Modelling Mobility Data during COVID-19 in Portugal with `R-INLA`

André Brito [1,2,4] Ausenda Machado [4,5] Ana Paula Rodrigues [4] Paula Patricio [1,2] Regina Bispo [1,3]

[1]Center for Mathematics and Applications (NOVA Math), NOVA School of Science and Technology (NOVA FCT), Portugal
[2]Department of Mathematics, NOVA School of Science and Technology (NOVA FCT), Portugal
[3]School of Mathematics and Statistics and Centre for Research into Ecological and Environmental Modelling University of St Andrews, UK
[4]Department of Epidemiology, National Institute of Health Doctor Ricardo Jorge, Portugal
[5]Comprehensive Health Research Center (CHRC), NOVA University Lisbon, Portugal

Background & Objective

The COVID-19 pandemic underscored the vital role of mathematical modelling in epidemiology, aiding in forecasting disease trends and informing public health strategies. In Portugal, models demonstrated the effectiveness of physical distancing in curbing transmission [1], and provided key insights into the impacts of deconfinement and vaccination on disease burden [2]. The period also saw increased use of non-traditional data, notably human mobility data, which served as a proxy for contact patterns and helped refine models of disease spread [3, 4].

We aim to model mobility patterns across districts in Portugal during the COVID-19 pandemic using the Google Mobility Reports. Our objective is to identify which covariates are relevant in predicting mobility dynamics during this period using a spatio-temporal modelling framework.

Data

Google Mobility Reports

The reports provide data on the percentage change in mobility relative to a baseline for different location categories. We considered Retail and Recreation, Transit Stations and Workplaces. The data is available from February 15th 2020 to October 15th 2022 and grouped by district. Regrettably, while the data remains accessible, updates have ceased.

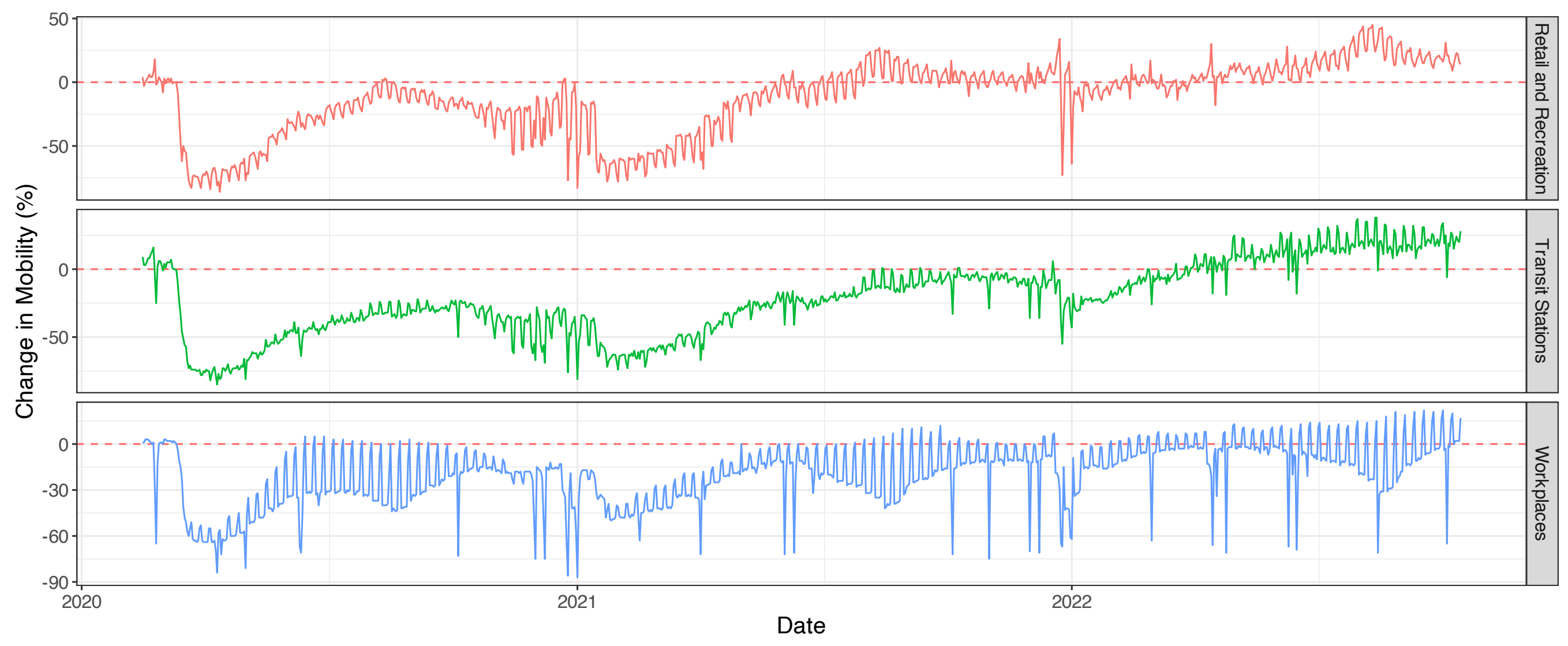

Figure 1. Overall mobility change per category in Portugal from Feb. 15th 2020 to Oct. 15th 2022.

COVID-19 Government Response Stringency Index

The Stringency Index was developed by the University of Oxford [5] and is a composite measure of response metrics, among them, school and workplace closures, public event cancellations, stay-at-home requirements, scored from 0 to 100.

Copernicus Temperature Data

We used daily gridded land-only observational temperature data from the Copernicus Climate Data Store. The dataset provides mean near-surface temperatures. We computed daily time series of averaged mean temperatures for each district. This process transforms point-based gridded data into spatially aggregated raster data.

Model

We modelled daily mobility in mainland Portugal across three categories—Workplaces (C_1), Retail and Recreation (C_2), and Transit Stations (C_3)—by district, using a spatio-temporal model to capture variations over time and space. It assumes mobility follows a normal distribution based on covariates and parameters.

$$
\begin{aligned}
Y_{j\,t\,c} = \beta_c + \texttt{temporal}_{j\,t\,c} + \texttt{lockdown}_{t\,c} + \texttt{stringency}_t + \texttt{temperature}_{j\,t} + \\
u_j + v_j + \gamma_{t\,|\,c} + \phi_{t\,|\,c} + \delta_{j\,t\,|\,c}
\end{aligned}
$$

where $Y_{j\,t\,c}$ is the percentage change in mobility for district j $(j = 1, \ldots, 18)$ on day t for category c $(c = C_1, C_2, C_3)$.

- β_c includes category fixed effect terms, $\beta_c = \sum_{i=1}^{3} \beta_{C_i} \mathbb{I}\,(c = C_i\,)$ where $\mathbb{I}\,(c = C_i\,)$ is the indicator function that takes the value 1 if c is the category C_i, and 0 otherwise.
- the $\texttt{temporal}_{j\,t\,c}$ covariates include a linear temporal term, two Fourier pairs with yearly and half-year frequencies to account for seasonal patterns for each category, a day of the week indicator variable, coded as weekday (d_1), Saturday (d_2) or Sunday (d_3), combined with a district- and category-specific random effect on day of the week. It also includes indicator variables to flag `national`, `commemorative` and `regional` holidays.
- the $\texttt{lockdown}_{t\,c} = \left(\psi + \tilde{\psi}_c\right) \mathbb{I}(t \in L)$ term is an indicator variable to identify the periods under full lockdown. The term also includes category-specific random effects, with L being the time intervals from March 18th to May 2nd 2020 and from January 14th to March 14th 2021.
- the $\texttt{stringency}_t = \zeta\, \texttt{stringency}_t$ includes the Stringency Index variable.
- the $\texttt{temperature}_{j\,t} = (\omega + \tilde{\omega}_j)\ \texttt{temperature}_{j\,t}$ term includes the daily mean temperature computed for the district j at day t. It includes a fixed effect combined with a district-specific random effects.
- $u_j + v_j$ constitutes the Besag–York–Molliè (BYM) model.
- $\gamma_{t\,|\,c}$ represents an autocorrelated random time effect given a category of mobility c.
- $\phi_{t\,|\,c}$ represents an uncorrelated time dependence.
- $\delta_{j\,t\,|\,c}$ is defined as a Type I space-time interaction. conditional to a category c.

All the fixed effect terms have been assigned with normal prior distributions $N\,(0, 100^2)$ and all the random effects were assigned with a default Gamma prior distribution for the precision parameter $\tau_{(.)} \sim \text{Gamma}\,(1, 10^{-3})$.

Results & Discussion

The training data included two complete years of data (Feb. 15th 2020 - Feb. 14th 2022) while the testing data comprised the last six months (Feb. 15th 2022 - Oct. 15th 2022), representing approximately 25% of the total dataset. Model estimation was done using `R-INLA`. For the spatial structure assumed that districts are neighbours, *i.e.* adjacent to each other, if their borders share any point of contact. To simplify the presentation of the results, Figure 2 shows the estimated linear and seasonal trends and Table 1 presents the posterior estimates and 95% [Q 2.5% - Q 97.5%] Bayesian Credible Intervals (BCI) for all the covariates' coefficients included in the model.

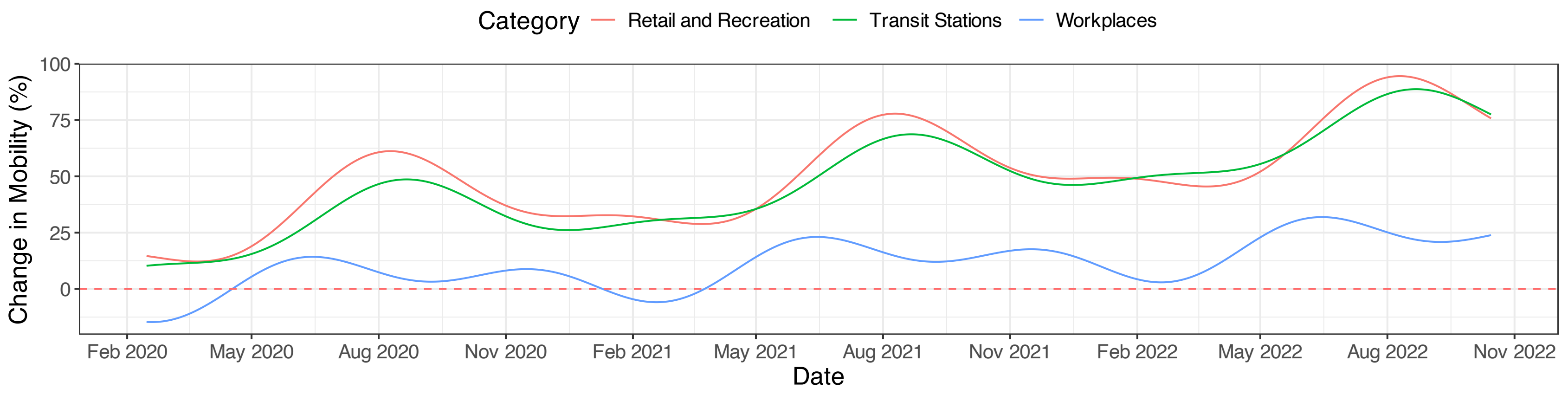

Figure 2. Time series of the estimated mobility seasonal patterns, includes a combination of a linear trend and two Fourier pairs with yearly and half-year frequencies, respectively, per category.

Table 1. Posterior estimates for the model, mean, standard deviation (SD), quantiles at 2.5%, 50% and 97.5%, and mode.

Covariate	Coefficient	Mean	SD	Q 2.5%	Q 50%	Q 97.5%	Mode
`Workplaces`	β_{C_1}	-15.76	5.46	-26.56	-15.73	-5.14	-15.73
`Retail and Recreation`	β_{C_2}	22.38	5.39	11.84	22.36	33.03	22.36
`Transit Stations`	β_{C_3}	15.52	5.28	5.26	15.49	25.97	15.49
`Saturdays vs. weekdays`	$\eta_{d_2\,vs.\,d_1}$	-0.30	1.19	-2.64	-0.30	2.05	-0.30
`Sundays vs. weekdays`	$\eta_{d_3\,vs.\,d_1}$	-0.34	2.22	-4.71	-0.34	4.02	-0.34
`national holiday`	θ_1	-29.47	1.25	-31.92	-29.47	-27.01	-29.47
`regional holiday`	θ_2	-6.78	1.35	-9.44	-6.78	-4.13	-6.78
`commemorative holiday`	θ_3	-12.99	2.01	-16.94	-12.99	-9.05	-12.99
`lockdown`	ψ	-12.02	2.77	-17.49	-12.02	-6.56	-12.02
`stringency`	ζ	-0.77	0.03	-0.83	-0.77	-0.72	-0.77
`temperature`	ω	0.68	0.17	0.35	0.68	1.02	0.68

- Different mobility categories, such as Retail, Workplaces, and Parks, demonstrated distinct seasonal patterns and justifying the category-specific modelling terms.
- Saturdays and Sundays did not significantly affect mobility compared to weekdays. However, holidays had a strong impact across all categories. National holidays caused the largest drop in mobility (−29.47%), followed by commemorative holidays (−12.99%) and regional holidays (−6.78%).
- The definition of a lockdown variable was crucial to fully capture the impact of complete lockdown restrictions on mobility. The estimated mean effect corresponds to a mobility decrease of approximately 12% across all categories.
- The link between mobility and the stringency index is small but statistically significant. On average, each one-point increase in stringency reduces mobility by about 0.77%. Since the index usually ranged between 50 and 85 during the pandemic — a 35-point swing — the maximum expected drop in mobility would be around 27%. This shows that while stringency measures did affect mobility, it only explains part of the overall changes. Especially during strict lockdowns, producing sharp declines in mobility without corresponding changes in the index.
- Temperature has a positive and significant effect on mobility, as higher temperatures are linked to more movement, even after accounting for other factors. However, the effect is fairly small: each 1°C increase leads to about a 0.68% rise in mobility. For instance, a 20°C rise - going from winter to summer - would increase mobility by roughly 14%. This means temperature does influence mobility, but its impact is moderate compared to other factors.
- A key issue identified in the analysis was the construction of the response variable. Google's mobility data is based on percentage changes from a fixed baseline set in early 2020 — a short five-week window in winter. This baseline is not representative of typical annual mobility, making it difficult to interpret changes over time. A more suitable baseline, such as an annual average from pre-pandemic years, would have allowed for clearer and more meaningful comparisons. As it stands, interpreting what it means for mobility to be, for instance, "150% higher in August 2022 compared to January 2020" is inherently ambiguous.

References

[1] Constantino Caetano, Maria Luísa Morgado, Paula Patrício, João F. Pereira, and Baltazar Nunes.
Mathematical modelling of the impact of non-pharmacological strategies to control the covid-19 epidemic in portugal.
Mathematics, 9, 5 2021.

[2] João Viana, Christiaan H. van Dorp, Ana Nunes, Manuel C. Gomes, Michiel van Boven, Mirjam E. Kretzschmar, Marc Veldhoen, and Ganna Rozhnova.
Controlling the pandemic during the sars-cov-2 vaccination rollout.
Nature Communications, 12, 12 2021.

[3] Kaylin Bolt, Diana Gil-González, and Nuria Oliver.
Unconventional data, unprecedented insights: leveraging non-traditional data during a pandemic.
Frontiers in Public Health, 12, 3 2024.

[4] Jeffrey E. Harris.
Mobility was a significant determinant of reported covid-19 incidence during the omicron surge in the most populous u.s. counties.
BMC Infectious Diseases, 22, 12 2022.

[5] Thomas Hale, Noam Angrist, Rafael Goldszmidt, Beatriz Kira, Anna Petherick, Toby Phillips, Samuel Webster, Emily Cameron-Blake, Laura Hallas, Saptarshi Majumdar, and Helen Tatlow.
Variation in government responses to covid-19.
(BSG-WP-2020/032), 2023.
Version 15, updated June 2023.

This work is funded by national funds through the FCT – Fundação para a Ciência e a Tecnologia, I.P., under the scope of the projects UIDB/00297/2020 (https://doi.org/10.54499/UIDB/00297/2020), UIDP/00297/2020 (https://doi.org/10.54499/UIDP/00297/2020) (Center for Mathematics and Applications) and 2024.00664.BDANA (PhD Scholarship)

NOVA Math, INSA

INLA Workshop | University of Glasgow

anm.brito@campus.fct.unl.pt