anselin, L. (1995). Local Indicators of Spatial Association-LISA. *Geographical Analysis*, 27(2), 93-115. DOI: <https://doi.org/10.1111/j.1538-4632.1995.tb00338.x>
- Anselin, L., & Hudak, S. (1992). Spatial econometrics in practice. A review of software options. *Regional Science and Urban Economics*, 22(3), 509-536. DOI: [https://doi.org/10.1016/0166-0462\(92\)90042-Y](https://doi.org/10.1016/0166-0462(92)90042-Y)
- Anselin, L., & Rey, S. J. (2014). *Modern spatial econometrics in practice: A guide to GeoDa, GeoDaSpace and PySAL*. Chicago, IL: GeoDa Press.
- Cavaco, C. (Coord.). (2016). *Habitat III – National Report Portugal*. Lisboa, Portugal: Direção Geral do Território. Retrieved from: http://habitat3.org/wp-content/uploads/PT_UN-HabitatIII_NationalReport_2016_08_04_EN.pdf

- Chandra, S., Kassens-Noor, E., Kuljanin, G., & Vertalka, J. (2013). A geographic analysis of population density thresholds in the influenza pandemic of 1918-19. *International Journal of Health Geographics*, 12(1), 9. DOI: <https://doi.org/10.1186/1476-072X-12-9>
- Cont, R., Kotlicki, A., & Xu, R. (2020). Modelling COVID-19 contagion: Risk assessment and targeted mitigation policies. *MedRxiv*. DOI: <https://doi.org/10.1101/2020.08.26.20182477>
- Davies, N. G., Klepac, P., Liu, Y., Prem, K., Jit, M., & CMMID COVID-19 working group, Eggo, R. M. (2020). Age-dependent effects in the transmission and control of COVID-19 epidemics. *MedRxiv*. DOI: <https://doi.org/10.1101/2020.03.24.20043018>
- Elhorst, J. P. (2014). *Spatial econometrics. From cross-sectional data to spatial panels. Springer Briefs in Regional Science*. Berlin, Germany: Springer.
- Gao, Q., Hu, Y., Dai, Z., Xiao, F., Wang, J., & Wu, J. (2020). The epidemiological characteristics of 2019 novel coronavirus diseases (COVID-19) in Jingmen, China. *SSRN Electronic Journal*, 2(8), 113-122. DOI: <https://doi.org/10.2139/ssrn.3548755>
- Guliyev, H. (2020). Determining the spatial effects of COVID-19 using the spatial panel data model. *Spatial Statistics*, 38, 100443. DOI: <https://doi.org/10.1016/j.spasta.2020.100443>
- Holko, A., Mędrek, M., Pastuszek, Z., & Phusavat, K. (2016). Epidemiological modeling with a population density map-based cellular automata simulation system. *Expert Systems with Applications*, 48, 1-8. DOI: <https://doi.org/10.1016/j.eswa.2015.08.018>
- Hu, H., Nigmatulina, K., & Eckhoff, P. (2013). The scaling of contact rates with population density for the infectious disease models. *Mathematical Biosciences*, 244(2), 125-134. DOI: <https://doi.org/10.1016/j.mbs.2013.04.013>
- INE. (2020). *COVID-19: uma leitura do contexto demográfico e da expressão territorial da pandemia - Dados até 16 de dezembro*. Lisboa, Portugal: Instituto Nacional de Estatística. Retrieved from: https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_destaques&DESTAQUESdest_boui=470299933&DESTAQUESstema=55481&DESTAQUESmodo=2
- Kang, D., Choi, H., Kim, J.-H., & Choi, J. (2020). Spatial epidemic dynamics of the COVID-19 outbreak in China. *International Journal of Infectious Diseases*. DOI: <https://doi.org/10.1016/j.ijid.2020.03.076>
- Krisztin, T., Piribauer, P., & Wögerer, M. (2020). The spatial econometrics of the coronavirus pandemic. *Letters in Spatial and Resource Sciences*, 13, 209-218. DOI: <https://doi.org/10.1007/s12076-020-00254-1>
- Le Gallo, J., & Ertur, C. (2003). Exploratory spatial data analysis of the distribution of regional per capita GDP in Europe, 1980-1995. *Papers in Regional Science*, 82, 175-201. Retrieved from: <https://link.springer.com/article/10.1007/s101100300145>
- Li, R., Richmond, P., & Roehner, B. M. (2018). Effect of population density on epidemics. *Physica A: Statistical Mechanics and Its Applications*, 510, 713-724. DOI: <https://doi.org/10.1016/j.physa.2018.07.025>
- Lima, I. D., Queiroz, J. W., Lacerda, H. G., Queiroz, P. V. S., Pontes, N. N., Barbosa, J. D. A., Martins, D. R., Weirahter, J. L., Pearson, R. C., Wilson, M. E., & Jeronimo, S. M. B. (2012). Leishmania infantum chagasi in northeastern Brazil: Asymptomatic infection at the urban perimeter. *American Journal of Tropical Medicine and Hygiene*, 86(1), 99-107. DOI: <https://doi.org/10.4269/ajtmh.2012.10-0492>
- Mansour, S., Al Kindi, A., Al-Said, A., Al-Said, A., & Atkinson, P. (2021). Sociodemographic determinants of COVID-19 incidence rates in Oman: Geospatial modelling using multiscale geographically weighted regression (MGWR). *Sustainable Cities and Society*, 65, 102627. DOI: <https://doi.org/10.1016/j.scs.2020.102627>
- Mollalo, A., Vahedi, B., & Rivera, K. M. (2020). GIS-based spatial modeling of COVID-19 incidence rate in the continental United States. *Science of the Total Environment*, 728, 138884. DOI: <https://doi.org/10.1016/j.scitotenv.2020.138884>
- Neves, S. (14 Abril 2020). Covid-19: TVI retira reportagem em que dizia que Norte é mais afectado por ter população “menos educada”. *Público*. Retrieved from: <https://www.publico.pt/2020/04/14/sociedade/noticia/covid19-tvi-retira-reportagem-dizia-norte-afectado-populacao-menos-educada-1912220>
- Oliva Denis, R. D., & Aldrey Vázquez, J. A. (2018). Padróns de distribución territorial da poboación estranxeira en Galicia, 1997-2017. *Revista Galega de Economía*, 27(2), 49-60. DOI: <https://doi.org/10.15304/rge.27.2.5657>
- Páez, A., López, F. A., Menezes, T., Cavalcanti, R., & Pitta, M. G. R. (2020). A spatio-temporal analysis of the environmental correlates of COVID-19 incidence in Spain. *Geographical Analysis*, 1-25. DOI: <https://doi.org/10.1111/gean.12241>
- Ramírez-Aldana, R., Gómez-Verjan, J. C., & Bello-Chavolla, O. Y. (2020). Spatial analysis of COVID-19 spread in Iran: Insights into geographical and structural transmission determinants at a province level. *PloS Neglected Tropical Diseases*, 14(11). DOI: <https://doi.org/10.1371/journal.pntd.0008875>

- Sannigrahi, S., Pilla, F., Basu, B., Sarkar Basu, A., & Molter, A. (2020). Examining the association between socio-demographic composition and COVID-19 fatalities in the European region using spatial regression approach. *Sustainable Cities and Society*, 62, 102418. DOI: <https://doi.org/10.1016/j.scs.2020.102418>
- Sayampanathan, A. A., Heng, C. S., Pin, P. H., Pang, J., Leong, T. Y., & Lee, V. J. (2021). Infectivity of asymptomatic versus symptomatic COVID-19. *The Lancet*, 397(10269), 93-94.
DOI: [https://doi.org/10.1016/S0140-6736\(20\)32651-9](https://doi.org/10.1016/S0140-6736(20)32651-9)
- Sarrias, M. (2020). *Spatial models*. Talca, Chile: Universidad de Talca.
- Singh, R., & Adhikari, R. (2020). Age-structured impact of social distancing on the COVID-19 epidemic in India. *arXiv:2003.12055 [q-bio.PE]*. Retrieved from: <https://arxiv.org/pdf/2003.12055.pdf>
- Sun, F., Matthews, S. A., Yang, T.-C., & Hu, M.-H. (2020). A spatial analysis of the COVID-19 period prevalence in US counties through June 28, 2020: Where geography matters? *Annals of Epidemiology*, 52, 54-59.e1.
DOI: <https://doi.org/10.1016/j.annepidem.2020.07.014>
- You, H., Wu, X., & Guo, X. (2020). Distribution of COVID-19 morbidity rate in association with social and economic factors in Wuhan, China: Implications for urban development. *International Journal of Environmental Research and Public Health*, 17(10), 3417. DOI: <https://doi.org/10.3390/ijerph17103417>
- Zhang, C. H., & Schwartz, G. G. (2020). Spatial disparities in coronavirus incidence and mortality in the United States: An ecological analysis as of May 2020. *The Journal of Rural Health*, 36(3), 433-445.
DOI: <https://doi.org/10.1111/jrh.12476>