



UNIVERSIDADE D
COIMBRA

Nuno Filipe Pocinho Gonçalves

ANÁLISE ESPACIAL DAS DETERMINANTES DA PROPAGAÇÃO DA COVID-19 EM PORTUGAL

Trabalho Projeto em Economia, orientado pelo Professor Doutor António Manuel Antunes Rafael Amaro e pelo Professor Doutor Helder Miguel Correia Virtuoso Sebastião e apresentado à Faculdade de Economia da Universidade de Coimbra

julho de 2021



FEUC FACULDADE DE ECONOMIA
UNIVERSIDADE DE COIMBRA

ANÁLISE ESPACIAL DAS DETERMINANTES DA PROPAGAÇÃO DA COVID-19 EM PORTUGAL

Nuno Filipe Pocinho Gonçalves

Trabalho Projeto do Mestrado em Economia com especializações em Economia do Crescimento e das Políticas Estruturais e em Economia Financeira. Orientado pelo Professor Doutor António Manuel Antunes Rafael Amaro e pelo Professor Doutor Helder Miguel Correia Virtuoso Sebastião. Apresentado à Faculdade de Economia da Universidade de Coimbra para obtenção do grau de Mestre.

julho de 2021

Agradecimentos

A meus pais, sem as oportunidades que me proporcionaram nunca seria o que sou hoje. Por todo o apoio incondicional e acima de tudo por todo o amor.

Um obrigado aos meus orientadores, pela confiança e disponibilidade. Ao professor António Rafael Amaro pela discussão e debate desenvolvido durante o último ano e pelo que me ensinou durante o meu percurso académico. Ao professor Helder Sebastião por ter aceitado fazer parte deste projeto, pelas suas sugestões e ajuda. Sem a contribuição deles este trabalho não era possível.

Por último, um agradecimento aos meus avós, em especial aos que não conseguirão testemunhar a conclusão de mais uma etapa da minha vida, obrigado pela construção das memórias mais belas. A todas as amigas que construí, em especial Filipe Murta e Sara Paz.

A todos só poderei prometer que enquanto houver capacidade e força continuarei o meu caminho!

Resumo

Este trabalho tem como objetivo identificar os fatores que influenciaram a propagação da Covid-19 nos 278 municípios de Portugal continental, durante a primeira fase da pandemia, entre os meses de março e outubro de 2020. Foram analisadas 56 variáveis, de diferentes dimensões, com destaque para as socioeconómicas, demográficas, infraestruturas de saúde ou étnico-raciais. A evolução da pandemia não é um fenómeno aleatório, a sua propagação é influenciada por múltiplos fatores, emergindo como principais determinantes as condições socioeconómicas, nomeadamente os níveis de desigualdade e pobreza, o contexto e os modelos de trabalho, a concentração demográfica e as condições de habitabilidade. Usámos modelos espaciais (Modelo de Desfasamento Espacial e Modelo Espacial dos Erros) e modelos locais (Regressão Ponderada Geograficamente e Regressão Multi-escalar Geograficamente Ponderada Geograficamente), complementando a análise com estatísticas espaciais (I de Moran), que provaram ser os instrumentos mais indicados para a compreensão e explicação da nossa problemática, nomeadamente por permitirem a interação entre as diferentes variáveis no espaço. Corroborando aquilo que a literatura mais recente vem defendendo, pudemos concluir com esta nossa investigação que a desigualdade de rendimentos, a pobreza, as condições de habitabilidade, as relações da população com o mercado de trabalho e a concentração urbana estiveram entre os principais determinantes da propagação do Covid-19 em Portugal na sua primeira fase.

Palavras-chave: COVID-19; Modelo Local; Modelo Espacial; Portugal, Determinantes Socioeconómicas

Classificação JEL: C21; I14; R2

Abstract

In this study, we aimed to identify the main determinants that influenced the spread of Covid-19 in the 278 municipalities of Continental Portugal, during the first phase of the disease outbreak, between the months of March and October 2020. A total of 56 variables of different dimensions were analyzed, with emphasis on socioeconomic, demographic, health infrastructures or ethnic-racial. The evolution of the pandemic is not a random phenomenon, the spread is influenced by multiple factors, emerging as the main determinants of socioeconomic conditions, namely the levels of inequality and poverty, the context and work models, demographic concentration and living conditions. Spatial models were used, namely Spatial Lag Model and Spatial Error Model, as well as local models, as Multiscale Geographically Weighted Regression and Geographically Weighted Regression, complementing the analysis with spatial statistics (Moran'I), proved to be the most indicated for the understanding and explanation of the problematic, allowing the interaction between the different variables in space. Corroborating what the most recent literature has been defending, we were able to conclude with our investigation that income inequality, poverty, living conditions, the population's relations with the labor market and urban concentration were among the main determinants of the propagation of Covid-19 in Portugal in its first phase.

Keywords: COVID-19; Local Model; Spatial Model; Portugal, Socioeconomic Determinants

JEL Classification: C21; I14; R2

Índice

Agradecimentos	iii
Resumo	iv
Abstract	v
1. Introdução	1
2. Revisão da literatura.....	2
3. Metodologia	6
3.1 Modelos e correlação espacial	6
3.1.1 Modelo introdutório.....	6
3.1.2 Modelos espaciais	7
3.1.3 Modelos locais.....	7
3.1.4 Correlação espacial	8
4. Dados e análise preliminar	10
4.1 Dados	10
4.2 Análise preliminar.....	12
5. Resultados.....	14
6. Discussão de resultados	19
6.1 Interpretação dos resultados.....	19
6.2 Robustez do modelo	23
7. Conclusão.....	24
Bibliografia	26
Anexos	31

1. Introdução

O primeiro caso oficialmente reportado de Covid-19, doença provocada pelo vírus Sars-Cov-2, foi registado na cidade de Wuhan, da província de Hubei, na China. A Organização Mundial de Saúde, em 20 de janeiro de 2020, classificou o surto como de Emergência de Saúde Pública Internacional e, dois meses depois, como pandemia. Desde o início ficou reconhecido que a propagação do vírus não teria fronteiras e que, no atual contexto de globalização, a velocidade de transmissão e de intensidade da doença seriam imparáveis.

Os primeiros casos confirmados de Covid-19 em Portugal remontam a 2 de março de 2020. A partir desta data, os casos foram aumentando, levando o Presidente da República no dia 18 de março a declarar Estado de Emergência que durou até ao terceiro dia do mês de maio. Para travar o avanço da pandemia foram encerrados serviços não-essenciais, impostas restrições à circulação, suspensão do direito à greve e quando possível a obrigatoriedade do teletrabalho. Assistiu-se a uma brusca e significativa queda da atividade económica portuguesa e do emprego. As empresas que beneficiaram de apoios estatais foram impedidas de despedir os seus trabalhadores. No entanto, apesar de todas as medidas de proteção e do confinamento obrigatório, o número de casos Covid-19 continuou a crescer.

A pandemia Covid-19 não atingiu todas as pessoas da mesma maneira. Como em todas as pandemias, o impacto foi também diferenciado entre continentes, países e regiões. São muito diversos os fatores que potenciam, como vários estudos epidemiológicos indicam, ou dificultam a propagação da doença. Por isso, as diferenças observadas não são um fenómeno aleatório. Com base nesta realidade, o principal objetivo deste trabalho consiste em identificar as determinantes que permitem explicar as diferenças da propagação da Covid-19 em Portugal continental, tendo como referência temporal os primeiros oito meses (entre 2 de março e 25 de outubro de 2020) da pandemia. Durante este período, procurámos analisar o impacto de 56 variáveis no número de casos acumulados de Covid-19 nos 278 municípios do continente, tendo em conta as dimensões económicas, sociais, demográficas, climáticas, condições de habitação, movimento pendulares da população, níveis de educação e de saúde.

No período em análise, foram contabilizados 121 133 casos confirmados de Covid-19, dos quais 69 956 recuperaram da doença e 2 343 terminaram em óbitos. Dos casos confirmados a maioria, 65 952, eram do sexo feminino. A população com idades compreendidas entre os 20 e 59 anos, era a que registava mais casos. Os óbitos observavam-se esmagadoramente na

população idosa com mais de 80 anos. O boletim diário, de dia 26 de outubro, indicava 48 834 casos ativos em Portugal, 1 672 pessoas em internamento e 240 internadas em unidades de cuidados intensivos. Relativamente ao dia anterior, somavam-se 1 341 novos casos, 27 óbitos e 108 pessoas em internamento. À data as regiões mais atingidas pela pandemia eram Lisboa e Norte, respetivamente com 53 412 e 51 932 casos confirmados. Os arquipélagos de Açores e Madeira eram as regiões menos afetadas, ambas com menos de 400 casos confirmados, sendo que a Madeira ainda não registava nenhum óbito.

Este documento prosseguirá com a revisão da literatura na secção 2, onde exponho os resultados e conclusões dos estudos mais recentes e relevantes acerca das determinantes da propagação da pandemia, também analiso os trabalhos que lidaram com a mesma metodologia usada neste estudo. Na secção 3 é explicada a metodologia, acerca dos modelos espaciais e locais, assim como das estatísticas espaciais. Na secção 4 são apresentados os dados e a análise preliminar aos mesmos. Os resultados são expostos na secção 5, a sua interpretação é dada na secção 6. Terminamos com a conclusão correspondente à secção 7.

2. Revisão da literatura

Apesar de o vírus se transmitir através de gotículas que contêm partículas virais libertadas por pessoas infetadas, que infetam quem estiver próximo, existem vários fatores que contribuem para explicar as disparidades observadas entre o número de infeções em várias zonas. Desde o início, como referiram Blundell et al. (2020), a capacidade de uma população para se proteger da doença é, em grande medida, influenciada e agravada pelas desigualdades registadas antes da pandemia. Ou seja, as diferenças na propagação e gravidade da pandemia a nível mundial evidenciavam a existência de condições prévias que as potenciavam e favoreciam. Alguns estudos (Gray et al., 2020) alertam para a existência de determinantes que influenciam a propagação do vírus, nomeadamente para fatores que causam divergências na qualidade da saúde da população: racismo, discriminação, pobreza, políticas sociais. Como consequência destas, temos determinantes de nível intermédio: instrução e formação, relações com o mercado de trabalho, condições habitacionais, estabilidade dos rendimentos, qualidade alimentar, confiança nas recomendações das entidades de saúde, acesso a cuidados de saúde, acesso a rastreamento e cura de Covid-19, acesso a máscaras e capacidade de manter distanciamento social. Por último, temos as determinantes que são, em grande medida, o resultado de todas as anteriores, com grande influência na mortalidade: doenças cardíacas, respiratórias, excesso de peso, diabetes, cancro, e disponibilidade hospitalar para tratar casos Covid-19.

Os fenómenos pandémicos estão também dependentes de características geográficas, nomeadamente a maior proximidade ou dispersão das populações e as relações de dependência entre elas. Literatura recente (Rahman et al., 2020; Guliyev, 2020; Kang et al., 2020; Sannigrahi et al., 2020; Ehlert, 2020; Y. Sun et al., 2021; Mansour et al., 2021; Amdaoud et al., 2021) tem procurado identificar as principais determinantes da maior ou menor propagação e gravidade da doença, com base em modelos espaciais e locais. Recorrendo também a estatísticas espaciais que permitem identificar e explicar os perfis e características das localidades afetadas pela Covid-19, bem como a evolução e determinantes da sua propagação e letalidade. A importância destes modelos fica bem demonstrada em estudos sobre a realidade dos Estados Unidos da América (EUA) (Coven & Gupta, 2020; Baum & Henry, 2020; Mollalo et al., 2020; Andersen et al., 2021; Maiti et al., 2021). Aqui, entre várias dimensões sociológicas que explicam uma maior propagação da doença, destaca-se a questão étnico-racial. A população afrodescendente e hispânica estadunidense é, entre outras coisas, desproporcionalmente afetada por doenças crónicas e, sobretudo, pelo desigual acesso a cuidados de saúde, tendo como consequência uma maior probabilidade de infeção grave (Tai et al., 2021). Estes grupos minoritários, socialmente mais desfavorecidos, estão também entre os mais contaminados pela doença pandémica (Yang et al., 2021). No caso dos EUA, é em zonas de cidades onde existe maior percentagem de população afrodescendente e hispânica que as taxas de incidência da doença são maiores, tal como as consequentes taxas de mortalidade (Coven & Gupta, 2020; F. Sun et al., 2020; Andersen et al., 2021; Khanijahani, 2021; Karmakar et al., 2021). Também o género tem impacto na propagação do vírus. Os condados mais afetados são aqueles onde a percentagem de mulheres negras é maior (Mollalo et al., 2020).

Estudos sobre Inglaterra também evidenciam o mesmo padrão: as zonas onde habita a população mais desfavorecida, como a população asiática, negra e desempregada imigrante, correspondem a zonas onde a mortalidade por Covid-19 é maior (Y. Sun et al., 2021).

As relações da população com o mercado de trabalho e o local onde vive desempenham um papel igualmente importante. As pessoas que vivem em zonas residenciais mais pobres são as mais afetadas; correspondendo à população que mais horas de trabalho presencial tem e, consequentemente, fazem parte dos grupos com trabalhos mais expostos ao risco de contágio (Coven & Gupta, 2020). As famílias com menos recursos são também aquelas que apresentam formação profissional inferior e que menos possibilidades têm de trabalhar a partir de casa, expondo-se desta forma mais ao risco de contágio (Blundell et al., 2020). De acordo com

Galasso (2020), os trabalhadores com menos qualificações, com empregos menos diferenciados, continuaram a trabalhar no local habitual de trabalho, expondo-se duplamente ao contágio: por um lado, são mais afetados devido a seu trabalho e à sua condição socioeconómica; por outro, vêm a sua situação agravada pelo facto de residirem em áreas pobres e povoadas por etnias minoritárias (Blundell et al., 2020; Adams-Prassl et al., 2020).

Os meios urbanos foram os mais afetados desde o início da pandemia (Andersen et al., 2021). A densidade populacional é, desde logo, um dos principais fatores para a transmissão do vírus em fases iniciais das pandemias (Gerritse, 2020). A cidade agrega vários fatores propensos à propagação do vírus. Desde logo, as pessoas que residem em áreas mais pobres e concentradas, de menores rendimentos e com piores condições de habitabilidade. Condicionismos agravados quando se acrescenta a tudo isto as dificuldades de acesso à informação e aos cuidados sanitários, aumentando exponencialmente as probabilidades de virem a testar positivo à Covid-19 (Borjas, 2020; Schmitt-Grohé et al., 2020). Para combater o avanço da pandemia numa cidade, o modelo de trabalho remoto provou ser a melhor solução, uma vez que as zonas onde maior parte da população tem um trabalho reproduzível remotamente estão associadas a menores taxas de incidência (Yang et al., 2021). Os grandes centros de atividade económica são locais onde as conexões com o resto do mundo são mais densas, e com maior atividade económica, sendo, por isso, os primeiros a terem mais casos de Covid-19 (Ascani et al., 2021).

Como era conjecturado, as determinantes socioeconómicas são das mais importantes, na explicação do avanço da pandemia. Os estudos que retratam a realidade estadunidense identificam as zonas de maior desigualdade de rendimento como aquelas onde se registam maiores taxas de incidência de Covid-19 (Mollalo et al., 2020; Karmakar et al., 2021). Também na Europa se registam as mesmas relações, as condições de pobreza potenciam a propagação do vírus (Sannigrahi et al., 2020). Aliás, sobre a realidade de alguns países da Europa do Sul verificou-se que o risco de infeção nas regiões mais pobres aumentou da primeira para a segunda vaga (Marí-Dell'olmo et al., 2021). Do mesmo modo, alguns autores encontraram uma correlação positiva entre o desemprego observado em algumas regiões com uma maior disseminação do vírus (Amdaoud et al., 2021). Em França, por exemplo, nas localidades onde existe maior desigualdade de rendimentos o vírus é mais letal (Ginsburgh et al., 2021). Já no Reino Unido, as regiões onde se observam maiores taxas de desemprego são também aquelas que evidenciam uma maior taxa de mortalidade (Y. Sun et al., 2021).

As determinantes demográficas são das mais estudadas para explicar a propagação e letalidade da Covid-19. Em Itália, Espanha, França e Reino Unido as regiões com uma percentagem de população masculina acima de 60 anos foi onde se registaram maior número de mortos (Sannigrahi et al., 2020). Nos países do Sul da Europa, o género e o grupo etário mais afetado no que toca a contaminações foi o das mulheres com menos de 64 anos (Marí-Dell’olmo et al., 2021). Vários estudos indicam também que à medida que a idade média da população aumenta, o número de infetados também (Ehlert, 2020; F. Sun et al., 2020; Baum & Henry, 2020). Na Europa, todos os estudos indicam que as maiores taxas mortalidade ficam associadas a territórios onde a percentagem da população mais velha é maior (Amdaoud et al., 2021).

No que diz respeito à condição de bem-estar da população, estudos revelam que a preexistência de doenças como a diabetes, hipertensão, cancro, doenças respiratórias, crônicas e cardiovasculares, determinam fortemente a taxa de morte e infeção por Covid-19 (Sannigrahi et al., 2020; Mansour et al. 2021). Ainda dentro da dimensão saúde, existem condições que podem influenciar a propagação do vírus: a existência ou não de infraestruturas de apoio aos sistemas de saúde; a percentagem de enfermeiros (Mollalo et al., 2020); o número de profissionais de saúde ou de cuidados à população sénior (Ehlert, 2020). Está comprovado que a incidência do vírus está positivamente relacionada com a proximidade de infraestruturas de saúde, já o combate à mortalidade está negativamente relacionado com a acessibilidade aos serviços de saúde (Y. Sun et al., 2021).

O clima, a mobilidade, os transportes utilizados, mostraram ser determinantes em diferentes estudos: a taxa de mortalidade, por exemplo, está negativamente relacionada com a humidade do ar (Y. Sun et al., 2021); a mobilidade da população e os fluxos de interação entre pessoas, correlacionados com o uso de transporte públicos, estão entre os principais fatores responsáveis pelas altas taxas de contaminação (Sá, 2020).

Os estudos das determinantes capazes de influenciar as taxas de disseminação do vírus e as taxas de morbilidade são muito diferentes entre eles. Procurámos, no entanto, com base em diferentes estudos, colocar em evidência como as condições económicas, sociais, de transporte e mobilidade, de saúde, de habitabilidade, de formação, demográficas e mesmo climatéricas acabam por influenciar a disseminação da Covid-19. Muitos dos ensaios compulsados utilizaram modelos espaciais e locais, outros estatísticas de análise espacial, adequados à realidade que pretendiam analisar. Tendo em conta os dados disponíveis, foi isso que também fizemos, adaptando à realidade portuguesa.

3. Metodologia

3.1 Modelos e correlação espacial

Neste estudo foram computados cinco modelos. Primeiro uma regressão multivariada através do método dos mínimos quadrados (MMQ), permitindo analisar a relação entre os casos acumulados de Covid-19 por município e as eventuais variáveis explicativas, sem entrar em consideração com interdependências espaciais. Dada a natureza dos fenômenos epidemiológicos é de esperar a existência de interdependências espaciais, se tal acontece os estimadores destes modelos, tendem a ser enviesados e inefficientes (Amdaoud et al., 2021). Pois, o número de casos acumulados observados num município, pode ser influenciado espacialmente de três maneiras diferentes: (1) O valor de y de um território pode impactar (ou estar relacionado) com o valor de y de um território vizinho, (2) os valores de uma determinante x de um território pode afetar (ou estar relacionado) com os valores de y de outro território, e (3) os resíduos de um território podem afetar os resíduos de outro território. Por estas razões os modelos mais indicados para o estudo da problemática são os modelos espaciais e locais.

Aquando da confirmação da interdependência espacial estimámos: um Modelo de Desfasamento Espacial (*Spatial Lag Model* - SLM), um Modelo Espacial dos Erros (*Spatial Error Model* - SEM); uma Regressão Ponderada Geograficamente (*Geographically Weighted Regression* - GWR) e uma Regressão Multi-escalar Ponderada Geograficamente (*Multiscale Geographically Weighted Regression* - MGWR). Os dois primeiros modelos mencionados são denominados de modelos espaciais, os últimos dois são modelos locais.

3.1.1 Modelo introdutório

O MMQ permite estabelecer relações *cross section*, considerando por hipótese que não existem quaisquer relações espaciais entre a variável dependente e as variáveis explicativas (Oshan et al., 2019; Sannigrahi et al., 2020; Mollalo et al., 2020). A sua forma geral é a seguinte:

$$y_i = \beta_0 + \mathbf{B}'\mathbf{X}_i + \varepsilon_i, \text{ com } i = 1, 2, \dots, n \quad (1)$$

Onde y_i representa os casos acumulados por Covid-19 por 100 mil residentes no município i ; β_0 é a constante; \mathbf{B} é o vetor coluna dos coeficientes da regressão, \mathbf{X}_i é o vetor coluna das variáveis explicativas no município i , e ε_i é o termo de erro. O MMQ permite determinar o vetor \mathbf{B} que minimiza a soma dos erros quadrados (Anselin & Arribas-Bel, 2013; Sannigrahi et al., 2020; Mollalo et al., 2020). Na aplicação do MMQ assume-se que as variáveis explicativas não estão correlacionadas com os erros e que os erros são independentes e igualmente

distribuídos (i.i.d.), portanto exclui-se a possibilidade de qualquer interdependência produzida pela proximidade geográfica (Anselin & Arribas-Bel, 2013; Sannigrahi et al., 2020; Mollalo et al., 2020).

3.1.2 Modelos espaciais

Modelo de Desfasamento Espacial (SLM)

O SLM assume a dependência espacial entre a variável dependente e as variáveis explicativas, incorporando o vetor da variável dependente, \mathbf{y} , defasado “espacialmente” (Anselin, 2003):

$$y_i = \beta_0 + \mathbf{B}'\mathbf{X}_i + \rho\mathbf{W}_i\mathbf{y} + \epsilon_i \quad (2)$$

Neste modelo, $\rho\mathbf{W}_i\mathbf{y}$ é o desfasamento espacial da variável dependente, onde ρ é o parâmetro espacial; $\mathbf{W}_i\mathbf{y}$ é o vetor da matriz espacial \mathbf{W} , que expressa a contiguidade entre o município i e os outros municípios (1 se for vizinho do município i e 0 se não for vizinho do município i), tal que $\mathbf{W}_i\mathbf{y} = \sum_{j=1}^n w_{ij}y_j$. Desta forma, o modelo incorpora a relação entre a variável dependente de um município e as variáveis explicativas e dependente dos municípios vizinhos (Anselin & Arribas-Bel, 2013). Se o parâmetro ρ for significativo então existe um processo de difusão espacial. Por esta razão os coeficientes não podem ser interpretados como efeitos marginais, uma vez que a variável dependente é também influenciada pelas variáveis de concelhos vizinhos (Mollalo et al., 2020).

Modelo Espacial dos Erros (SEM)

Este modelo assume a dependência espacial a partir do termo de erro da Equação (1), em duas partes, isto é: $\epsilon_i = \lambda\mathbf{W}_i\boldsymbol{\epsilon} + \xi_i$ (Anselin, 2003; Mollalo et al., 2020; Sannigrahi et al., 2020):

$$y_i = \beta_0 + \mathbf{B}'\mathbf{X}_i + \lambda\mathbf{W}_i\boldsymbol{\epsilon} + \xi_i \quad (3)$$

Neste modelo as relações espaciais são incorporadas através do vetor dos termos de erro $\boldsymbol{\epsilon}$, onde $\lambda\mathbf{W}_i\boldsymbol{\epsilon}$ é o termo de erro espacial, λ indica o nível de correlação entre as componentes e ξ_i é o termo de erro usualmente assumido como sendo i.i.d.

3.1.3 Modelos locais

Regressão Ponderada Geograficamente (GWR)

Os modelos espaciais anteriores assumem que a relação entre a variável dependente e as variáveis explicativas é constante e, portanto, válida globalmente. O modelo GWR permite que

os parâmetros variem no espaço (Brunsdon et al., 1996, 1998), permitindo uma flexibilização dos modelos anteriores:

$$y_i = \beta_{i0} + \mathbf{B}'_i \mathbf{X}_i + \varepsilon_i \quad (4)$$

Isto significa que o modelo GWR estima os parâmetros locais para cada município individualmente (Fotheringham & Oshan, 2016):

$$\hat{\mathbf{B}}_i^* = [\hat{\beta}_{i0} \hat{\mathbf{B}}'_i] = (\mathbf{X}'\mathbf{W}_i^*\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}_i^*\mathbf{y} \quad (5)$$

Agora \mathbf{W}_i^* é uma matriz dos pesos espaciais, tendo em conta a distância entre o município i e os restantes municípios (Brunsdon et al., 1998; Fotheringham & Oshan, 2016). Para calcular \mathbf{W}_i^* normalmente são escolhidas funções de kernel bi-quadrado ou gaussiana. A banda do kernel afeta a abrangência dos municípios utilizados no esquema de ponderação local e pode ser definida para os concelhos mais próximos ou distância entre eles (Mollalo et al., 2020).

Regressão Multi-escalar Ponderada Geograficamente (MGWR)

Enquanto que no GWR as escalas de banda larga são homogêneas, o MGWR é uma extensão do GWR, permitindo que a escala dos efeitos das variáveis explicativas varie (Fotheringham et al., 2017; Mollalo et al., 2020; Mansour et al., 2021).

$$y_i = \beta_{i0}(bm_0) + \mathbf{B}_i(\mathbf{bm})\mathbf{X}_i + \varepsilon_i \quad (6)$$

\mathbf{bm} é o vetor de bandas de *kernel* utilizadas para calibrar a relação entre a variável dependente e cada uma das variáveis explicativas (Fotheringham et al., 2017). O modelo tem assim um conjunto de bandas largas, uma para cada uma das variáveis explicativas. As diferentes bandas, as quais são endogenamente estimadas pelo modelo, representam diferenças nas escalas espaciais, permitindo capturar melhor a heterogeneidade espacial (Fotheringham et al., 2017; Mollalo et al., 2020).

3.1.4 Correlação espacial

A correlação espacial é aqui analisada através da estatística I de Moran (Moran, 1948). Existem várias versões do I de Moran tendo em conta se a estatística é global ou local e se considera uma ou duas variáveis. As estatísticas locais também são por vezes designadas pelo acrónimo LISA (*Local Indicator of Spatial Association*).

Seja y a variável de interesse e $\tilde{y}_i = y_i - \bar{y}$ e $\tilde{y}_j = y_j - \bar{y}$, $\forall i, j = 1, 2, \dots, n$, com $i \neq j$ (por exemplo, y_i é a observação da variável y , número de casos acumulados de Covid-19 por 100

mil habitantes no município i) onde \bar{y} é a média da variável y . O I de Moran global (com uma variável) é dado por:

$$I = \frac{\sum_i \sum_j w_{ij} \tilde{y}_i \tilde{y}_j}{\sum_i \tilde{y}_i^2} \quad (7)$$

Onde w_{ij} são os elementos da matriz espacial \mathbf{W} .

Na versão local do I de Moran, a estatística é calculada para cada município i :

$$I_i = \frac{\sum_j w_{ij} \tilde{y}_i \tilde{y}_j}{\sum_i \tilde{y}_i^2} = \frac{\tilde{y}_i}{\sum_i \tilde{y}_i^2} \sum_j w_{ij} \tilde{y}_j \quad (8)$$

O I de Moran também pode ser aplicado a duas variáveis, \tilde{y} e \tilde{x} , global ou localmente, permitindo assim observar o grau de correlação entre duas variáveis globalmente ou entre concelhos vizinhos. O I de Moran global com duas variáveis é calculado pela seguinte expressão:

$$I^B = \frac{\sum_i \sum_j w_{ij} \tilde{y}_i \tilde{x}_j}{\sum_i \tilde{y}_i^2} \quad (9)$$

Enquanto o I de Moran local com duas variáveis é dado por:

$$I_i^B = \frac{\tilde{y}_i}{\sum_i \tilde{y}_i^2} \sum_j w_{ij} \tilde{x}_j \quad (10)$$

Para realizar inferências estatísticas sobre o I de Moran global e local usualmente é obtida a sua distribuição através de uma simulação aleatória, considerando um dado número de permutações (neste estudo são utilizadas 999 permutações). A determinação do seu valor-p é feita situando o valor de I de Moran observado na distribuição simulada (Amdaoud et al., 2021).

A distribuição espacial dos I de Moran e I de Moran com duas variáveis locais permite identificar diferentes relações espaciais entre as variáveis pelo território e o *clustering* (Anselin, 1995). No presente estudo, os I de Moran locais são calculados com recurso ao software GeoDa e os resultados são apresentados graficamente através de um código de cores em mapas de Portugal continental. Basicamente, o código de cores é o seguinte: Vermelho-escuro (*high-high*) indica que o valor da variável identificada no município i em análise e nos municípios vizinhos estão acima da média nacional; vermelho-claro (*high-low*) indica que o valor do município i está acima da média enquanto os valores dos municípios vizinhos está abaixo da média; azul-escuro (*low-low*) indica que os valores do município i e dos municípios vizinhos

estão abaixo da média, e, finalmente, azul-claro (*low-high*) indica que os valores do município *i* está abaixo da média, mas os dos município vizinhos estão acima da média. Municípios sem cores indicam a inexistência de qualquer relação espacial significativa.

4. Dados e análise preliminar

4.1 Dados

O estudo tem em conta os casos acumulados por concelho registados entre o dia 2 de março e dia 25 de outubro de 2020¹. Para a implementação deste estudo foi recolhida informação sobre 56 variáveis que potencialmente podem ter influenciado a propagação da Covid-19 nos 278 municípios de Portugal continental (Tabela 1). Foram excluídos os municípios dos arquipélagos dos Açores e Madeira por estarem isolados do continente. A seleção das variáveis teve por base a sua utilização documentada nos artigos referidos na revisão da literatura e a sua disponibilidade em fontes estatísticas nacionais fiáveis.

Nem todas as variáveis têm disponíveis valores municipais, sendo nestes casos dado o valor da unidade territorial mais próxima, nomeadamente NUT III e NUT II. Em particular, as variáveis *emvn* e *emv65* correspondem à média das NUT III a que pertence cada concelho; enquanto as variáveis *riscopob*, *privm*, *privms*, *ipm* e *riscopobs* correspondem à média das respetivas NUT II. Para as variáveis *gini*, *p8010* e *p9010*, dos municípios de Alcoutim, Alvito, Barrancos, Monforte, Mourão e Penedono, abrangidos por confidencialidade estatística, também se utilizou os valores médios das respetivas NUT III. Também de notar que a variável *des*, *proxy* da taxa de desemprego (número de pessoas inscritas à procura de emprego sobre a população residente) foi construída através da agregação dos dados mensais recolhidos através do IEFP. Mais precisamente, estes dados foram somados e de seguida divididos pelo número de meses de março a outubro.

¹ Não é usado um arco temporal mais extenso visto que, a DGS deixou de publicar informação acerca do número de casos confirmados por concelho. Sendo o boletim de dia 26 de outubro o último com esta informação.

Tabela 1 – Caracterização das variáveis

ACRÓNIMO	DESCRIÇÃO DA VARIÁVEL	DATA	FONTE	GRUPO
<i>ca100m</i>	Número de casos acumulados por 100 mil habitantes	26/10/2020 ²	DGS	Variável dependente
<i>gini</i>	Coefficiente de Gini do rendimento bruto declarado por agregado fiscal			
<i>p8020</i>	Número de vezes que o rendimento do agregado fiscal situado no percentil 80/90 é superior ao rendimento do percentil 20/10, do rendimento bruto declarado do agregado fiscal	29/07/2020	INE	Desigualdade
<i>p9010</i>				
<i>dispghm</i>	Disparidade no ganho médio mensal entre profissões (%)	12/11/2020		
<i>pcompra</i>	Poder de compra (índice)	08/11/2019	INE	Rendimento médio
<i>rbrutopc</i>	Rendimento Bruto per capita	29/07/2020		
<i>riscpob</i>	População residente em risco de pobreza ou exclusão social (NUT II)(%)			
<i>privm</i>	Taxa de privação material (NUT II)(%)			
<i>privms</i>	Taxa de privação material severa (NUT II)(%)	18/12/2019	INE	Pobreza
<i>ipm</i>	Intensidade da privação material (NUT II)			
<i>riscpobts</i>	Taxa de risco de pobreza após transferências sociais (NUT II)(%)			
<i>psdoe</i>	População residente beneficiária do subsídio de doença (%)	13/11/2020		
<i>psdes</i>	População residente beneficiária do subsídio de desemprego (%)	11/11/2020	INE	Proteção Social
<i>prsi</i>	População residente beneficiária do rendimento social de inserção (%)	14/11/2020		
<i>pcid</i>	População residente que vive em cidades (%)	11/03/2019		
<i>purb</i>	População residente que vive em meio predominantemente urbano (%)			
<i>prur</i>	População residente que vive no meio predominantemente rural (%)	10/03/2017	INE	Urbano/Rural
<i>p2m</i>	População residente que vive em aglomerado com mais de 2 mil pessoas (%)	29/01/2016		
<i>des</i>	Pessoas inscritas (no IEFP) à procura de trabalho pela população residente (%)	31/12/2020	IEFP	Emprego
<i>denspop</i>	Densidade Populacional	17/06/2020	INE	Demografia
<i>trcol</i>	População residente que usa meio de transporte coletivo (%)	04/01/2017	INE	Movimentos pendulares
<i>tempm</i>	Temperatura média em grau Celsius	26/06/2020	INE	Clima
<i>dmf</i>	Dimensão média das famílias	26/05/2015		
<i>asup</i>	Proporção de alojamentos sobrelotados (%)	12/03/2013		
<i>bs</i>	Número de bairros sociais	18/07/2012	INE	Habitação
<i>aseb</i>	Proporção de alojamentos sem pelo menos uma infraestrutura básica (%)			
<i>anr</i>	Proporção de alojamentos que necessitam de reparações (%)	14/01/2013		
<i>anqr</i>	Proporção de alojamentos que necessitam de grandes reparações (%)	29/01/2016		
<i>ie</i>	Índice de envelhecimento- quociente entre o número de pessoas com 65 ou mais anos e o número de pessoas com idades compreendidas entre os 0 e os 14 anos	26/06/2015		
<i>emv65</i>	Esperança média de vida aos 65 anos	24/09/2020		
<i>emvn</i>	Esperança média de vida á nascença (NUTS III)	18/12/2019		
<i>idi</i>	Índice de dependência de idosos (%)	26/06/2015	INE	Estrutura etária
<i>ge1t; ge2t; ge3t; ge1h; ge2h; ge3h; ge1m; ge2m; ge3m; h; m</i>	População: t(masculina e feminina); h(masculina); m(feminina) - com idades compreendidas entre os 1 (0-29); 2 (30-64); 3 (+65) (%)	15/06/2020		
<i>se</i>	População residente sem escolaridade (%)			
<i>b1c</i>	População residente com ensino básico 1º ciclo (%)			
<i>b2c</i>	População residente com ensino básico 2º ciclo (%)	26/06/2015	INE	Educação
<i>b3c</i>	População residente com ensino básico 3º ciclo (%)			
<i>sec</i>	População residente com ensino secundário (%)			
<i>sup</i>	População residente com ensino superior (%)			
<i>med</i>	Número de médicos por 100 mil habitantes	17/08/2020	INE	Saúde
<i>enf</i>	Número de enfermeiros por 100 mil habitantes			
<i>ipss</i>	Número de Instituições Particulares de Solidariedade Social	02/03/2021	Seg. Social	
<i>br</i>	População residente proveniente do Brasil (%)			
<i>el</i>	População residente proveniente da Europa do Leste (Ucrânia, Roménia, Moldávia) (%)	18/12/2019	INE	População estrangeira
<i>af</i>	População residente proveniente de África (Angola, Cabo Verde, Guiné-Bissau e São Tomé e Príncipe) (%)			
<i>pe</i>	População residente estrangeira (%)	10/10/2013		

*Notas: Nesta tabela são apresentadas as variáveis recolhidas para analisar as determinantes da propagação da Covid-19 pelo território de Portugal continental com escala ao nível municipal. Além do acrónimo da variável e sua descrição, a tabela também apresenta a data da última informação disponível, a fonte e o grupo a que pertence. Na descrição, “(%)” significa que a variável está apresentada em percentagem da população residente ou em percentagem do valor total. A variável dependente, casos acumulados por 100 mil habitantes é denotada por *ca100m*. Todas as outras são variáveis explicativas.*

² Os dados acerca do número de casos acumulados por concelho são retirados do boletim diário da DGS de dia 26 de outubro de 2020, no entanto os dados aí expostos correspondem a casos confirmados até ao final do dia 25 de outubro de 2020.

As estatísticas descritivas das 56 variáveis encontram-se em anexo na Tabela A.1. A variável dependente (*ca100m*) possuiu valores entre 0 e 3 252,362, média de 596,37 e o seu coeficiente de variação é inferior a 1 (de 0,817). As variáveis com maior variabilidade são *br*, *el*, *af* e *denspop*, todas elas com valores superiores a 2; outras quatro variáveis possuem um coeficiente maiores que 1, são elas *pcid*, *bs*, *pe*, *med* e *ipss*. As variáveis com menor dispersão, com coeficiente de variação inferior a 0,010, são *ipm* e *emvn*.

Antes de avançar é justificado analisar a Figura 1, acerca da distribuição de casos acumulados por 100 mil habitantes por município registados até ao dia 25 de outubro de 2020.

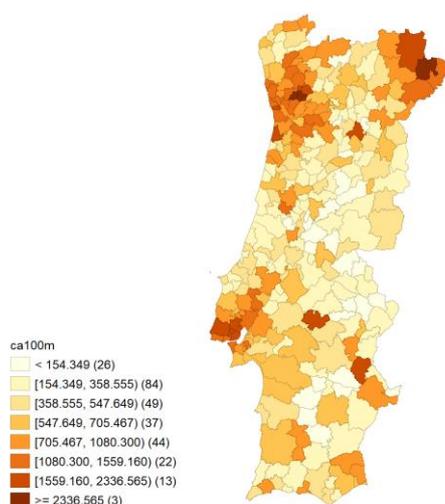


Figura 1- Distribuição do número de casos acumulados por 100 000 habitantes

À data, os concelhos mais afetados pela Covid, tendo em conta o número de casos acumulados por 100 mil habitantes, eram: Paços de Ferreira (3252,36), Lousada (2679,93), Vimioso (2336,56), Amadora (2021,66), Loures (1948,64), Reguengos de Monsaraz (1917,12), Sintra (1904,44), Sernancelhe (1853,22), Bragança (1838,90) e Lisboa (1806,03). Oito concelhos ainda não tinham qualquer caso registado: Alvito, Barrancos, Castanheira de Pêra, Castelo de Vide, Fronteira, Monforte, Penalva do Castelo e Penedono.

4.2 Análise preliminar

Devido à elevada dimensionalidade do espaço de regressores procedeu-se a uma seleção prévia destes de acordo com os seguintes critérios. (1) Dentro de cada grupo é selecionada a variável que apresenta uma maior correlação com a variável dependente, *ca100m*. Mais precisamente dentro de cada grupo é estimada uma regressão simples de *ca100m* sobre cada variável explicativa, selecionando-se aquela com maior R^2 . (2) É analisada a correlação entre as

variáveis explicativas. Se duas variáveis têm entre si um coeficiente igual ou superior a 0,75, então é descartada aquela variável que tem uma menor correlação com a variável dependente. (3) São excluídas da análise as variáveis não significativas, ao nível de significância de 10%, no modelo introdutório³.

Foram 15 as variáveis explicativas que passaram o critério (1) de seleção prévia (ver Tabela A.2, em anexo): *p9010*, *pcompra*, *privms*, *psdoe*, *purb*, *des*, *denspop*, *trcol*, *tempm*, *asup*, *idi*, *se*, *ipss*, e *br* e *af*⁴. Pela aplicação do critério (2) (ver Tabela A.3 em anexo) foram descartadas 3 variáveis *se*, *ipss* e *idi*. Finalmente, revelaram-se não significativas no modelo introdutório outras 3 variáveis *psdoe*, *trcol* e *br*. Portanto, após a aplicação dos 3 critérios, o espaço de regressores ficou reduzido às seguintes 9 variáveis: *p9010* (indicador de desigualdade de rendimento), *pcompra* (indicador de rendimento médio), *privms* (indicador de pobreza), *purb* (indicador urbano/rural), *des* (indicador de emprego); *denspop* (indicador demográfico), *tempm* (indicador de clima), *asup* (indicador de habitação), e *af* (indicador de população estrangeira).

O fator de inflação da variância (VIF) aplicado à regressão multivariada pelo MMQ aponta o maior valor individual (3,48) à variável *pcompra*, enquanto o VIF global do modelo é apenas 2,17. Estes valores, substancialmente inferiores ao limite de 10, são indicativos de inexistência de multicolinearidade

A estimação da regressão multivariada pelo MMQ considerando apenas as variáveis significativas está apresentada na Tabela 2. A regressão MMQ consegue explicar 43,10% da variabilidade do número de casos acumulados de Covid-19 por cada 100 mil habitantes por município. As variáveis *p9010*, *des* e *tempm* apresentam sinais negativos.

³ A presente análise preliminar foi efetuada através do *software* Stata 15.

⁴ Do grupo População Estrangeira foram selecionadas as variáveis *br* e *af*, as quais têm R² semelhantes. As duas variáveis analisam grupos de população diferentes e com uma distribuição diferenciada no território. Estas variáveis possuem entre si um coeficiente de correlação inferior a 0,75. A inclusão da variável *af* é em parte justificada pela possibilidade de segregação destas comunidades em resultados de comportamentos racistas (Amnistia Internacional, 2021).

Tabela 2 – Estimação da regressão multivariada pelo MMQ

	MMQ
<i>constante</i>	515,694
<i>p9010</i>	-65,710***
<i>pcompra</i>	4,569**
<i>privms</i>	158,364***
<i>purb</i>	1,709*
<i>des</i>	-42,508*
<i>asup</i>	36,353***
<i>denspop</i>	0,123***
<i>tempm</i>	-62,849***
<i>af</i>	21,902***
R²*	0,431

Nota: Apresenta significância estatística ao nível de 10 % (*); 5% (**) e 1% (***).

5. Resultados

Os testes de diagnóstico espacial e de Lagrange/especificação, as estimações dos modelos espaciais e as estatísticas de I de Moran local e I de Moran local com duas variáveis foram calculados no software GeoDa 1.18⁵. Respeitando a primeira regra da geografia de Tobler (1970), todas as coisas estão relacionadas umas com as outras, no entanto as mais próximas estão mais relacionadas do que as mais distantes, foi construída uma matriz com relação de continuidade de rainha de primeira ordem, isto é, uma matriz de 1's e 0's, sendo 1 para os concelhos vizinhos do município *i*, 0 para os concelhos não vizinhos do município *i*. Para cálculo dos modelos locais foi usado software MGWR 2.2⁶, e foi usado um kernel bi-quadrado para minimizar o Critério de informação de Akaike corrigido (AICc). Os coeficientes das regressões locais foram padronizados⁷ para permitir comparações.

O teste de diagnóstico espacial, I de Moran (erros), confirma a presença de efeitos espaciais. Os testes de Lagrange ou de especificação, indicam que SLM é o melhor modelo, embora com resultados similares com o SEM.

Os modelos espaciais são executados uma primeira vez com todas as variáveis selecionadas previamente, depois são retiradas as variáveis não significativas (modelo parcimonioso). No modelo SLM parcimonioso foram retiradas as variáveis *purb*, *des*, *tempm*, *asup* e *denspop*. O modelo SEM parcimonioso não inclui as variáveis *purb*, *des*, *tempm* e *asup*. Em ambos os modelos o *rho* e *lambda* são significativos e positivos, demonstrando o efeito de difusão existente

⁵ <https://geodacenter.github.io/GeoDaSpace/>

⁶ <https://sgsup.asu.edu/sparc/multiscale-gwr>

⁷ Variáveis padronizadas de média 0 e desvio-padrão 1.

entre municípios. O pseudo-R² dos modelos varia entre 52.4% e 54.1%. Dos modelos estimados, SLM possui o maior (pseudo) R², e SLM parcimonioso o menor Critério de Informação de Akaike⁸.

Tabela 3 – Regressões espaciais

	SLM	SLM parcimonioso	SEM	SEM parcimonioso
<i>constante</i>	237,932	-635,046	218,951	-604,372**
<i>p9010</i>	-60,554***	-62,868**	-57,500***	-59,405***
<i>pcompra</i>	4,422**	6,251***	5,342***	5,730***
<i>privms</i>	107,494***	152,585***	168,398***	203,626***
<i>purb</i>	1,070	----	0,944	----
<i>des</i>	-15,960	----	-36,186	----
<i>asup</i>	15,841	----	26,270**	----
<i>denspop</i>	0,056	----	0,113***	0,117***
<i>tempm</i>	-38,560**	----	-48,682*	----
<i>af</i>	17,300**	22,424***	19,209***	21,789***
<i>rho</i>	0,484***	0,556***	n.a.	n.a.
<i>lambda</i>	n.a.	n.a.	0,492***	0,554***
Pseudo-R²	0,541	0,533	0,529	0,524
Critério de informação de Akaike	4049,46	4049,23	4055,07	4054,27

Nota: Apresenta significância estatística ao nível de 10 % (*); 5% (**) e 1% (***).

As estatísticas I de Moran local (Figura 22, na Discussão dos Resultados) e I de Moran local com duas variáveis (Figuras A.1 a A.8, em anexo) são calculadas possibilitando uma melhor análise pelo território e identificação de *clusters*. Também são calculados os I de Moran local e com 2 variáveis da AM de Lisboa e Porto (Figuras A.9 a A.13 em anexo). Todas as estatísticas mencionadas são significantes (valor-p < 0.05).

Por último, são estimados os modelos locais. O MGWR possui o menor AICc, de 599,944, contra 626,818 do modelo GWR. Os valores do AIC dos modelos locais, são muito maiores comparados aos dos modelos locais, MGWR com 593,371 e GWR com 610,155. O R² de GWR é 61,6%, já o MGWR possui um R² de 59,6%.

Observando os resultados do R² e dos coeficientes padronizados das variáveis (Figura 2 a 21), vemos que os resultados são semelhantes. As diferenças existentes são entre os coeficientes das variáveis *p9010* e *pcompra*.

⁸ O ideal era analisar o Critério de Informação de Akaike corrigido (AICc), indicado para modelos com poucas observações. No entanto não é possível obter esta informação usando o GeoDa.

Com os coeficientes padronizados é possível comparar o impacto das diferentes variáveis por todo o território. Também nos permitirá identificar a variável a nível local mais importante para a propagação da Covid-19. Fizemos também a análise complementar com os resultados do I de Moran local, permitindo-nos verificar se existe ou não um efeito de contágio provocado pelas variáveis a nível local.

Usando o modelo MGWR com variáveis padronizadas obtém-se informação acerca das bandas largas, portanto a escala espacial da relação entre a variável dependente e explicativa. O mesmo permite-nos concluir que, das nove variáveis, apenas *asup* é uma variável com influência local, a banda larga possui informação de 43 municípios. As bandas largas das restantes variáveis possuem informação de 275 ou mais municípios.

Os resultados dos modelos locais estão expostos a seguir:

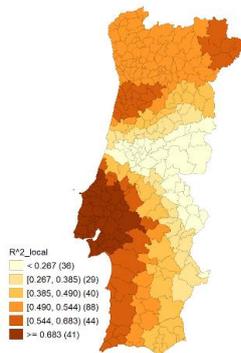


Figura 2 – R² do modelo MGWR

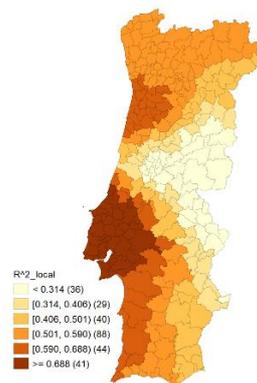


Figura 3 – R² do modelo GWR

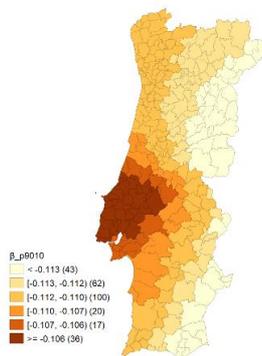


Figura 4 – Coeficiente padronizado da variável *p9010* do modelo MGWR

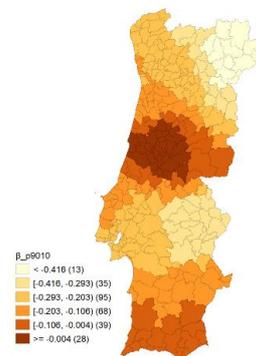


Figura 5 – Coeficiente padronizado da variável *p9010* do modelo GWR

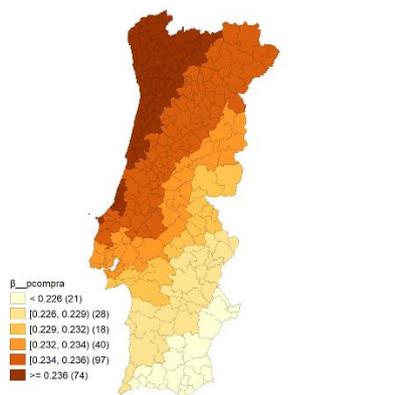


Figura 6 – Coeficiente padronizado da variável *pcompra* do modelo MGWR

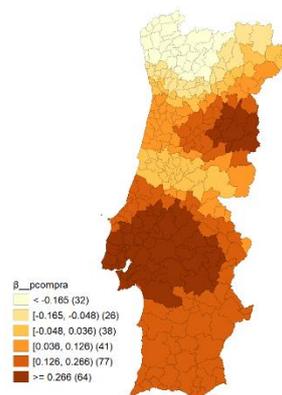


Figura 7 – Coeficiente padronizado da variável *pcompra* do modelo GWR

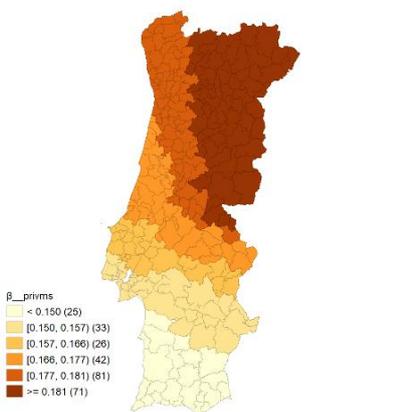


Figura 8 – Coeficiente padronizado da variável *privms* do modelo MGWR

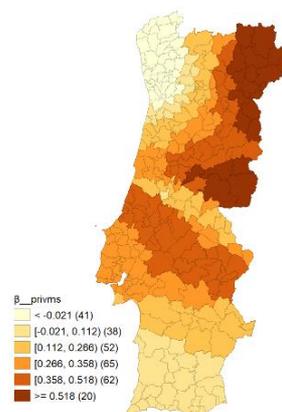


Figura 9 – Coeficiente padronizado da variável *privms* do modelo GWR

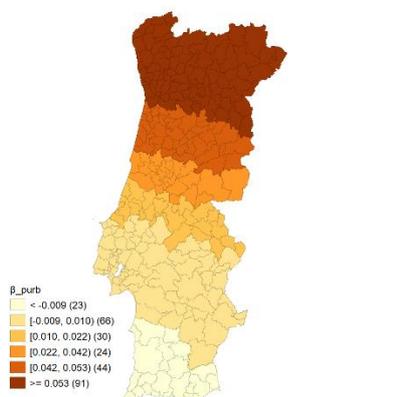


Figura 10 – Coeficiente padronizado da variável *purb* do modelo MGWR

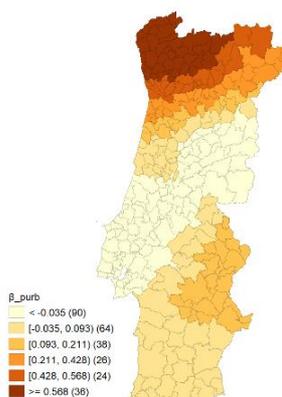


Figura 11 – Coeficiente padronizado da variável *purb* do modelo GWR

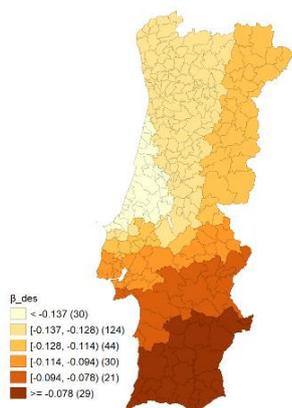


Figura 12 – Coeficiente padronizado da variável *des* do modelo MGWR

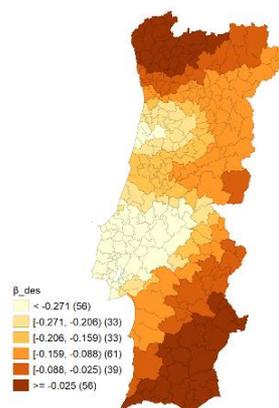


Figura 13 – Coeficiente padronizado da variável *des* do modelo GWR

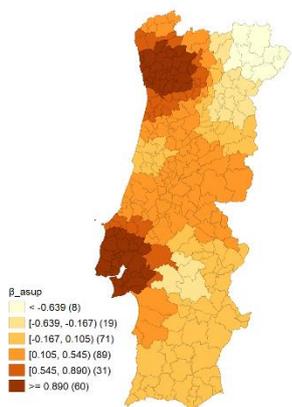


Figura 14 – Coeficiente padronizado da variável *asup* do modelo MGWR

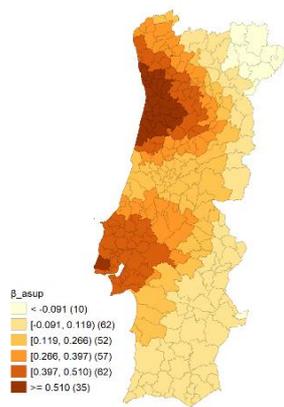


Figura 15 – Coeficiente padronizado da variável *asup* do modelo GWR

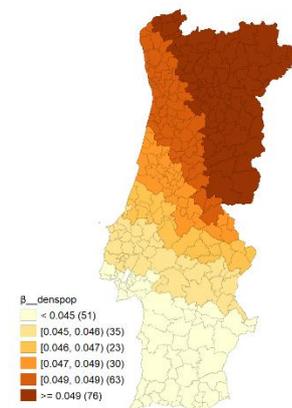


Figura 16 – Coeficiente padronizado da variável *denspop* do modelo MGWR

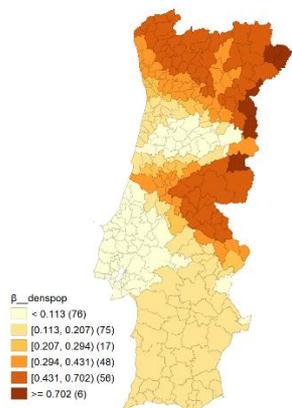


Figura 17 – Coeficiente padronizado da variável *denspop* do modelo GWR

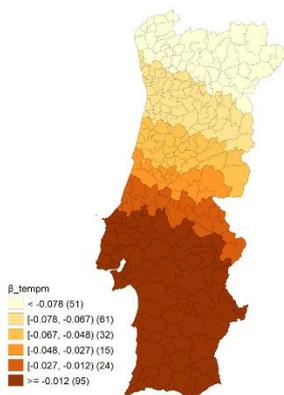


Figura 18 – coeficiente padronizado da variável *tempm* do modelo MGWR

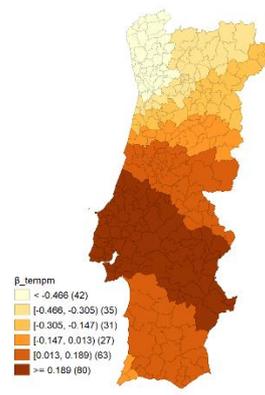


Figura 19 – Coeficiente padronizado da variável *tempm* do modelo GWR

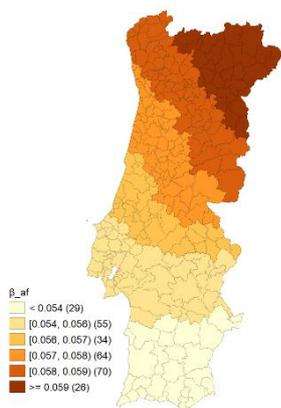


Figura 20 – Coeficiente padronizado da variável *af* do modelo MGWR

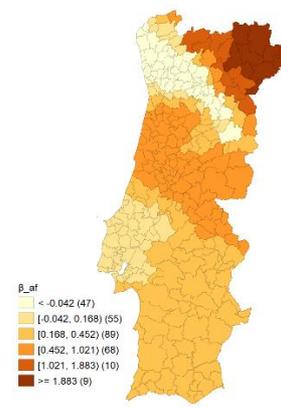


Figura 21 – Coeficiente padronizado da variável *af* do modelo GWR

6. Discussão de resultados

6.1 Interpretação dos resultados

Antes de partir para a análise das determinantes da propagação da Covid-19, é igualmente importante, através das Figuras 1 e 22, identificar quais os locais mais afetados. Como era de esperar, a maior concentração verifica-se nos municípios das Áreas Metropolitanas de Lisboa e do Porto. Uma vez que a transmissão está dependente da proximidade e interação entre pessoas é expectável que os epicentros da pandemia sejam grandes zonas urbanas, de elevada densidade populacional e de grande atividade económica. Os resultados confirmam aquilo que a literatura registou para outros países. As duas áreas metropolitanas são os principais centros económicos do país e possuem maior interação com o resto do mundo. Os principais focos de casos acumulados por 100 mil habitantes eram à data a Área Metropolitana de Lisboa, Porto e

sua periferia, mas também encontramos vários concelhos fora destas zonas, a vermelho-escuro, no distrito de Bragança (Figura 22). Os concelhos identificados, a vermelho-escuro, apresentam uma forte correlação com os casos dos concelhos vizinhos. No interior do país é onde se verifica menos efeito de contágio, a azul-escuro. O restante território possuirá uma distribuição de casos acumulados menos influenciada (mais próxima de uma distribuição aleatória) pelos casos acumulados de municípios vizinhos (Figura 22).

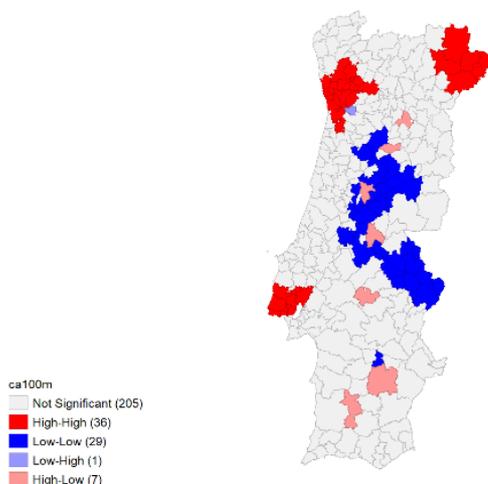


Figura 22- I de Moran local de *ca100m*

Foi ainda calculado o I de Moran local e do I de Moran local com duas variáveis para as Áreas Metropolitanas de Lisboa e do Porto (Figuras A.9 a A.13, em anexo). Em Lisboa, os principais epicentros da pandemia foram Sintra e Odivelas. Odivelas é ainda o concelho mais afetado pela propagação do vírus provocada por alojamentos sobrelotados e situações de fragilidade de proteção face à pandemia por parte da população com origem em países africanos. Foram ainda registadas duas relações *high-high* entre *ca100m* e *denspop* nos concelhos de Lisboa e Amadora. Na Área Metropolitana do Porto, apenas de realçar o facto de Paredes ser o único concelho mais afetado pela quantidade de alojamentos sobrelotados dos concelhos vizinhos.

Passando à análise das determinantes e suportando-me nos resultados dos vários modelos, as condições socioeconómicas demonstram ser das principais determinantes da propagação da Covid-19. Dos modelos espaciais (SLM e SEM), a variável que demonstra ser mais importante

para explicar a propagação da Covid-19 é a privação material severa⁹, uma vez que possui o maior coeficiente. O indicador permite identificar situações de vulnerabilidade socioeconómica importantes, por exemplo, não ter um carro individual obriga as pessoas e famílias a usarem transportes públicos ou coletivos. É provável que quem esteja numa condição de privação material severa, não possua uma boa alimentação, não tenha acesso a cuidados de saúde de maneira célere e gratuita durante a pandemia, ou não apresente capacidade para se precaver contra uma situação de perda inesperada de rendimento. Dados recentemente publicados, com base num inquérito do Instituto Nacional de Estatística, afirmam que um em cada cinco portugueses está em situação de pobreza, sendo que a maior parte possui um emprego sem termo, trabalho precário ou está em idade de reforma. Este indicador (*privms*) está interligado com a variável *asup*, sendo um dos itens a capacidade para pagar atempadamente rendas, a incapacidade de tal pode colocar um agregado em más condições de habitabilidade, ou mesmo num alojamento superlotado. Problemas com a habitação, neste caso superlotação, ocorrem em municípios de Lisboa, Porto e sua periferia (ver Figuras 14, 15 nos resultados e A.13 em anexo). O modelo local permite confirmar em linha com os resultados da estatística I de Moran, que o efeito contágio provocado pela sobrelotação de alojamento é um fenómeno local, exclusivo das Áreas Metropolitanas de Lisboa e do Porto.

Outro fator socioeconómico importante são as desigualdades. A relação entre o número de casos acumulados e as desigualdades (*p9010*) na regressão simples é positiva. Os municípios onde as desigualdades são maiores, também são os que apresentam maiores valores de casos acumulados por 100 mil habitantes. Pelo litoral português e região de Lisboa e Vale do Tejo, as desigualdades são das variáveis mais importantes para a propagação do vírus (ver Figura 4, nos resultados). Os resultados são corroborados pela estatística do I de Moran (Figura A.1, em anexo). As principais determinantes da propagação da Covid-19 dos modelos são as socioeconómicas, o que demonstra que as mesmas potenciaram a propagação do vírus pelo território. A propagação da pandemia é influenciada por condicionantes anteriores à mesma – em linha, também neste caso, com aquilo que a literatura refere, nomeadamente o facto de a

⁹ Uma situação de privação material severa verifica-se quando pelo menos quatro dos seguintes itens, não é observado devido a dificuldades económicas: capacidade para assegurar o pagamento imediato de uma despesa inesperada e próxima do valor mensal da linha de pobreza; capacidade para pagar uma semana de férias, por ano, fora de casa, suportando a despesa de alojamento e viagem para todos os membros do agregado; capacidade para pagar atempadamente rendas, prestações de crédito ou despesas correntes da residência principal, ou outras despesas não relacionadas com a residência principal; capacidade para ter uma refeição de carne ou de peixe (ou equivalente vegetariano), pelo menos de 2 em 2 dias; capacidade para manter a casa adequadamente aquecida; capacidade para ter máquina de lavar roupa; capacidade para ter televisão a cores; capacidade para ter telefone fixo ou telemóvel; capacidade para ter automóvel.

população residente em localidades onde existem maiores índices de desigualdade e pobreza estarem mais expostas ao risco de contágio (Ahmed et al, 2020).

A *proxy* calculada para o desemprego registado nos municípios possui um coeficiente de sinal negativo. Contudo, a análise da Figura A.5 do I de Moran de duas variáveis permite identificar três locais diferentes onde se observa uma relação *high-high* (vermelho-escuro). Temos a norte do país a maioria dos municípios identificados a pertencerem à região do Tâmega e Sousa, no Alentejo Central o município de Reguengos de Monsaraz e no Baixo Alentejo o município de Moura, na região algarvia, Albufeira. Segundo os resultados dos modelos locais MGWR e GWR (Figuras 12 e 13, nos Resultados) o desemprego é das variáveis mais importantes para explicar a lastração do vírus na zona sul de Portugal, acrescentado a região do Alto Minho e Cavado do modelo GWR.

A densidade populacional e a população urbana são fatores importantes para o estudo da propagação do vírus pelo território muitas vezes esquecido em vários estudos (Liu et al., 2021). A inclusão destas variáveis no modelo é importante. Com efeito, as zonas urbanas apresentam maiores densidades populacionais e taxas de atividade económica mais altas, variáveis que facilitam a transmissão de Covid-19 (Ramírez-Aldana et al., 2020; Hamidi et al., 2020). A densidade populacional assim como a percentagem de população urbana são dos principais fatores de transmissão do vírus entre municípios das grandes áreas metropolitanas (Figuras A.4 e A.6 em anexo). As Figuras 16 e 17, indicam que a densidade populacional é dos principais fatores para explicar a propagação do vírus no interior do país, apesar de a mesma neste território ser baixa. Nas zonas de elevada densidade populacional existem outras determinantes que demonstram ter maior importância, como as desigualdades, a pobreza ou o poder de compra. Nas localidades onde o poder de compra é maior também é onde existe maior atividade económica e concentração de população, facilitando assim o contágio (Figuras 6 e 7 nos resultados, Figura A.2 em anexo).

O recente relatório da Amnistia Internacional expressa preocupação com a discriminação contra a população afrodescendente, no acesso à educação, emprego e à habitação, também ela é objeto de discurso e crimes de ódio (Amnistia Internacional, 2021). Tendo em conta os resultados, zonas onde a percentagem de população com origem em países africanos de expressão portuguesa é maior, são também zonas mais afetadas pela Covid-19. A literatura existente (Gray et al. 2020) pode explicar este resultado, a discriminação racial pode pôr em situação vulnerável as suas vítimas, podendo assim criar desigualdades de rendimento, más condições

habitacionais, falta de qualidade alimentar, difícil acesso a cuidados de saúde, ou perda de bem-estar ou saúde. Em linha com o que foi dito, esta população está associada a zonas com população em situações de pobreza, desigualdades e com problemas na habitação. Mais especificamente nos concelhos da periferia da cidade de Lisboa, por exemplo da população residente na Amadora, 45,6% da população possui origens de países africanos de expressão portuguesa. Sendo esse um dos municípios identificados na Figura A.7, (em anexo) a vermelho-escuro, os concelhos identificados possuem percentagem de população africana e número de casos acumulados por Covid-19, acima da média nacional. Os modelos locais (Figuras 20 e 21) identificam a variável *af* como das principais variáveis para explicar a propagação da Covid-19 em vários concelhos de Trás-os-Montes. Entre esses concelhos estão Bragança, onde 2% da população é afrodescendente e Mirandela com 0,33%. Apesar de os valores da percentagem da população afrodescendente serem reduzidos, esta variável mostra ter mais importância que as restantes para explicar o fenómeno na região.

Os modelos locais permitem analisar a heterogeneidade do território nacional através da análise do R^2 local. As nove variáveis do modelo permitem obter uma grande capacidade explicativa propagação da Covid-19 na região de Lisboa e Vale do Tejo, Norte e Litoral e menor no interior do país (Figuras 2 e 3, nos resultados). O interior de Portugal tem sido menos atingido pela Covid-19, também é no interior onde se encontram menores níveis de desigualdades, menos proporção de alojamentos sobrelotados, percentagem de população que vive em zonas urbanas, percentagem de população com origens em países africanos e densidade populacional (ver Figuras A.1 a A.8).

6.2 Robustez do modelo

Quando comparados os resultados obtidos com o que aconteceu durante a pandemia em Portugal, o estudo identifica o território e características da população afetada. Um dos primeiros surtos que se registou em Portugal foi numa fábrica em Felgueiras¹⁰ perto do Porto. Um município em que a atividade económica desenvolve-se em torno da indústria, onde grande parte dos trabalhos não podem ser feitos de remotamente. O município situa-se numa zona de alojamentos sobrelotados, o que potencia a propagação do vírus. Os primeiros surtos da pandemia surgem em concelhos limítrofes da capital portuguesa¹¹. Onde há mais pobreza, mais

¹⁰<https://www.publico.pt/2020/11/09/infografia/caso-fabrica-calcado-desenrolou-surtos-covid19-portugal-519>

¹¹<https://www.publico.pt/2020/05/27/sociedade/noticia/chegada-surtos-zonas-maior-pobreza-torna-contencao-dificil-1918211>

desigualdades, onde uma parte significativa da população residente tem origem em países africanos de expressão portuguesa. As condições de habitabilidade são más e, devido à sua condição social, grande parte da população possui mais do que um trabalho, o que leva que haja mais deslocações, consequentemente maior número de contactos e aumento do risco de contágio. Outros casos surgiram, como o episódio da cerca sanitária da comunidade cigana de Moura no Alentejo¹². Mais recentemente no litoral alentejano¹³, comunidades imigrantes que trabalham em explorações agrícolas, com trabalhos mal remunerados, clara exploração laboral, com habitação sobrelotadas e sem as mais elementares condições de habitabilidade, têm sido as principais vítimas desta região e de outros locais¹⁴. Um dos setores de atividade onde se registou maior número de trabalhadores infetados foi o da construção civil¹⁵, neste ramo não há possibilidade de o trabalho ser feito remotamente, as condições de trabalho e o recurso a mão-de-obra imigrante sujeita a situações de trabalho e de habitabilidade que favorecem a transmissão do vírus, são determinantes que favoreceram a expansão da pandemia neste setor.

O modelo não identifica os equipamentos de solidariedade social como lares, nem as faixas etárias dos que possuem mais de 65 anos como os principais determinantes da propagação de Covid-19. Existiram vários surtos em lares por todo o país¹⁶, no entanto estes fatores podem não ter sido um dos principais para a propagação da Covid-19 por Portugal continental.

7. Conclusão

Este trabalho, a nosso ver, conseguiu responder às principais questões colocadas no início da nossa investigação. Sendo identificadas, com uma regressão multivariada estimada pelo MMQ, nove possíveis determinantes cruciais para explicar a propagação da Covid-19 em Portugal continental, durante a primeira fase da pandemia. As variáveis identificadas pelo modelo são: poder de compra, desigualdade de rendimentos dos agregados familiares entre o percentil 90 e 10, privação material severa, alojamentos superlotados, desemprego, população urbana,

¹²https://www.rtp.pt/noticias/pais/comunidade-cigana-de-moura-isolada-apos-varios-casos-de-covid-19_v1220193

¹³<https://www.publico.pt/2021/04/30/local/noticia/infetados-trancas-porta-insalubres-alojamentos-imigrantes-1960691>

¹⁴<https://expresso.pt/coronavirus/2021-05-11-Covid-19.-Surto-entre-trabalhadores-agricolas-em-Torres-Vedras-aumenta-para-37-casos-ativos-892099f4>

¹⁵https://www.rtp.pt/noticias/pais/ha-trabalhadores-da-construcao-civil-a-morrer-com-covid-19-alerta-sindicato_a1215923 e <https://tvi24.iol.pt/sociedade/dgs/covid-19-ministra-confirma-13-surtos-preocupantes-em-lisboa-loures-odivelas-amadora-e-sintra>

¹⁶<https://expresso.pt/sociedade/2020-06-20-Covid-19-surto-em-lar-de-Reguengos-de-Monsaraz-pode-ja-estar-na-comunidade>; <https://www.jornaldeleiria.pt/noticia/covid-19-dois-surtos-em-lares-de-pombal-com-117-infetados> e <https://tvi24.iol.pt/sociedade/coronavirus/covid-19-surto-em-lar-da-guarda-com-130-casos>

densidade populacional, temperatura média e percentagem população com origem em países africanos de expressão portuguesa. Entre elas, as variáveis socioeconómicas (a variável de desigualdade e de pobreza) mostraram ser as mais importantes e fulcrais da disseminação da Covid-19.

Para incorporar as dependências espaciais foram usados modelos espaciais e locais. O modelo SLM, apresentou o maior R^2 de 54%, ou seja, conseguimos explicar mais de metade dos casos acumulados de Covid-19. Complementando o nosso estudo com estatísticas espaciais, como o I de Moran, conseguimos identificar *clusters* nacionais onde existe maior efeito propagação (a norte de Portugal e em Lisboa e Vale do Tejo) e menor efeito de propagação (centro-interior de Portugal); com o, I de Moran local com duas variáveis, identificar a nível local se as variáveis podem ou não ter efeito no número de casos acumulados. Os modelos locais permitiram aferir com mais precisão o diferente impacto (não-estacionariedade espacial) das variáveis pelo território e confirmar como a heterogeneidade do território continental teve influência no fenómeno pandémico.

As nossas conclusões corroboram com o que tem sido exposto de outros estudos acerca do impacto das determinantes socioeconómicas, de habitabilidade, demográficas e étnico-raciais, na explicação da propagação da pandemia da Covid-19. Este estudo também possuiu uma pequena limitação: para se conseguir estimar os modelos espaciais e locais tivemos de preencher dados em falta, usando valores de NUT II e NUT III, podendo alterar um pouco os resultados finais. O presente estudo pode ser incluído na literatura acerca da economia da saúde e de estudos com modelos espaciais e locais. Ainda prova ser uma ferramenta útil para a decisão política no combate à pandemia, fragilidades sociais e mitigação das desigualdades territoriais.

Bibliografia

- Adams-Prassl, A., Boneva, T., Golin, M., & Rauh, C. (2020). Inequality in the impact of the coronavirus shock: Evidence from real time surveys. *Journal of Public Economics*, 189, 104245. <https://doi.org/10.1016/j.jpubeco.2020.104245>
- Ahmed, F., Ahmed, N., Pissarides, C., & Stiglitz, J. (2020). Why inequality could spread COVID-19. *The Lancet Public Health*, 5(5), e240. [https://doi.org/10.1016/S2468-2667\(20\)30085-2](https://doi.org/10.1016/S2468-2667(20)30085-2)
- Amdaoud, M., Arcuri, G., & Levratto, N. (2021). Are regions equal in adversity? A spatial analysis of spread and dynamics of COVID-19 in Europe. *European Journal of Health Economics*, 1–14. <https://doi.org/10.1007/s10198-021-01280-6>
- Amnesty Internacional Report 2020/21 - The state of the world's human rights. (2021). In *Amnesty Internacional*. Acedido no dia 7 de abril de 2020 através do sítio <https://www.amnesty.org/download/Documents/POL1032022021BRAZILIANPORTUGUESE.PDF>
- Andersen, L. M., Harden, S. R., Sugg, M. M., Runkle, J. D., & Lundquist, T. E. (2021). Analyzing the spatial determinants of local Covid-19 transmission in the United States. *Science of the Total Environment*, 754, 142396. <https://doi.org/10.1016/j.scitotenv.2020.142396>
- Anselin, L. (1995). Local Indicators of Spatial Association—LISA. *Geographical Analysis*, 27(2), 93–115. <https://doi.org/10.1111/j.1538-4632.1995.tb00338.x>
- Anselin, L. (2003). Spatial externalities, spatial multipliers, and spatial econometrics. *International Regional Science Review*, 26(2), 153–166. <https://doi.org/10.1177/0160017602250972>
- Anselin, L., & Arribas-Bel, D. (2013). Spatial fixed effects and spatial dependence in a single cross-section. *Papers in Regional Science*, 92(1), 3–17. <https://doi.org/10.1111/j.1435-5957.2012.00480.x>
- Ascani, A., Faggian, A., & Montresor, S. (2021). The geography of COVID-19 and the structure of local economies: The case of Italy. *Journal of Regional Science*, 61(2), 407–441. <https://doi.org/10.1111/jors.12510>
- Baum, C. F., & Henry, M. (2020). Socioeconomic Factors influencing the Spatial Spread of COVID-19 in the United States. *Miguel, Socioeconomic Factors Influencing the Spatial Spread of COVID-19 in the United States*, (May 29, 2020). <https://doi.org/10.2139/ssrn.3614877>

- Blundell, R., Costa Dias, M., Joyce, R., & Xu, X. (2020). COVID-19 and Inequalities. *Fiscal Studies*, 41(2), 291–319. <https://doi.org/10.1111/1475-5890.12232>
- Borjas, G. J. (2020). Demographic Determinants of Testing Incidence and COVID-19 Infections in New York City Neighborhoods. *National Bureau of Economic Research*. <https://doi.org/10.2139/ssrn.3572329>
- Brunsdon, C., Fotheringham, S., & Charlton, M. (1996). Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity. *International Encyclopedia of Human Geography*, 28(4), 407–414. <https://doi.org/10.1016/B978-008044910-4.00447-8>
- Brunsdon, C., Fotheringham, S., & Charlton, M. (1998). Geographically weighted regression-modelling spatial non-stationarity. *Journal of the Royal Statistical Society*, 47(3), 431–443. <https://doi.org/https://doi.org/10.1111/j.1538-4632.1996.tb00936.x>
- Coven, J., & Gupta, A. (2020). Disparities in mobility responses to COVID-19. In *New York University*. <https://static1.squarespace.com/static/56086d00e4b0fb7874bc2d42/t/5ebf201183c6f016ca3abd91/1589583893816/DemographicCovid.pdf>
- Ehlert, A. (2020). The socio-economic determinants of COVID-19: A spatial analysis of German county level data. *Socio-Economic Planning Sciences*, 101083. <https://doi.org/10.1101/2020.06.25.20140459>
- Fotheringham, A. S., & Oshan, T. M. (2016). Geographically weighted regression and multicollinearity: dispelling the myth. *Journal of Geographical Systems*, 18(4), 303–329. <https://doi.org/10.1007/s10109-016-0239-5>
- Fotheringham, A. S., Yang, W., & Kang, W. (2017). Multiscale Geographically Weighted Regression (MGWR). *Annals of the American Association of Geographers*, 107(6), 1247–1265. <https://doi.org/10.1080/24694452.2017.1352480>
- Galasso, V. (2020). COVID: Not a great equalizer. *CESifo Economic Studies*, 66(4), 376–393. <https://doi.org/10.1093/cesifo/ifaa019>
- Gerritse, M. (2020). Cities and COVID-19 infections: Population density, transmission speeds and sheltering responses. *Covid Economics*, 37, 1–26.
- Ginsburgh, V., Magerman, G., & Natali, I. (2021). COVID-19 and the role of inequality in French regional departments. *European Journal of Health Economics*, 22(2), 311–327. <https://doi.org/10.1007/s10198-020-01254-0>
- Gray, D. M., Anyane-Yeboah, A., Balzora, S., Issaka, R. B., & May, F. P. (2020). COVID-19

- and the other pandemic: populations made vulnerable by systemic inequity. *Nature Reviews Gastroenterology and Hepatology*, 17(9), 520–522.
<https://doi.org/10.1038/s41575-020-0330-8>
- Guliyev, H. (2020). Determining the spatial effects of COVID-19 using the spatial panel data model. *Spatial Statistics*, 38, 1004433. <https://doi.org/10.1016/j.spasta.2020.100443>
- Hamidi, S., Sabouri, S., & Ewing, R. (2020). Does Density Aggravate the COVID-19 Pandemic?: Early Findings and Lessons for Planners. *Journal of the American Planning Association*, 86(4), 495–509. <https://doi.org/10.1080/01944363.2020.1777891>
- 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*, 94, 96–102.
<https://doi.org/10.1016/j.ijid.2020.03.076>
- Karmakar, M., Lantz, P. M., & Tipirneni, R. (2021). Association of Social and Demographic Factors With COVID-19 Incidence and Death Rates in the US. *JAMA Network Open*, 4(1), e2036462. <https://doi.org/10.1001/jamanetworkopen.2020.36462>
- Khanijahani, A. (2021). Racial, ethnic, and socioeconomic disparities in confirmed COVID-19 cases and deaths in the United States: a county-level analysis as of November 2020. *Ethnicity and Health*, 26(1), 22–35. <https://doi.org/10.1080/13557858.2020.1853067>
- Liu, S. H., Li, Y., & Liu, B. (2021). RE: Association Between State-Level Income Inequality and COVID-19 Cases and Mortality in the USA. *Journal of General Internal Medicine*, 1–1. <https://doi.org/10.1007/s11606-021-06604-z>
- Maiti, A., Zhang, Q., Sannigrahi, S., Pramanik, S., Chakraborti, S., Cerda, A., & Pilla, F. (2021). Exploring spatiotemporal effects of the driving factors on COVID-19 incidences in the contiguous United States. *Sustainable Cities and Society*, 68, 102784.
<https://doi.org/10.1016/j.scs.2021.102784>
- 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. <https://doi.org/10.1016/j.scs.2020.102627>
- Marí-Dell’olmo, M., Gotsens, M., Pasarín, M. I., Rodríguez-Sanz, M., Artazcoz, L., de Olalla, P. G., Rius, C., & Borrell, C. (2021). Socioeconomic inequalities in COVID-19 in a European urban area: Two waves, two patterns. *International Journal of Environmental Research and Public Health*, 18(3), 1–12. <https://doi.org/10.3390/ijerph18031256>
- 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. <https://doi.org/10.1016/j.scitotenv.2020.138884>
- Moran, P. (1948). The interpretation of statistical maps. *Journal of the Royal Statistical Society: Series B (Methodological)*, 10(2), 243–251. <https://doi.org/https://doi.org/10.1111/j.2517-6161.1948.tb00012.x>
- Oshan, T. M., Li, Z., Kang, W., Wolf, L. J., & Stewart Fotheringham, A. (2019). MGWR: A python implementation of multiscale geographically weighted regression for investigating process spatial heterogeneity and scale. *ISPRS International Journal of Geo-Information*, 8(6), 269. <https://doi.org/10.3390/ijgi8060269>
- Rahman, M. H., Zafri, N. M., Ashik, F., & Waliullah, M. (2020). Gis-Based Spatial Modeling to Identify Factors Affecting COVID-19 Incidence Rates in Bangladesh. *SSRN Electronic Journal*, 1–36. <https://doi.org/10.2139/ssrn.3674984>
- Ramírez-Aldana, R., Gomez-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), e0008875. <https://doi.org/10.1101/2020.04.19.20071605>
- Sá, F. (2020). Socioeconomic Determinants of COVID-19 Infections and Mortality: Evidence from England and Wales. *Policy Paper Series*, 159. <http://ftp.iza.org/pp159.pdf>
- Sannigrahi, S., Pilla, F., Basu, B., & Basu, A. S. (2020). The overall mortality caused by COVID-19 in the European region is highly associated with demographic composition: A spatial regression-based approach. *ArXiv*, 62(July), 102418. <https://doi.org/10.1016/j.scs.2020.102418>
- Schmitt-Grohé, S., Teoh, K., & Uribe, M. (2020). Covid-19: Testing Inequality in New York City. *National Bureau of Economic Research*. <https://www.nber.org/papers/w27019>
- Sun, F., Matthews, S. A., Yang, T. C., & Hu, M. H. (2020). A spatial analysis of the COVID-19 period prevalence in U.S. counties through June 28, 2020: where geography matters? *Annals of Epidemiology*, 52, 54–59. <https://doi.org/10.1016/j.annepidem.2020.07.014>
- Sun, Y., Hu, X., & Xie, J. (2021). Spatial inequalities of COVID-19 mortality rate in relation to socioeconomic and environmental factors across England. *Science of the Total Environment*, 758, 143595. <https://doi.org/10.1016/j.scitotenv.2020.143595>
- Tai, D. B. G., Shah, A., Doubeni, C. A., Sia, I. G., & Wieland, M. L. (2021). The Disproportionate Impact of COVID-19 on Racial and Ethnic Minorities in the United States. *Clinical Infectious Diseases : An Official Publication of the Infectious Diseases*

Society of America, 72(4), 703–706. <https://doi.org/10.1093/cid/ciaa815>

Tobler, W. R. (1970). A Computer Movie Simulating Urban Growth in the Detroit Region.

Economic Geography, 46(sup1), 234–240. <https://doi.org/10.2307/143141>

Yang, T.-C., Kim, S., Zhao, Y., & Choi, S. E. (2021). Examining spatial inequality in COVID-19 positivity rates across New York City ZIP codes. *Health & Place*, 69, 102574. <https://doi.org/10.1016/j.healthplace.2021.102574>

Anexos

Tabela A.1- Estatísticas Descritivas

<i>Variável</i>	Média	Desvio- Padrão	Mínimo	Máximo	Coefficiente de variação
<i>ca100m</i>	596,370	487,471	0,000	3252,362	0,817
<i>gini</i>	30,536	2,261	24,900	41,100	0,074
<i>p8020</i>	3,265	0,406	2,500	5,600	0,124
<i>p9010</i>	7,230	1,445	4,700	17,400	0,200
<i>dispgmm</i>	27,628	8,846	12,400	121,000	0,320
<i>pcompra</i>	80,515	18,679	55,320	219,630	0,232
<i>rbrutopc</i>	7608,817	1726,297	4352,000	19574,000	0,227
<i>riscpob</i>	22,128	1,608	16,700	23,000	0,073
<i>privm</i>	15,678	1,723	13,300	18,100	0,110
<i>privms</i>	5,437	0,814	4,500	6,600	0,150
<i>ipm</i>	3,506	0,023	3,500	3,600	0,007
<i>riscpobts</i>	17,837	1,613	12,300	18,600	0,090
<i>psdoe</i>	6,506	1,672	2,632	11,491	0,257
<i>psdes</i>	3,111	1,230	0,959	13,131	0,395
<i>prsi</i>	2,403	1,665	0,343	12,735	0,693
<i>pcid</i>	21,313	27,919	0,000	100,000	1,310
<i>purb</i>	41,520	36,629	0,000	100,000	0,882
<i>prur</i>	34,419	26,494	0,000	96,863	0,770
<i>p2m</i>	37,333	30,208	0,000	100,000	0,809
<i>des</i>	3,604	1,256	1,434	10,991	0,349
<i>denspop</i>	306,134	852,459	3,800	7740,500	2,785
<i>trcol</i>	15,927	6,099	2,900	37,600	0,383
<i>tempm</i>	15,433	1,367	11,700	18,000	0,089
<i>dmf</i>	2,542	0,203	2,100	3,100	0,080
<i>asup</i>	8,919	3,042	3,510	21,990	0,341
<i>bs</i>	6,421	10,359	0,000	78,000	1,613
<i>aseb</i>	2,838	1,672	0,360	10,200	0,590
<i>anr</i>	28,825	6,474	9,780	52,140	0,225
<i>anr</i>	4,678	2,036	0,570	15,350	0,435
<i>ie</i>	197,054	101,342	59,300	597,800	0,514
<i>emv65</i>	19,756	0,408	18,450	20,270	0,021
<i>emvn</i>	80,595	0,690	78,760	81,810	0,009
<i>idi</i>	39,555	14,641	15,700	91,200	0,370
<i>ge1t</i>	27,184	3,482	17,048	34,523	0,128
<i>ge2t</i>	47,147	2,514	36,406	52,383	0,053
<i>ge3t</i>	25,669	5,546	14,195	46,271	0,216
<i>ge1h</i>	29,159	3,646	17,850	36,068	0,125
<i>ge2h</i>	48,275	1,958	39,648	54,124	0,041
<i>ge3h</i>	22,565	4,745	12,951	42,480	0,210
<i>ge1m</i>	25,411	3,445	16,361	33,104	0,136
<i>ge2m</i>	46,127	3,358	33,480	52,860	0,073
<i>ge3m</i>	28,462	6,341	15,343	49,692	0,223
<i>h</i>	47,414	0,981	43,470	53,209	0,021
<i>m</i>	52,586	0,981	46,791	56,530	0,019
<i>se</i>	14,932	5,151	4,900	29,200	0,345
<i>b1c</i>	31,113	4,831	16,100	43,700	0,155
<i>b2c</i>	12,622	2,506	7,800	24,000	0,199
<i>b3c</i>	17,632	2,634	9,900	24,500	0,150
<i>sec</i>	13,542	3,054	8,000	21,500	0,226
<i>sup</i>	9,340	4,345	2,800	31,100	0,465
<i>med</i>	278,982	303,230	38,256	3449,458	1,087
<i>enf</i>	501,343	411,243	117,398	3216,910	0,820
<i>ipss</i>	18,295	36,282	1,000	520,000	1,983
<i>br</i>	1,055	3,095	0,000	34,219	2,934
<i>el</i>	0,658	1,509	0,000	11,224	2,293
<i>af</i>	0,792	3,775	0,000	45,599	4,766
<i>pe</i>	2,729	2,919	0,310	21,630	1,070

Tabela A.2 - Regressões Simples

Variável	Coefficiente	Valor-p	R ²	Grupo	
<i>gini</i>	54,226	0,000	0,063	Desigualdade	
<i>p8020</i>	258,292	0,000	0,046		
<i>p9010</i>	86,025	0,000	0,065		
<i>dispgmm</i>	9,372	0,004	0,029		
<i>pcompa</i>	8,629	0,000	0,109	Rendimento médio	
<i>rbrutopc</i>	0,067	0,000	0,057	Pobreza	
<i>riscpob</i>	-76,657	0,000	0,064		
<i>privm</i>	89,126	0,000	0,099		
<i>privms</i>	235,051	0,000	0,154		
<i>ipm</i>	-834,464	0,507	0,002		
<i>riscpobts</i>	-73,71	0,000	0,059	Proteção Social	
<i>psdoe</i>	68,569	0,000	0,055		
<i>psdes</i>	44,999	0,059	0,013		
<i>prsi</i>	1,424	0,936	0,000	Urbano/Rural	
<i>pcid</i>	5,295	0,000	0,092		
<i>purb</i>	5,336	0,000	0,161		
<i>prur</i>	-5,897	0,000	0,103		
<i>p2m</i>	5,618	0,000	0,121		
<i>des</i>	42,532	0,068	0,012	Emprego	
<i>denspop</i>	0,277	0,000	0,235	Demografia	
<i>trcol</i>	26,939	0,000	0,114	Movimentos pendulares	
<i>tempm</i>	-18,368	0,392	0,003	Clima	
<i>dmf</i>	735,914	0,000	0,094	Habitação	
<i>asup</i>	70,343	0,000	0,193		
<i>bs</i>	18,713	0,000	0,158		
<i>aseb</i>	-50,347	0,004	0,030		
<i>anr</i>	6,398	0,158	0,007		
<i>angr</i>	-20,121	0,162	0,007		
<i>ie</i>	-1,345	0,000	0,078		
<i>emv65</i>	98,369	0,171	0,007		
<i>emvn</i>	140,164	0,001	0,039		
<i>idi</i>	-11,734	0,000	0,124		
<i>ge1t</i>	46,103	0,000	0,108	Estrutura Etária	
<i>ge2t</i>	48,601	0,000	0,063		
<i>ge3t</i>	-28,150	0,000	0,103		
<i>ge1h</i>	42,073	0,000	0,099		
<i>ge2h</i>	27,914	0,062	0,013		
<i>ge3h</i>	-29,595	0,000	0,083		
<i>ge1m</i>	47,125	0,000	0,111		
<i>ge2m</i>	42,953	0,000	0,088		
<i>ge3m</i>	-25,962	0,000	0,114		
<i>h</i>	31,669	0,290	0,004		
<i>m</i>	31,669	0,290	0,004		
<i>se</i>	-36,234	0,000	0,147		Educação
<i>b1c</i>	-23,215	0,000	0,053		
<i>b2c</i>	11,327	0,000	0,064		
<i>b3c</i>	35,872	0,001	0,038		
<i>sec</i>	30,814	0,001	0,037		
<i>sup</i>	33,702	0,000	0,090		
<i>med</i>	0,355	0,000	0,049	Saúde	
<i>enf</i>	0,160	0,024	0,018		
<i>ipss</i>	4,065	0,000	0,092		
<i>br</i>	64,459	0,000	0,167	População Estrangeira	
<i>el</i>	97,844	0,000	0,092		
<i>af</i>	50,693	0,000	0,154		
<i>pe</i>	17,855	0,075	0,011		

Tabela A.3- Matriz de Correlação

	p9010	pcompra	privms	psdoe	purb	des	trcol	asup	denspop	tempm	idi	br	af	se	ipss
p9010	1,000														
pcompra	0,568	1,000													
privms	0,409	0,008	1,000												
psdoe	-0,127	0,342	-0,013	1,000											
purb	0,377	0,699	0,137	0,537	1,000										
des	0,257	0,100	0,366	-0,106	0,133	1,000									
trcol	0,215	-0,030	0,347	-0,031	-0,000	0,135	1,000								
asup	0,345	0,353	0,432	0,324	0,484	0,510	0,356	1,000							
denspop	0,447	0,562	0,207	0,145	0,435	0,088	0,439	0,393	1,000						
tempm	0,020	0,405	-0,400	0,136	0,283	0,009	-0,114	0,243	0,182	1,000					
idi	-0,342	-0,465	-0,293	-0,585	-0,685	-0,274	-0,174	-0,619	-0,285	-0,055	1,000				
br	0,611	0,688	0,157	0,062	0,404	0,110	0,346	0,349	0,676	0,231	-0,260	1,000			
af	0,305	0,331	0,093	-0,002	0,287	0,028	0,409	0,327	0,595	0,203	-0,184	0,694	1,000		
se	-0,397	-0,671	-0,207	-0,595	-0,770	-0,095	-0,070	-0,462	-0,449	-0,193	0,839	-0,432	-0,300	1,000	
ipss	0,544	0,651	0,111	0,097	0,317	0,047	0,231	0,176	0,520	0,063	-0,184	0,839	0,439	-0,333	1,000

Resultados do I de Moran local com duas variáveis

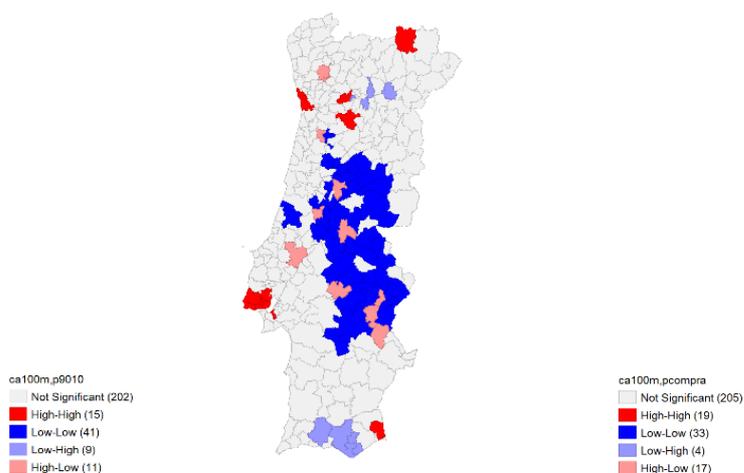


Figura A.1- I de Moran local com duas variáveis: *ca100m* e *p9010*

Figura A.2- I de Moran local com duas variáveis: *ca100m* e *pcompra*

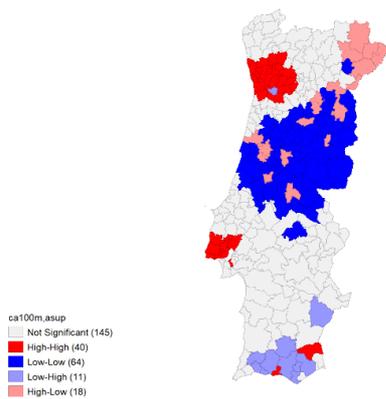


Figura A.3- I de Moran local com duas variáveis:
ca100m e *asup*

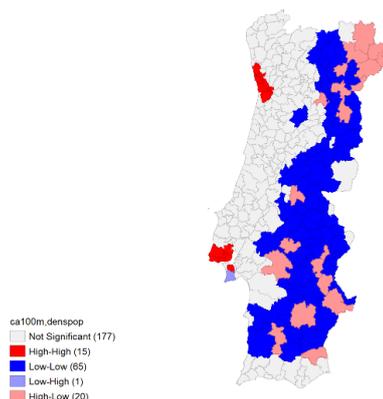


Figura A.4- I de Moran local de duas variáveis:
ca100m e *denspop*

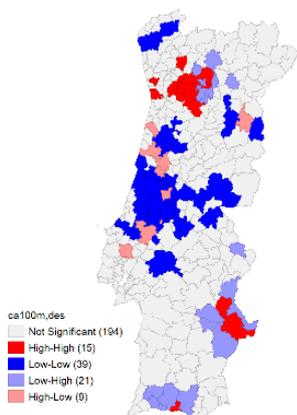


Figura A.5 – I de Moran local com duas variáveis:
ca100m e *des*

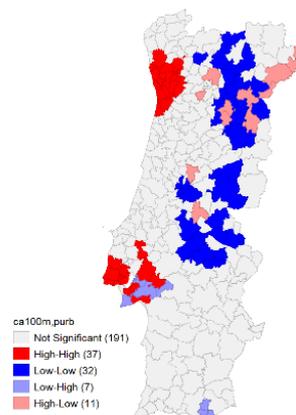


Figura A.6- I de Moran local com duas variáveis:
ca100m e *purb*

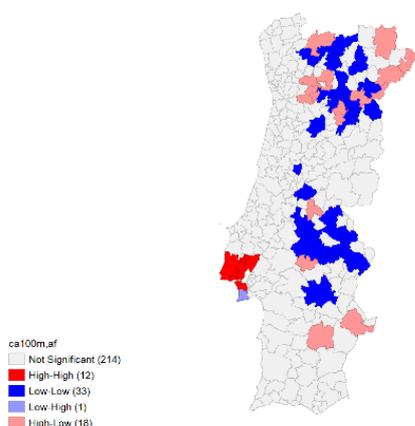


Figura A.7 – I de Moran local com duas variáveis:
ca100m e *af*

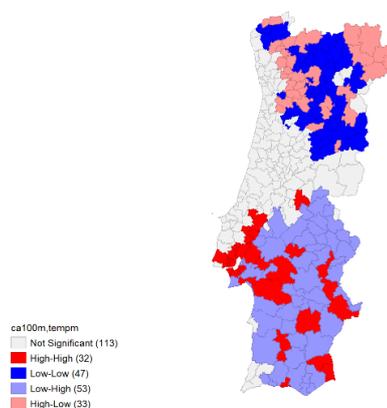


Figura A.8 – I de Moran local com duas variáveis:
ca100m e *tempm*

Resultados do I de Moran local e I de Moran local com duas variáveis da Área Metropolitana (AM) de Lisboa e Porto

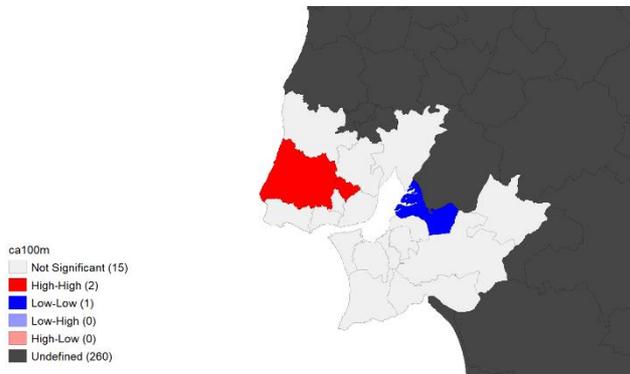


Figura A.9- I de Moran local de *ca100m* da AM de Lisboa

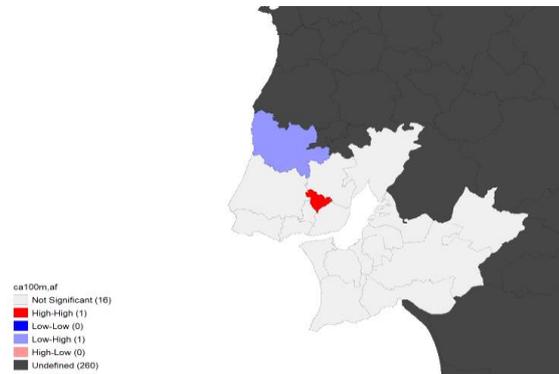


Figura A.10- I de Moran local com duas variáveis: *ca100m* e *af* da AM de Lisboa

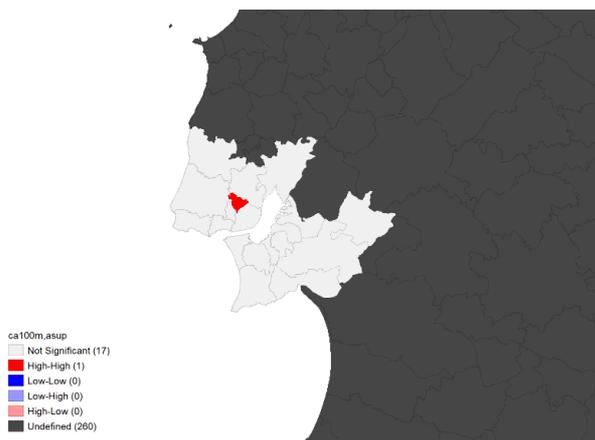


Figura A.11- I de Moran local com duas variáveis: *ca100m* e *asup* da AM de Lisboa

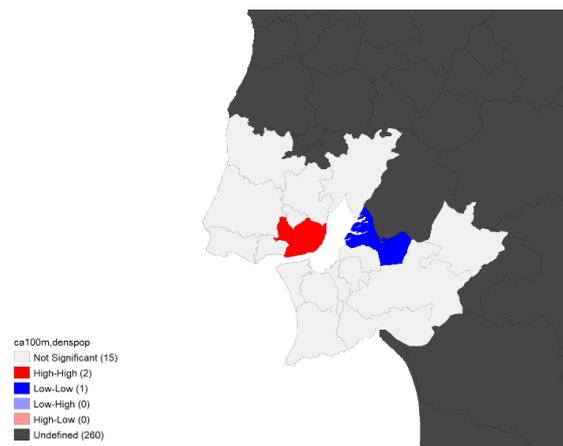


Figura A.12- I de Moran local com duas variáveis: *ca100m* e *denspop* da AM de Lisboa

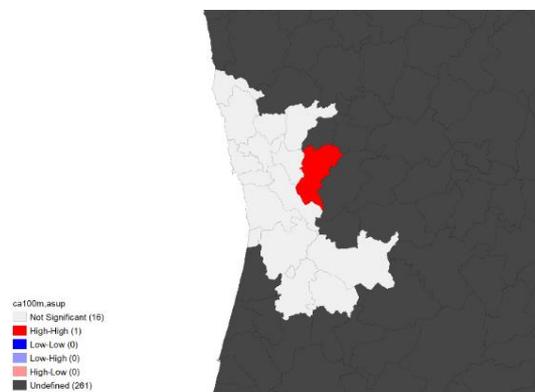


Figura A.13- I de Moran local com duas variáveis: *ca100m* e *asup* da AM do Porto