

MOBILIDADE E PROPAGAÇÃO DO SARS-COV-2 EM PORTUGAL CONTINENTAL – MODELO EXPLICATIVO TERRITORIALIZADO EM CONTEXTO ANTERIOR À VACINAÇÃO

Marques da Costa, Nuno ^{1*}; Mileu, Nelson ²; Marques da Costa, Eduarda ³; Alves, André ⁴; et al. ⁵

1 CEG-IGOT. ULisboa; nunocosta@campus.ul.pt

2 CEG-IGOT. ULisboa; nmileu@campus.ul.pt

3 CEG-IGOT. ULisboa; eduarda.costa@campus.ul.pt

4 IGOT. ULisboa; andrejoelalves@campus.ul.pt

5 IGOT - Freitas, Carlos; Rocha, J.; Abrantes, P.; Jonas, Eduardo; ENSP – Sousa, Paulo; Campos Fernandes, Adalberto; FFUL – Gaspar, Rogério; Duarte-Ramos, Filipa; FMUL – Portugal, Rui

*Autor correspondente: nunocosta@campus.ul.pt

Resumo: A difusão do novo coronavírus tem importantes associações com a mobilidade da população. Recorrendo a modelos de regressão linear múltipla, com informação epidemiológica da Direção-Geral da Saúde (DGS) e dados de mobilidade disponibilizados pela Google, captura-se a relação causal existente entre alterações na mobilidade e a tendência de incidência de COVID-19 para várias escalas em Portugal, demonstrando-se que maior mobilidade está associada a maior número de novos casos. Com base nesta premissa e recorrendo aos padrões de mobilidade da população é possível desenvolver um modelo de previsão do número de infeções futuras, com pelo menos 14 dias de antecedência.

Palavras-chave: Mobilidade; COVID-19; Modelo de propagação; Modelo preditivo

Abstract: The diffusion of the new coronavirus has important associations with population mobility. Using multiple linear regression models, with epidemiological information and mobility data from Google, a causal relationship between changes in mobility and the trend of COVID 19 incidence for several scales in Portugal has been identified, showing that greater mobility is associated with a higher number of new cases. Based on this premise and using information on mobility patterns of the population, it is possible to develop a model to predict the number of future infections, at least 14 days in advance.

Keywords: Mobility; COVID-19; Propagation model; Predictive Model

1. Introdução

A presente comunicação enquadra-se no projeto COMPRI_MOv - Conhecer Mais PaRa Intervir melhor no contexto da MObilidade (FCT-ID:613765655, Rede:IGOT-UL Coord.: FFUL; ENSP-UNL; ACES-Lx-Central; FM-UL; ALTICE-PT; AML; AMP; ACSS, <http://www.comprime-compri-mov.com/index.html>), projeto que teve como objetivo esboçar um modelo de simulação de propagação do vírus SARS-CoV-2 com base na mobilidade, tendo em conta a intensidade, motivações e padrão geográfico dos fluxos.

O processo pandémico, resultante da propagação do vírus SARS-CoV-2, colocou em evidência a necessidade de obtenção de informação relevante para a compreensão das tendências, processos e padrões de difusão no território por forma a apoiar a decisão em matéria de saúde pública com vista ao controlo da COVID-19. O lançamento do segundo período do “RESEARCH 4 COVID 19”, centrado na previsão da evolução futura da pandemia e do seu impacto nos serviços de saúde e na sociedade, bem

como na otimização da resposta adequada à pandemia a diversos níveis, levou a que fosse proposto o projeto COMPRI_MOv, “CONhecer Mais PaRa Intervir melhor no contexto da MObilidade” (ID: 613765655), na sequência do trabalho desenvolvido no projeto CONhecer Mais PaRa Intervir MELhor COMPRI (ID: 596685735), aprovado no primeiro período do “RESEARCH 4 COVID 19”.

A importância da mobilidade enquanto fator determinante da difusão de doenças transmissíveis é inegável, já que os movimentos da população são indispensáveis para ocorrer contacto social. Num mundo cada vez mais globalizado, a mobilidade é responsável por reduzir as diferenças epidemiológicas entre regiões no mundo (Gushulak & MacPherson, 2000), razão pela qual as restrições às deslocações são das primeiras medidas de contenção adotadas em contexto de surtos epidémicos para reduzir a transmissão e disseminação dos agentes patogénicos (Kraemer et al., 2020; Mateus et al., 2014).

A relevância de considerar a mobilidade em estudos epidemiológicos é comprovada pela boa performance de modelos de estimação da transmissão com base nos quantitativos dos movimentos da população, nomeadamente com informação “em tempo real”, por exemplo a partir da “geolocalização” de telemóveis, o que permite aferir não só a tendência como os seus padrões no território (Barlacchi et al., 2017; Engebretsen et al., 2020). Assim, os fluxos de mobilidade da população são “*epidemiologically informative*” (Jia et al., 2020).

No caso da COVID-19 vários estudos demonstram a indissociabilidade entre os padrões de mobilidade e a propagação territorial da doença (Hamidi et al., 2020; Jia et al., 2020; Severo et al., 2021; Zhang et al., 2021). O aumento da mobilidade apresenta uma relação forte com a variação positiva da incidência (Badr et al., 2020; Carteni et al., 2020) e a redução da mesma aquando da queda dos movimentos populacionais, devido a medidas de restrição às deslocações, comprova a importância da mobilidade (Kraemer et al., 2020). Além disso, o volume de mobilidade influencia a magnitude epidémica, com a não adesão da população ao recolhimento domiciliário a ser um fator para uma evolução descontrolada da pandemia (Yilmazkuday, 2020).

Alguns autores têm suportado a análise da relação entre a incidência de COVID-19 e a mobilidade com um desfasamento temporal entre as séries de dados. Ou seja, a mobilidade registada hoje apenas terá influência no número de casos futuros. Neste contexto, o *lag* temporal comumente adotado, de forma a albergar o período de incubação da doença e o tempo necessário para a testagem, reporte e comunicação oficial dos casos, é entre 7 e 28 dias (Badr et al., 2020; Carteni et al., 2020; Sulyok & Walker, 2020; Yilmazkuday, 2021).

Nesta sequência, a mobilidade constitui um bom preditor do número de casos de COVID-19, com vários autores a desenvolver metodologias de previsão do número de casos a partir da mobilidade da população. Alguns estudos têm sustentado as suas análises nos relatórios “*COVID-19 Community Mobility Reports*” da Google (Sulyok & Walker, 2020; Wang & Yamamoto, 2020; Yilmazkuday, 2021), o que é indicativo da qualidade da informação da Google enquanto representativa dos movimentos da população. Outros autores utilizam informação relativa à utilização de transportes públicos e de tráfego rodoviário (Carteni et al., 2020), de redes sociais como o Facebook (Ilin et al., 2021) e a partir da geolocalização de telemóveis (Guan et al., 2021).

2. Métodos

A metodologia desenvolvida tem por base a aplicação de regressões lineares múltiplas entre a incidência de COVID-19 e os dados de mobilidade da Google (Google, 2020), com um desfasamento temporal de 14 dias. Este valor foi selecionado com base em valores de referência sugeridos na literatura e em vários testes realizados, que permitiram concluir esta janela temporal como a mais adequada à realidade portuguesa.

Os procedimentos foram desenvolvidos em linguagem de programação *Python* para permitir maior celebridade na execução, já que é importante que o desfasamento temporal definido se ajuste sequencialmente com o avançar dos dias (Figura 1). Assim, utilizando os valores de mobilidade para blocos de 14 dias, é possível estimar o número de novos casos com duas semanas de antecedência.

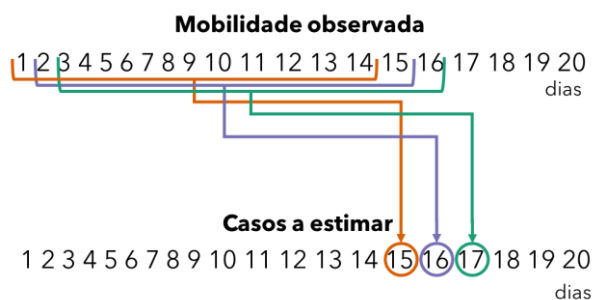


Figura 1 – Conceitualização das datas pelo modelo

Para garantir a execução da metodologia, desde a importação da informação, à manipulação de tabelas, à análise de regressão até aos valores estimados de *output*, foram utilizados os seguintes módulos em *Python*: *pandas* (<https://pandas.pydata.org/>), *scikit-learn* (<https://scikit-learn.org/stable/>), *statsmodels* (<https://www.statsmodels.org/stable/index.html>) e *numpy* (<https://numpy.org/>).

Variável dependente

A variável dependente é o número de novos casos confirmados diariamente. Esta informação epidemiológica foi recolhida dos relatórios diários da Direção-Geral da Saúde (DGS, 2020). O modo de disponibilização desta informação constitui uma limitação ao seu uso, já que tem de ser recolhida manualmente a partir dos relatórios, suscetível a erros de introdução, e requer a edição manual da informação previamente a qualquer tipo de processo de análise e modelação. A disponibilização da informação pela DGS não constitui o melhor exemplo de uma política de comunicação e de partilha de dados elementares, quando comparada com outras autoridades de saúde na Europa. Exemplo de excelência é o repositório de dados da Proteção Civil Italiana (<https://github.com/pcm-dpc/COVID-19>) que disponibiliza vários indicadores epidemiológicos em formatos de dados manipuláveis e prontos a analisar, assim como integráveis em sistemas de informação geográfica (CSV/Excel/JSON). Além disso, devido a alterações no Sistema Nacional de Vigilância Epidemiológica (SINAVE), a informação dos casos de COVID-19 relativa aos municípios deixou de ser fornecida diariamente, e mais tarde semanalmente, como o número acumulado de casos confirmados e foi substituída pela incidência cumulativa a 14 dias por 100 mil habitantes. O novo formato de disponibilização ocorre semanalmente, contudo as séries de 14 dias originam sobreposições de informação que não podem ser eliminadas e obrigam a cálculos adicionais para obter apenas o número aproximado de novos casos diários. Outras limitações relativamente à qualidade, resolução espacial, inconsistências e sincronização dos dados da DGS são discutidas em Marques da Costa et al. (2021).

Variáveis independentes

As seis variáveis independentes provêm do repositório da Google com dados de mobilidade (<https://www.google.com/covid19/mobility/>). Esta informação, disponível para diferentes escalas (nacional, regional, municipal) de diferentes países representa a percentagem de variação na mobilidade, baseada na mediana das primeiras 5 semanas de 2020 (3 de janeiro a 6 de fevereiro), consideradas como referência dos valores de mobilidade pré-pandémica. Esta informação é recolhida pela Google a partir de utilizadores que permitem o histórico de localização em aplicações de tecnologia Google, sendo, posteriormente, a informação agregada e anonimizada.

A percentagem de variação é determinada pelas visitas e período de estadia em diferentes tipos de atividades como *retail & recreation*, *grocery & pharmacy*, *parks*, *transit stations*, *workplaces* e *residential*. Estas 6 categorias foram utilizadas como variáveis independentes no modelo, avaliando-se de que forma a sua variação explica a variabilidade da incidência.

No entanto, a informação para alguns distritos e municípios apresenta lacunas, com ausência de valores para uma ou mais categorias de locais em algumas datas. Assim, sempre que possível estas ausências foram recalculadas garantindo a continuidade da série. Nos casos em que as falhas eram mais generalizadas o processo de modelação não pôde ser aplicado, pois não existe forma minimamente

fiável de as reconstruir.

3. Resultados e discussão

A associação entre as variáveis de mobilidade e a ocorrência de novos casos de COVID-19 é comprovada pela forte relação entre os dados reais e os estimados, de acordo com os valores de alteração de mobilidade dos 14 dias anteriores (Figura 2a). Em momentos de maior mobilidade da população, ou seja, menor permanência na residência, verifica-se o aumento do número de casos de COVID-19. A estimativa do número de novos casos a partir da mobilidade acompanha a tendência de facto registada. A validação com base nos coeficientes de determinação reforçam o ajustamento do modelo à realidade e a importância que este tipo de informação em “tempo real” tem para a estimativa do processo pandémico. Os resultados estão em linha com a literatura já referida anteriormente e a adequabilidade dos modelos ao contexto português validam não só a importância que a recolha e disponibilização de informação acerca das deslocações da população pode ter, mas também são uma validação da qualidade da mesma.

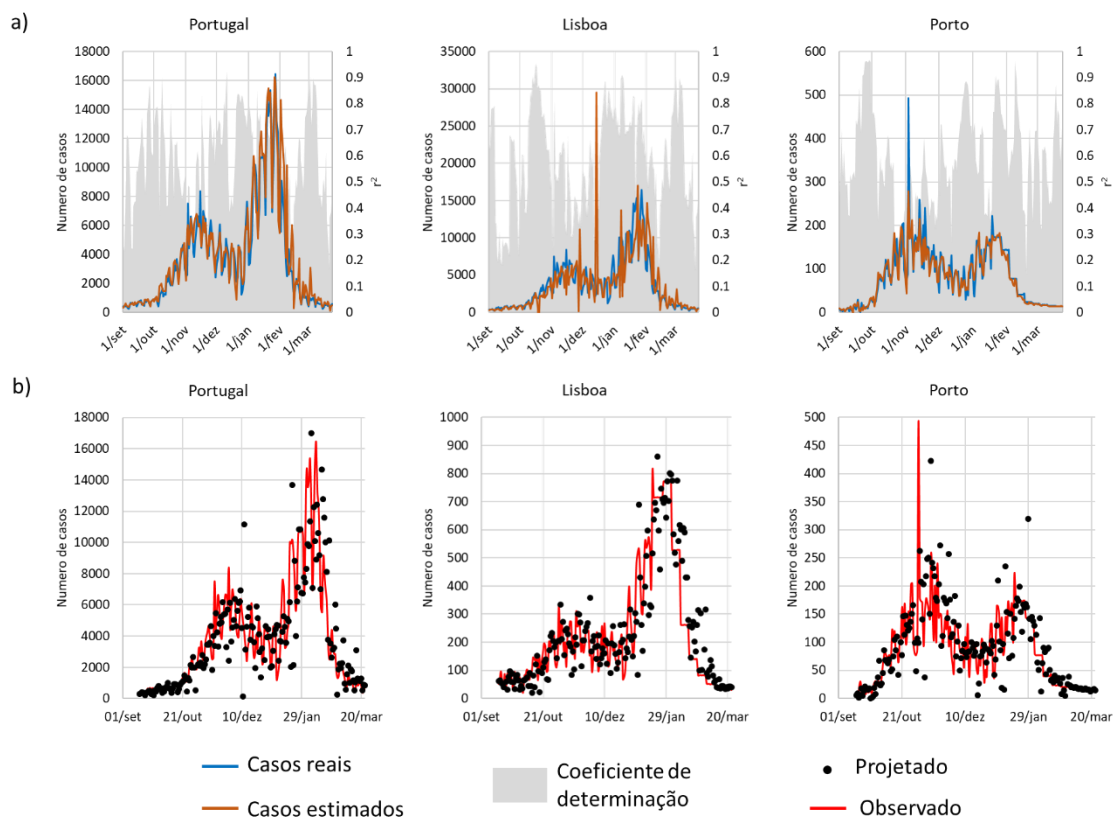


Figura 2. Resultados do modelo: (a) Estimativa e validação; (b) Projeção.

Os resultados variam de acordo com o grau de agregação dos dados e a dimensão das unidades territoriais, com melhores ajustamentos para o município do Porto (Figura 2a).

Relativamente ao método de predição para os 14 dias seguintes (Figura 2b), os modelos revelam maior dificuldade a ajustarem-se à realidade, todavia os resultados são significativos e acompanham a tendência de evolução, excetuando-se momentos onde a incidência atingiu valores mais extremos, permitindo um grau de conhecimento suficiente sobre a evolução futura da pandemia essencial para a sua gestão no que concerne ao desenho e execução de medidas de saúde pública.

Importa referir que os procedimentos executados se referem a uma situação onde a cobertura vacinal era ainda reduzida em Portugal. A posterior aplicação da mesma metodologia num contexto de vacinação completa em grande parte da população poderá ser indicativo da eficácia do processo vacinal

se o ajustamento dos modelos for menor, ou seja, se existir sobrestimação do número de novos casos. Esta comparação parte também do princípio de que o período de 14 dias se mantém adequado à realidade perante possíveis alterações na gestão e disponibilização da informação por parte da DGS.

4. Conclusões

O forte ajustamento entre os dados de mobilidade e os epidemiológicos, com um desfasamento de 14 dias, permitem aferir a boa performance da metodologia apresentada enquanto modelo de previsão da propagação da COVID-19 com base na mobilidade da população. A forte relação verificada evidencia que o aumento da mobilidade da população contribuiu para o crescimento da variação do número de novos casos.

Assim, não só é novamente comprovada a importância da mobilidade enquanto fator determinante para a transmissão de doenças infecciosas, especificamente para o caso da transmissão do SARS-CoV-2, como a informação acerca do grau e local de deslocação da população tem uma grande importância para o apoio à decisão em contextos epidémicos. A incorporação de metodologias semelhantes com capacidade preditiva nos sistemas de vigilância epidemiológica é crucial para o controlo da pandemia e fenómenos epidémicos futuros.

Por fim, a qualidade e granularidade da informação, tanto epidemiológica como de mobilidade, são propriedades fundamentais para a execução generalizada da metodologia e, no futuro, a qualidade da mesma, em termos de resolução espacial, modo de divulgação e periodicidade, deve constituir uma preocupação para que se consigam alcançar melhores resultados e evidências para o estudo de fenómenos epidémicos.

Bibliografia

- Badr, H. S., Du, H., Marshall, M., Dong, E., Squire, M. M., & Gardner, L. M. (2020). Association between mobility patterns and COVID-19 transmission in the USA: a mathematical modelling study. *The Lancet Infectious Diseases*, 20(11), 1247–1254. [https://doi.org/10.1016/S1473-3099\(20\)30553-3](https://doi.org/10.1016/S1473-3099(20)30553-3)
- Barlacchi, G., Perentis, C., Mehrotra, A., Musolesi, M., & Lepri, B. (2017). Are you getting sick? Predicting influenza-like symptoms using human mobility behaviors. *EPJ Data Science*, 6(1), 27. <https://doi.org/10.1140/epjds/s13688-017-0124-6>
- Carteni, A., Di Francesco, L., & Martino, M. (2020). How mobility habits influenced the spread of the COVID-19 pandemic: Results from the Italian case study. *Science of the Total Environment*, 741, 140489. <https://doi.org/10.1016/j.scitotenv.2020.140489>
- DGS - Direção-Geral da Saúde. (2020). *Relatório de Situação - COVID-19*. <https://covid19.min-saude.pt/relatorio-de-situacao/>
- Engelbrechtsen, S., Engø-Monsen, K., Aleem, M. A., Gurley, E. S., Frigessi, A., & de Blasio, B. F. (2020). Time-aggregated mobile phone mobility data are sufficient for modelling influenza spread: the case of Bangladesh. *Journal of The Royal Society Interface*, 17(167), 20190809. <https://doi.org/10.1098/rsif.2019.0809>
- Google. (2020). *Relatórios de mobilidade da comunidade da COVID-19*. <https://www.google.com/covid19/mobility/>
- Guan, G., Dery, Y., Yechezkel, M., Ben-Gal, I., Yamin, D., & Brandeau, M. L. (2021). Early detection of COVID-19 outbreaks using human mobility data. *PLOS ONE*, 16(7), e0253865. <https://doi.org/10.1371/journal.pone.0253865>
- Gushulak, B. D., & MacPherson, D. W. (2000). Population Mobility and Infectious Diseases: The Diminishing Impact of Classical Infectious Diseases and New Approaches for the 21st Century. *Clinical Infectious Diseases*, 31(3), 776–780. <https://doi.org/10.1086/313998>
- 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>
- Ilin, C., Annan-Phan, S., Tai, X. H., Mehra, S., Hsiang, S., & Blumenstock, J. E. (2021). Public mobility data enables COVID-19 forecasting and management at local and global scales. *Scientific Reports*, 11(1), 13531. <https://doi.org/10.1038/s41598-021-92892-8>

- Jia, J. S., Lu, X., Yuan, Y., Xu, G., Jia, J., & Christakis, N. A. (2020). Population flow drives spatio-temporal distribution of COVID-19 in China. *Nature*, 582(7812), 389–394. <https://doi.org/10.1038/s41586-020-2284-y>
- Kraemer, M. U. G., Yang, C.-H., Gutierrez, B., Wu, C.-H., Klein, B., Pigott, D. M., Covid-19 Data, O., Group, W., Du Plessis, L., Faria, N. R., Li, R., Hanage, W. P., Brownstein, J. S., Layan, M., Vespignani, A., Tian, H., Dye, C., Pybus, O. G., & Scarpino, S. V. (2020). *The effect of human mobility and control measures on the COVID-19 epidemic in China*. <http://science.sciencemag.org/>
- Marques da Costa, N., Mileu, N., & Alves, A. (2021). Dashboard comprime_compri_mov: Multiscalar spatio-temporal monitoring of the covid-19 pandemic in Portugal. *Future Internet*, 13(2), 1–17. <https://doi.org/10.3390/fi13020045>
- Mateus, A. L. P., Otete, H. E., Beck, C. R., Dolan, G. P., & Nguyen-Van-Tam, J. S. (2014). Effectiveness of travel restrictions in the rapid containment of human influenza: a systematic review. *Bulletin of the World Health Organization*, 92(12), 868–880D. <https://doi.org/10.2471/BLT.14.135590>
- Severo, M., Ribeiro, A. I., Lucas, R., Leão, T., & Barros, H. (2021). Urban Rail Transportation and SARS-Cov-2 Infections: An Ecological Study in the Lisbon Metropolitan Area. *Frontiers in Public Health*, 9(February). <https://doi.org/10.3389/fpubh.2021.611565>
- Sulyok, M., & Walker, M. (2020). Community movement and covid-19: A global study using google's community mobility reports. *Epidemiology and Infection*. <https://doi.org/10.1017/S0950268820002757>
- Wang, H., & Yamamoto, N. (2020). Using a partial differential equation with Google Mobility data to predict COVID-19 in Arizona. *Mathematical Biosciences and Engineering*, 17(5), 4891–4904. <https://doi.org/10.3934/mbe.2020266>
- Yilmazkuday, H. (2020). COVID-19 spread and inter-county travel: Daily evidence from the U.S. *Transportation Research Interdisciplinary Perspectives*, 8, 100244. <https://doi.org/10.1016/j.trip.2020.100244>
- Yilmazkuday, H. (2021). Stay-at-home works to fight against COVID-19: International evidence from Google mobility data. *Journal of Human Behavior in the Social Environment*, 00(00), 1–11. <https://doi.org/10.1080/10911359.2020.1845903>
- Zhang, J., Feng, B., Wu, Y., Xu, P., Ke, R., & Dong, N. (2021). The effect of human mobility and control measures on traffic safety during COVID-19 pandemic. *PLoS ONE*, 16(3 March 2021), 493–497. <https://doi.org/10.1371/journal.pone.0243263>