medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21250113; this version posted February 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Association between COVID-19, mobility and environment in São Paulo, Brazil

Authors and affiliation

Sergio Ibarra-Espinosa
Departamento de Ciências Atmosféricas, Instituto de Astronomia, Geofísica e Ciências Atmosféricas, Universidade de São Paulo, Brazil

Edmilson Dias de Freitas
Departamento de Ciências Atmosféricas, Instituto de Astronomia, Geofísica e Ciências Atmosféricas, Universidade de São Paulo, Brazil

Karl Ropkins
Institute for Transport Studies, University of Leeds, UK

Francesca Dominici
Harvard Data Science Initiative, Harvard University, Boston, MA 02138, USA

Amanda Rehbein
Departamento de Ciências Atmosféricas, Instituto de Astronomia, Geofísica e Ciências Atmosféricas, Universidade de São Paulo, Brazil

Corresponding author

Sergio Ibarra-Espinosa
Departamento de Ciências Atmosféricas, Instituto de Astronomia, Geofísica e Ciências Atmosféricas, Universidade de São Paulo, Brazil
Phone +55 (11) 934 909 778
E-mail: sergio.ibarra@usp.br

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21250113; this version posted February 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Association between COVID-19, mobility and environment in São Paulo, Brazil

Sergio Ibarra-Espinosa[1], Edmilson Dias de Freitas[1], Karl Ropkins[2], Francesca Dominici[3], Amanda Rehbein[1]

[1] Departamento de Ciências Atmosféricas, Instituto de Astronomia, Geofísica e Ciências Atmosféricas, Universidade de São Paulo, Brazil

[2] Institute for Transport Studies, University of Leeds, UK

[3] Harvard Data Science Initiative, Harvard University, Boston, MA 02138, USA

Author correspondence: sergio.ibarra@usp.br

Abstract

Background

Brazil, the country most impacted by the coronavirus disease 2019 (COVID-19) on the southern hemisphere, uses mobility indices to monitor quarantines. In this study we quantified the associations between residential mobility index (RMI), air pollution, meteorology, and daily cases and deaths of COVID-19 in São Paulo, Brazil

Objectives

To acquire time-series data to estimate the associations between daily residential mobility index (RMI), air pollution, and meteorology, and daily cases and deaths for COVID-19 in São Paulo, Brazil.

Methods

We applied a semiparametric generalized additive model (GAM) to estimate: 1) the association between residential mobility index and cases and deaths due to COVID-19, accounting for ambient particulate matter ($PM_{2.5}$), ozone (O_3), relative humidity, temperature and delayed exposure between 3-21 days and 2) the association between exposure to for ambient particulate matter ($PM_{2.5}$), ozone (O_3), accounting for relative humidity, temperature and mobility.

Results

We found an RMI of 45.28% results in 1,212 cases (95% CI: 1,189 to 1,235) and 44 deaths (95% CI: 40 to 47). Reducing mobility 5% would avoid 438 cases and 21 deaths. Also, we found that an increment of 10 $\mu g \cdot m^{-3}$ of $PM_{2.5}$ risk of 1.140 (95% CI: 1.021 to 1.274) for cases and of 1.086 (95% CI: 1.008 to 1.170) for deaths, while O_3 produces

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21250113; this version posted February 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

a relative risk of 1.075 (95% CI: 1.006 to 1.150) for cases and 1.063 (95% CI: 1.006 to 1.124) for deaths, respectively.

Discussion

We compared our results with observations and literature review, finding well agreement. These results implicate that authorities and policymakers can use such mobility indices as tools to support social distance activities and assess their effectiveness in the coming weeks and months. Small increments of air pollution pose a risk of COVID-19 cases.

Conclusion

Spatial distancing is a determinant factor to control cases and deaths for COVID-19. Small increments of air pollution result in a high number of COVID-19 cases and deaths. $PM_{2.5}$ has higher relative risks for COVID-19 than O_3.

Keywords: COVID-19, SARS-CoV 2, residential mobility index, air pollution, meteorology

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21250113; this version posted February 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Introduction

The world has been facing an unprecedent critical health crisis due to COVID-19 pandemic caused by the zoonotic severe acute respiratory syndrome coronavirus 2 (SARS-CoV2). According to the Johns Hopkins University (JHU) Center for Systems Science and Engineering (CSSE), on January 04, 2021, there were 85,510,573 confirmed cases of COVID-19 worldwide, and 1,849,054 deaths (Dong et al. 2020). Many countries implemented social isolation and quarantine strategies, and internet companies released aggregated location data to provide information on the effectiveness of quarantine and isolation (Google 2021). This new disease demonstrated that developed countries such as Belgium, Italy, and Spain were unprepared, which resulted in infection fatality rates (deaths/cases) of 15.4%, 13.5%, and 10.2%, respectively. Furthermore, new mutations of SARS-CoV2 would present higher transmissibility and current vaccines might not offer protection against it (Gupta et al. 2021; Kupferschmidt 2021). Therefore, it is urgent and crucial to conduct more research to better understand the relationships and associations between SARS-CoV2 transmissibility and environmental factors.

COVID-19 spread very quickly across Latin America (Bolaño-Ortiz et al. 2020). As of January 4, 2021, Brazil is the third country with more cases (7,733,746) and deaths (196,018) (Dong et al. 2020). Brazil is a continental country with an area of 8.5 million km² and a population of 208 million. The city of São Paulo is the highest populated city in Brazil, with 11.8 million people (IBGE 2014), and also, with the highest number of COVID-19 cases (404,025) and deaths (15,725 (https://covid.saude.gov.br/). In Brazil, only symptomatic cases are tested, hence, the real number of SARS-CoV2 infections could be much higher. A recent study shows that COVID-19 death notification in Brazil is underreported (Alves et al. 2020). On March 24, 2020 (GESP 2020), São Paulo's government recommended social distancing for suspected cases and introduced a local quarantine to reduce virus transmission. Brazil adopted containment measures such as close contact and limited mobility as protective measures, however "quarantine" was the official term used by São Paulo state government. During the São Paulo quarantine, the concentrations of carbon monoxide (CO), nitrogen dioxide (NO_2), $PM_{10,}$ and $PM_{2.5}$ were reduced by -35.70%, -29.56%, -17.80%, and -25.02%, respectively. However, O_3 increased by 53.25% between 21:00-03:00 Local Time (LT) (Dantas et al. 2020; Freitas et al. 2020; Nakada and Urban 2020). Other studies showed reductions in China (30% of NO_2 and 25% of CO_2) (Dutheil et al. 2020). However, there are indications that COVID-19 and air pollution interactions may be more complex. Some early reports suggest that longer-term exposure to air pollution increases susceptibility and severity upon infection (Tosepu et al. 2020; Wu et al. 2020). Similarly, meteorological

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21250113; this version posted February 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

conditions are likely to be important but not easily predicted. For example, some studies have suggested that high humidity and temperature would reduce virus transmission (Wang et al. 2020a, 2020b). Furthermore, the biological inactivation rate of SARS-CoV2 is sensitive to humidity, with minimum half-life at 65% and higher at 40% and 85%, and its decays at 27 □C is 10 times faster than at 10 □C (Morris et al. 2020). However, estimations for the Amazonian city of Manaus, which have a climatology[1] temperature 27 □C of and relative humidity 80.2% for July, show that between 44% and 66% of the population was infected with SARS-CoV-2 by July (Buss et al. 2020).

Few studies have evaluated the effect of mobility on COVID-19, mainly in China. Kraemer et al., (2020) found correlation of 0.94 between real-time mobility data and COVID-19 cases from Wuhan, China, confirming the exportation of cases from Wuhan to other provinces and the effectiveness of the sanitaire cordon. Tian et al (2020) found that Wuhan shutdown delayed the appearance of COVID-19 in 2.9 days. In the US, Badr et al (2020) associated county-level origin-destination matrices and COVID-19 cases finding that counties with more mobility presented higher number of cases. In a related correspondence, Gatalo et al (2020) found that the absence of a strong correlation between growth of cases and mobility may be related to other factors, such as wearing mask, keeping distance in encounters and also, the existence of superspreading events.

Our study investigates the associations between RMIand COVID-19, and air pollution and COVID-19 on a particular day, having accounted for environmental and meteorological factors. We analyzed the effects of mobility, air pollution, and meteorology on the daily COVID-19 cases and deaths in the city of São Paulo. We applied semiparametric generalized additive models to study the effect of each predictor by isolating confounding factors (Peng and Dominici 2008). To the best of our knowledge, few stutudies have associated mobility air pollution, meteorological factors, and COVID-19 simultaneously especially in Latin Ameirca.

Material and Methods

The Brazilian Ministry of Health reports the official daily time-series of cases and mortality associated with COVID-19 at https://covid.saude.gov.br/ updated once a day around 19:00 Brazilian official time (-3 GMT). This information is gathered from the 26 states and the federal district health secretaries and provided at the national, state, and municipality levels. The availability of data using the web site was interrupted on June 07, 2020, and

[1] Climatology between 1981-2010 https://portal.inmet.gov.br/

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21250113; this version posted February 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

returned on June 08, 2020 after judicial demand[1]. When returned, it included the data obtained when the web site was offline. To check the trustworthiness of that data, we compared it with the dataset from BrasilIO (https://brasil.io/covid19/). The BrasilIO is an independent organization made of voluntaries that also gathers COVID-19 data from state health secretaries. Both datasets are very similar, in effect, we applied the Wilcoxon test (R Core Team 2020) for cases and deaths getting p-values of 0.7798 and 0.922, meaning that there are no significant differences between the two datasets. Therefore, we used the official data from the Brazilian Ministry of Health. The day July 29 was reported cases and deaths including July 28, therefore, we distributed the number of cases and deaths equally in both days. Figure 1 shows the daily time series and boxplots of cases and deaths from COVID-19 in São Paulo State between March 27, 2020 and January 03, 2021. The mean of daily cases is 1432 and for daily deaths is 55, with a maximum of 7063 on August 13, 2020 and 179 deaths on June 23, 2020. Furthermore, the variance of cases is 1511183 and deaths 1889.89, indicating that overdispersion in the data. In effect, the dispersion parameter for cases is 1035.86 and for deaths 33.84 (both p-value<0.05) (Kleiber and Zeileis 2008). Currently, São Paulo is experiencing a second wave of COVID19.

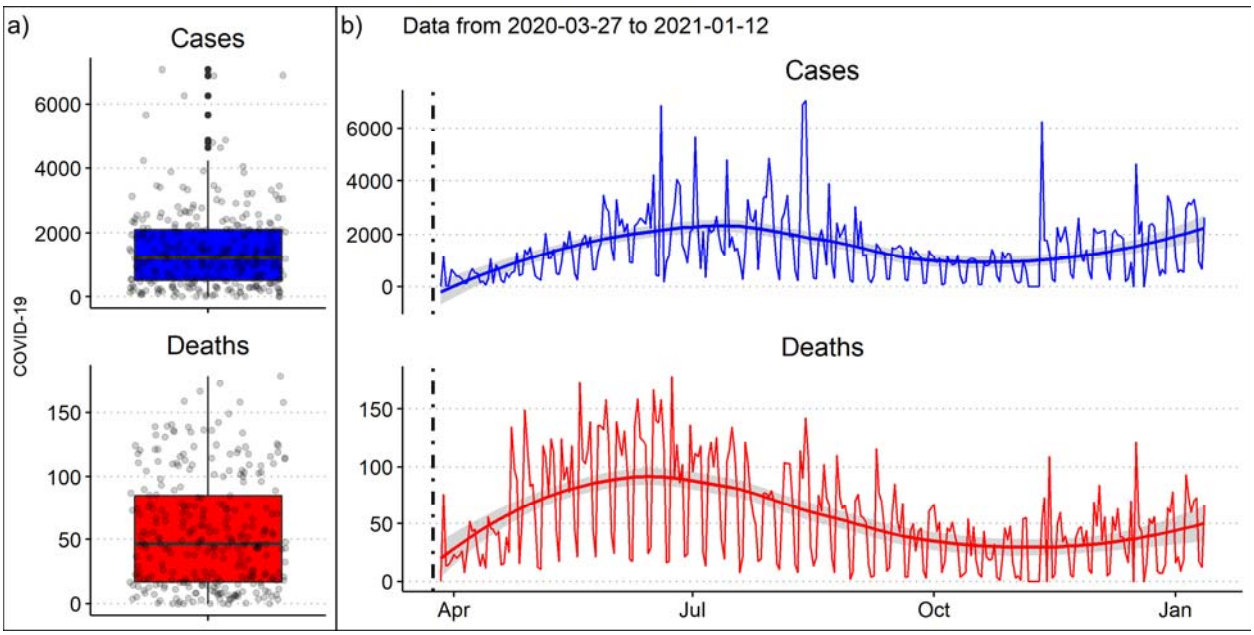

Figure 1. Daily cases and deaths of COVID-19 in (a) boxplots and (b) time series in São Paulo, Brazil, between March 27, 2020 and January 12, 2021 (Saúde 2021). The smooth lines on panel b) are the LOESS regression made with ggplot2 and R (R Core Team 2020; Wickham 2016).

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21250113; this version posted February 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

We obtained mobility indices datasets from Google mobility trends (Google, 2020) and the Intelligent Monitoring System for the city of São Paulo (SIMI-SP). The Google data is based on the use of smart devices such as cellphones, vehicle trackers, and other GPS enabled systems. The data reported by Google consists in mobility trends related to places as percentage change from baseline, with baseline as the median value, for each day of the week, during the 5-week period January 3 and February 6, 2020. We selected the mobility trend "residential", because it better represent the situation of spatial isolation (Hereafter Residential Mobility Index, RMI). The SIMI-SP data is collected by network companies which receive cellphone signals by triangulating with the nearest cellphone communication tower[2]. This means that SIMI-SP does not need internet to collect mobility information. Both indices are shown on Figure 2, meaning that higher values represent more stay-at-home and while lower values pre-quarantine conditions. RMI is an indicator that represent exposure in two ways, first, staying out of home increases the chances of getting infections with SARS-CoV2 present in aerosols (Morawska and Milton 2020) and second, increases the exposure to air pollution with deleterious effect on human health which could pose a synergetic effect. The trend on both RMI shows steady decline ahead of the start of quarantine on March 24, 2020 (black vertical line) (GESP 2020). The median RMI for Google data 13.51% and for RMI SIMI-SP is 44.75% for the whole period. The median RMI for the pre-quarantine period, that is the first 15 days of March 2020, are -1% for Google and 31.30% for SIMI-SP and after the quarantine 14% for Google and 45.48% for SIMI-SP.

[2] https://www.saopaulo.sp.gov.br/coronavirus/isolamento

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21250113; this version posted February 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

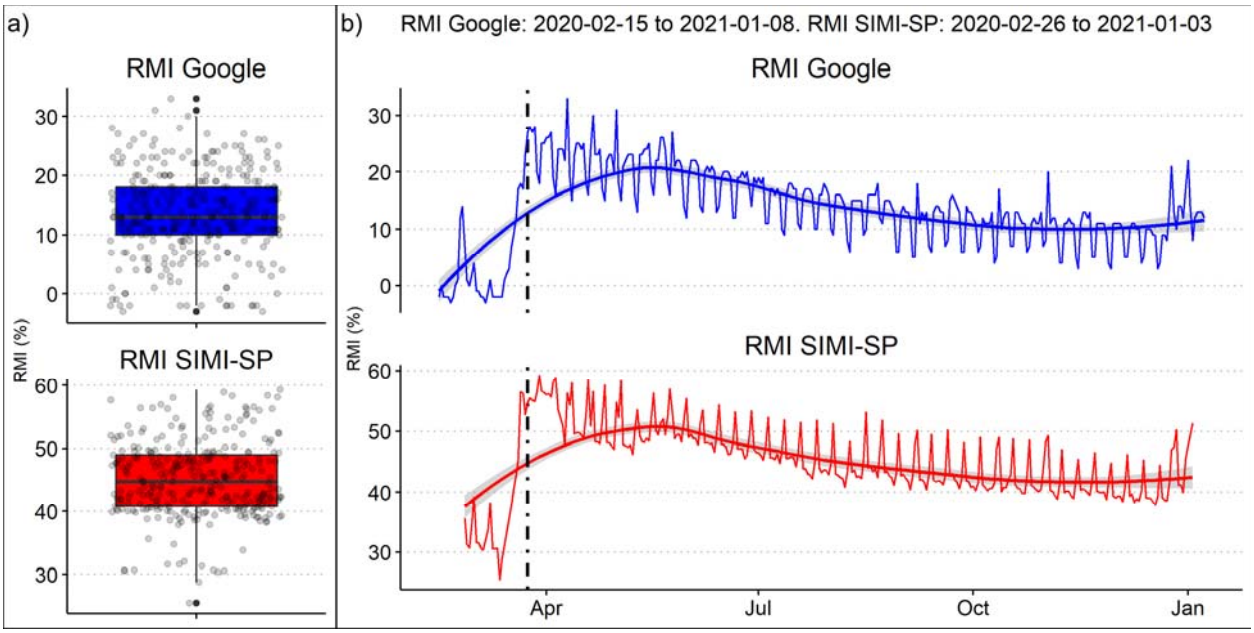

Figure 2. RMI values in São Paulo, Brazil, in a) boxplots and b) times series between January 15 and December 29, 2020 for RMI Google (Google 2021) and between February 26, 2020 and January 3, 2021. Black vertical line shows when started the quarantine in São Paulo on March, 24. The smooth lines on panel b) are the LOESS regression made with ggplot2 and R (R Core Team 2020; Wickham 2016).

Air pollution and meteorology hourly surface measurements were obtained from the air quality system (QUALAR) operated by the Environmental Agency of Sao Paulo State (CETESB, 2020). QUALAR archives air quality real-time data using several stations spread in São Paulo. Hourly averages of O_3 (ug/m³), $PM_{2.5}$ (ug/m³), Relative Humidity (%) and Temperature (°C) are shown in Figure 3, between January 01st-, 2019 and January 4, 2021. Locally estimated scatterplot smoothing (or local regression LOESS) show that that the O_3 incremented after the quarantine and remained higher than 2019 during most of 2020. The concentrations of PM2.5 during 2020 were lower than 2019, , related to the decrease in human activity and also, as reported by other studies (Bolaño-Ortiz et al. 2020; Debone et al. 2020; Nakada and Urban 2020), and just in September concentration increased. The quarantine beginning coincided with the dry season beginning at São Paulo, presenting a decline in temperature compared to the early months of 2020. The most significant feature of the São Paulo dry season (though April to November), the wet season counterpart (Vera et al. 2006), is the non-significative precipitation amount, caused by weak isolated events, and long periods with no precipitation occurring in between (Rehbein et al. 2018), and the very low relative humidity during the daytime (climate reports from the Climate Group of Studies - GrEC/USP, http://www.grec.iag.usp.br). Also, a general decreasing in temperature occurs according to the austral winter and presenting drops in temperature, associated to the synoptic systems crossing São Paulo and generally are not able to

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21250113; this version posted February 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

organize convection (GrEC/USP, http://www.grec.iag.usp.br) or even sea breezes that eventually reaches the interior of São Paulo (Freitas et al. 2007). In 2020, for instance, from April 3-May 3, 2020 there was very few (2mm at São Paulo-SESC Interlagos station, in the south of São Paulo) or no precipitation (at the São Paulo-Mirante de Santana station, in the north of São Paulo) according to the official meteorological stations (www.inmet.gov.br), while synoptic systems (such as cold fronts) were observed (CPTEC/INPE, https://www.cptec.inpe.br; GrEC/USP, http://www.grec.iag.usp.br). The first semester of 2020 was drier and colder than 2019, but during the second semester, relative humidity and temperature remained similar.

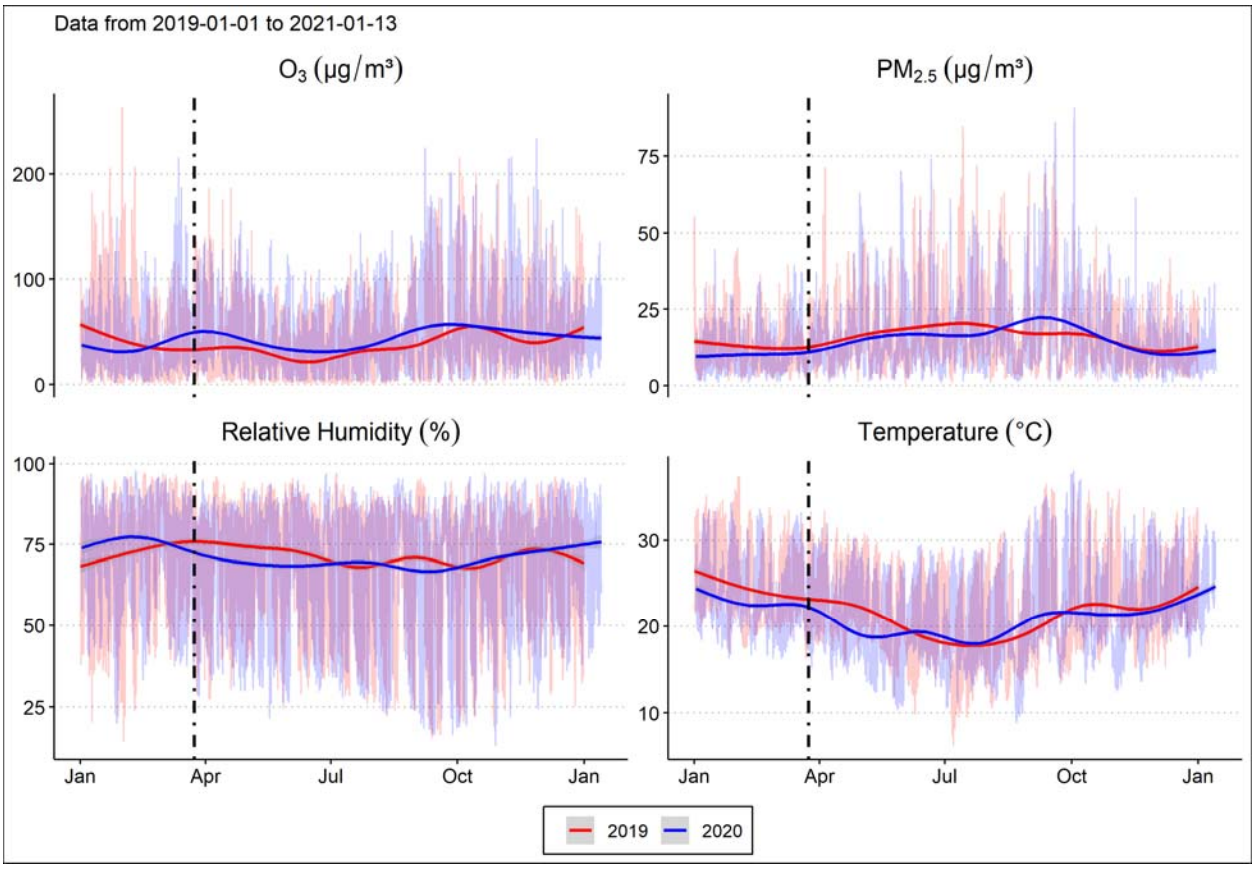

Figure 3. Hourly means of O_3 (µg/m³), $PM_{2.5}$ (µg/m³) and Air Temperature (°C) between January 1st 2020 and January 4 2021 for the city of São Paulo, Brazil (CETESB 2021). The means considered the parameters from the stations Congonhas, Cid.Universitária-USP-Ipen, Santana, Ibirapuera, Mooca, Pinheiros and Parque D. Pedro II. The red and blue lines are the automatic LOESS regression made with ggplot2 and R (R Core Team 2020; Wickham 2016).

Association between mobility and COVID-19

The statistical analyses consisted of the application of the generalized additive model (GAM) (Hastie and Tibshirani 1990). One of the most common applications of this framework consists of a semi-parametric model in

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21250113; this version posted February 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

environmental studies to understand the associations between air pollution and health outcomes by controlling other factors such as meteorology (Dominici et al. 2004; Peng and Dominici 2008). Recently, a study has shown that air pollution can increase up to 15% of COVID-19 mortality and worldwide, 27% in East Asia, 19% in Europe and 17% in North America (Pozzer et al. 2020). It has been shown that health effects of air pollution is related to previous days' exposure (Abrutzky et al., 2013; Carracedo-Martínez et al., 2010; Leitte et al., 2009). In this study we need to consider the incubation period for COVID-19, that is the period of time between the exposure to SARS CoV2 and the symptom onset. Furthermore, in this study we want to characterize how the exposure, measured by the mobility indices, is associated with COVID-19. It has been reported that the incubation period for COVID-19 is 5.1 days (95% CI, 4.5 to 5.8 days (Kraemer et al. 2020; Lauer et al. 2020), Lai et al (2020) found between 2 and 14 days with a mean of 6.4 days another (Lai et al. 2020). Therefore, we calculated moving averages between 4 and 21 days of mobility and environmental factors, and study possible associations with COVID-19. We used thin plate splines for accounting confounding factors of $PM_{2.5}$, O_3, temperature, relative humidity, day of the week and time, including interactions between the variables with quasi-poisson and negative binomial distributions to capture over-dispersion (Wood 2017; Zeileis et al. 2008). For instance, tropospheric O_3 is a secondary pollutant generated by reactions between NO_X, Volatile Organic Compounds and solar radiation (Jacob 1999), and as the diurnal cycle of temperature follows solar cycle, we would expect statistical interactions between O_3 and temperature. To identify associations between mobility and COID-19, by controlling confounding factors, we used the general equation 1. We performed a detailed sensitivity analyses between the variables, shown supplementary material S1.

$$(1) \qquad log(u_i) = \beta_0 + \beta_1 * RMI_{m,n} + s\left(PM_{2.5_m}\right) + s\left(O_{3_m}\right) + s(RH_m) + s(temp_m) + s(time) + s(dow)$$

Where log(u) is the log-transform of the daily cases and deaths of COVID-19 with a quasi-poisson or negative binomial distribution *i*, β_0 is the intercept β_1 is the coefficient that represents the association of $RMI_{m,n}$ moving average m using mobility data n from Google or SIMI-SP on cases, *s* the thin plate, and *temp* temperature, RH relative humidity, $PM_{2.5}$ and O_3 atmospheric pollutants, *time* represents each day to account unobserved factors, s(dow) is the cubic spline function with dimension of 7 to account for each day of the week. We used thin-plate splines to avoid knot placement and, therefore, avoid overfitting (Wood 2003). The predicted number of COVID-19 cases is then $exp(\beta_0 + \beta_1 * RMI_l)$, which is conditionally to the other predictors. Although more pollutants are reported by São Paulo (QUALAR) air quality stations, e.g., NO_2, CO, and PM_{10}, we limited the model to $PM_{2.5}$ and

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21250113; this version posted February 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

O_3 to avoid multicollinearity between $PM_{2.5}$ and these other species. As the objective of this study is to evaluate the effect of the quarantine measure, the data were filtered starting on March 24, 2020. In order to apply GAM, we used the R programming language and the library mgcv (v1.8.31)(R Core Team 2020; Wood 2017). One limitation of our method is the limited sample, consisting of 292 days between 2020-03-27 and 2020-12-30. Nevertheless, we performed a comprehensive sensitivity analyses, as we evaluated 18 equations with 4-21 moving average lag periods, and compared RMI data from Google and SIMI-SP with quasi-poisson and negative binomial distributions, this resulted in 1296 regressions. To ensure consistency, we repeated the analyses with data only until November 2020 finding similar results.

Association between air pollution and COVID-19

We also study the effects of air pollution on COVID-19. In this case, we are interested in evaluating the effect of specific level of air pollution present on the same or previous specific days, rather than the moving average of air pollution. Many papers have used lag models to identify associations with air pollution and health outcomes (Gasparrini 2011; Peng and Dominici 2008). Therefore, we used lags to account for the delayed effect of air pollution on COVID-19 with a quasi-poisson and negative binomial distributions (Wood 2017; Zeileis et al. 2008). Thus, we used single-lag generalized additive models and different configuration for confounding variables with thin plate splines. To identify associations between air pollution and COVID-19, by controlling confounding factors, we used the general equation 2. A sensitivity analyses is available on supplementary material S2.

$$log(u_i) = \beta_0 + \beta_1 * P_L + s(RMI_L) + s(temp_L) + s(time) + s(dow) \quad (2)$$

Where P represent the air pollutant concentrations of $PM_{2.5_L}$ or O_{3_L}, $temp_L$, RH_L, and RMI_L are the delayed environmental quantities at different lags l. The methodology is similar as we presented to calculate the effect of the RMI index, but in this case, we controlled all the variables except air pollution. This is useful to compare with other studies and to see the importance of exposure to air pollution and its effects on COVID-19 cases. Then, we calculated the relative risks of new cases by the increment of 10 $\mu g \cdot m^{-3}$air pollution with the expression $\exp(\beta_1 * pollutant$.

Results

Mobility and COVID-19 cases and deaths

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21250113; this version posted February 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

The associations between mobility indices RMI SIMI-SP and Google, parameter β_1 on equations 1-21, are shown on Figure 4. The x-axis represents the model configurations from equations 1-20 and the facet labels 4-21 the delayed effect of exposure as moving average. RMI Google was not significantly associated with COVID-19 in any model. This can be explained by the low correlation between RMI Google with Cases -0.19 and with Deaths 0.1 (not significant), as shown on **Error! Reference source not found.**. Au contraire, we did find statistical association between RMI SIMI-SP and COVID-19. It is evident increasing the mobility, that is staying out-of-home which results in lower RMI, have a delayed effect increasing COVID-19 cases after four to nine days of exposure. While some studies have shown that the incubation period for COVID-19 is between 4.5-5.8 days or 6.4 days (Kraemer et al. 2020; Lai et al. 2020; Lauer et al. 2020) here we found more days to confirm the case appearance in Brazil the Reverse transcription polymerase chain reaction (RT-PCR) tests to detect virus take between 3 days and one week. This explanation also applies to the expected cases after seven, eight, nine and 18 days of exposure. Regarding the deaths, we found more associations after 18-21 days of exposure, which makes sense because exposure between the exposure and the deaths there is the severe disease, which naturally implies more days. We also ound increased deaths after 4 days of exposure, but the magnitude is also lower, nearer to zero. Furthermore, there is a trend after 14 days of exposure decreasing the β_1, which means this association gains strength after more days, which after 21 days of exposure most of models signalized the association with mobility and COVID-19 death. These associations were found with negative binomial and quasi-poisson regressions, and the reader can reproduce these results following the instructions available in this public repository https://gitlab.com/ibarraespinosa/covid191.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21250113; this version posted February 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

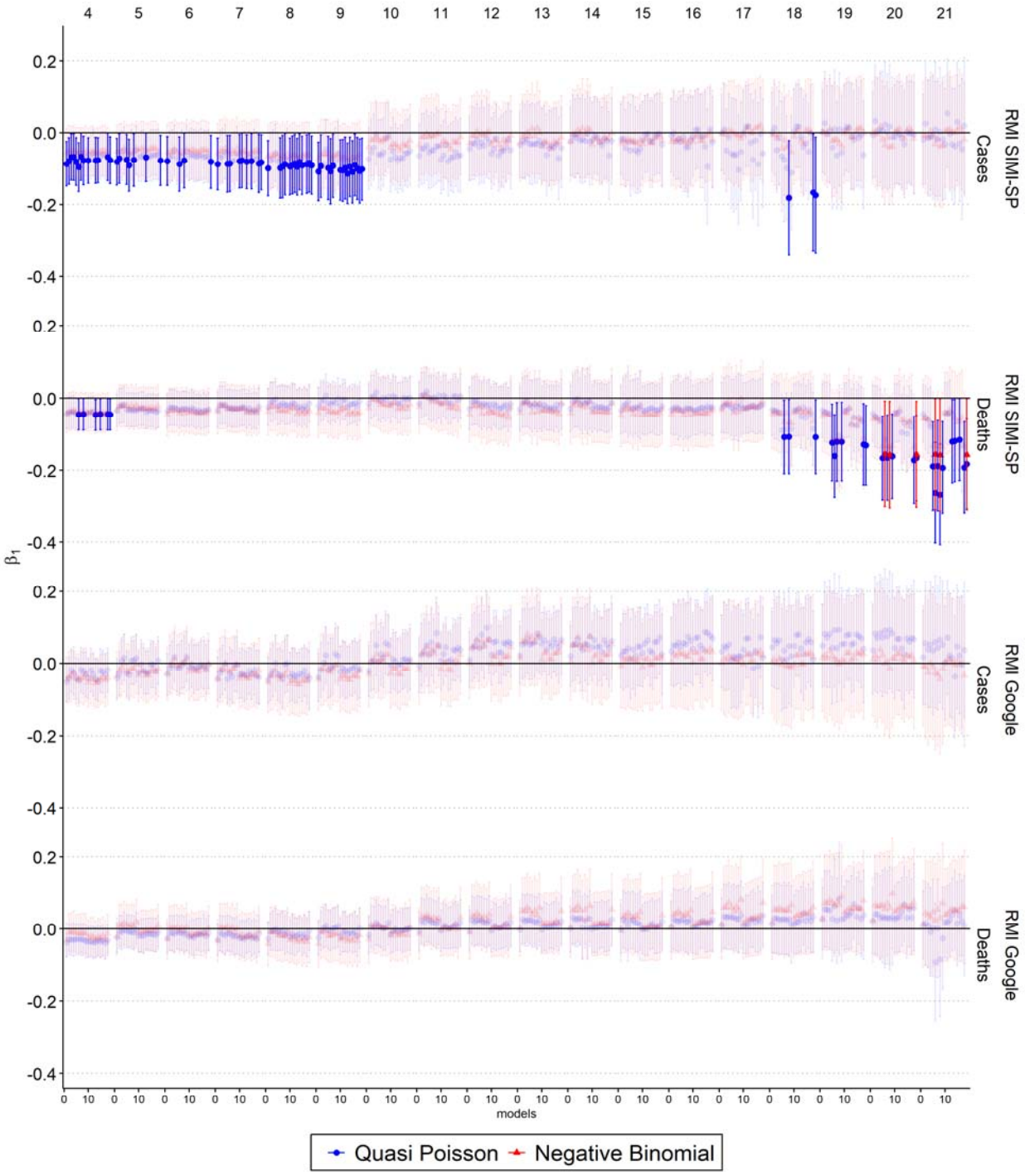

Figure 4. Coefficient of association between RMI SIMI-SP (%) and Google (%) and COVID-19 cases and deaths with a moving average of four to 21 days of delayed exposure with quasi-poisson and negative binomial distribution under different model configuration. All non-significant associations (p-value > 0.05) of β_1 were plotted transparent (alpha = 0.2) so that we can clearly see the significant associations. (R Core Team 2020; Wickham 2016).

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21250113; this version posted February 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

After analyzing the coefficients of association RMI and COVID-19, we calculated the expected outcomes after moving average periods. As the RMI produce significant results after several days of exposure, we grouped the results to obtain one curve exposure-response for cases and deaths, as shown on Figure 5. They grey points represents the expected cases and deaths under different exposure levels, expressed as RMI SIMI-SP and the red curves a gam regression with 95% confidence interval. We found strong associations between moving average of delayed exposure and COVID-19 cases and deaths. In other words, when people stayed home COVID-19 cases and deaths decreased. Likewise, with less RMI or increased outdoor activities, COVID-19 cases and resulting deaths increased. The mean RMI SIMI-SP during pre-quarantine (before March 15) was 31.10% and just after the quarantine (between March 27 and April 15) was 54.77%, reducing the mobility. Then,he association between RMI and cases was assessed with the median RMI SIMI-SP of 45.28% post-quarantine, presented as the second vertical black line in Figure 5. Under this RMI we would expect 1,212 cases (95% CI: 1,189 to 1,235) and 44 deaths (95% CI: 40 to 47). We applied the expected outcomes for several RMI values to evaluate the resulting cases for RMI extremes. For example, under the first quantile of RMI SIMI-SP, first vertical black line, that is 41.38%, it would result in 1,757 cases (95% CI: 1,734 to 1,780) and 80 deaths (95% CI: 77 to 84) and under the third quantile of 48.87%, third vertical black line, it would result in 846 cases (95% CI: 823 to 869) and 25 deaths (95% CI: 22 to 29). Analyzing the extreme values shows that, if the RMI SIMI-SP were 37.82% would result in 2,311 cases (95% CI: 2,285 to 2,338) and 127 deaths (95% CI: 122 to 131) and with RMI SIMI-SP of 59.25%, 351 cases (95% CI: 325 to 378) and 8 deaths (95% CI: 4 to 13). Therefore, avoiding unnecessary outdoor activities and staying at home would result in a reduction in expected cases and deaths. We added a repeated histogram of RMI SIMI-SP on the top of Figure 5 so that it is easier to see the quantiles and histograms. This means that most of time RMI values are below 49%, resulting that mobility was increased contributing to the current second COVID-19 wave occurring in São Paulo.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21250113; this version posted February 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

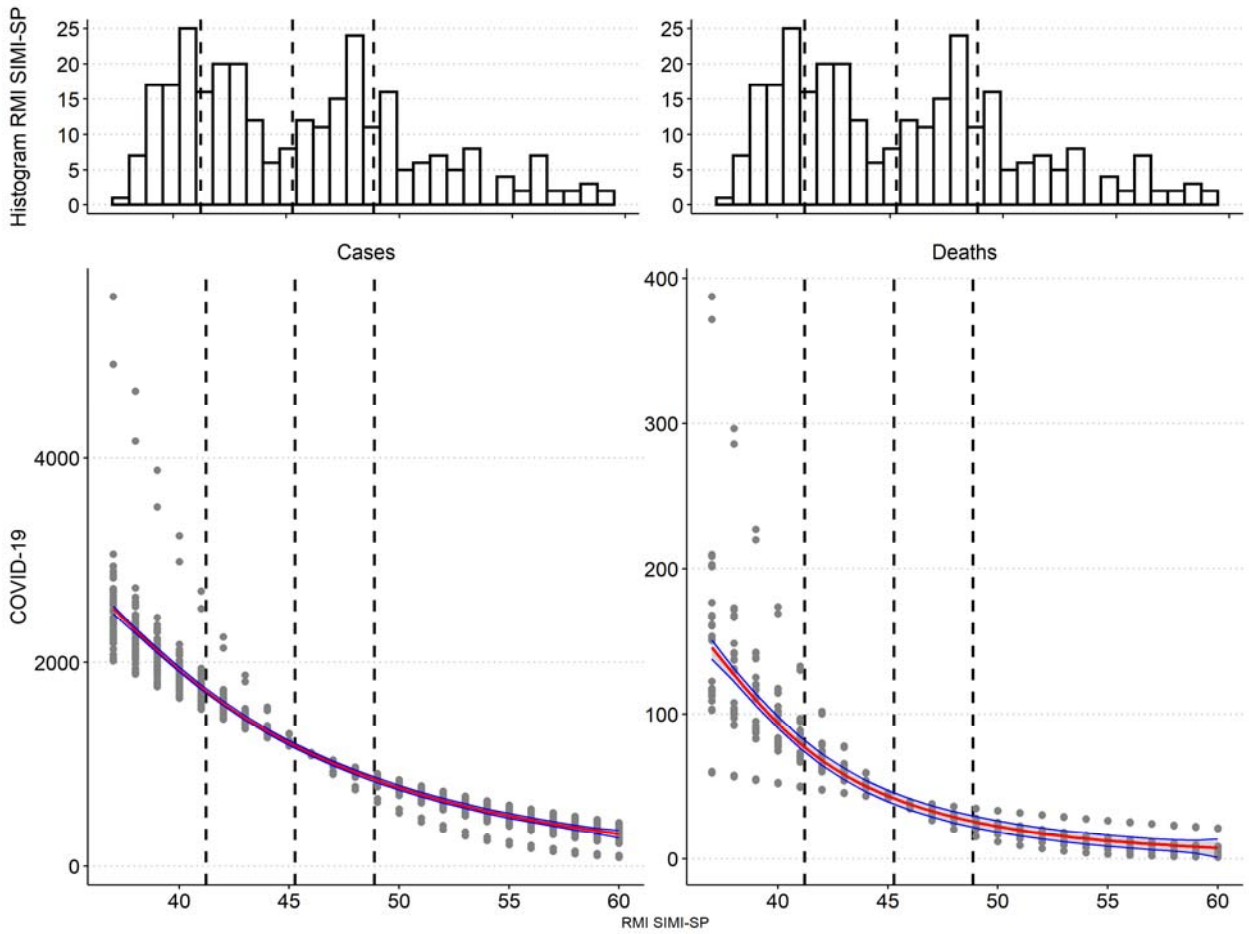

Figure 5. Association between COVID-19 cases and deaths and the percentage of RMI SIMI-SP (%). The grey points show the association between COVID-19 and RMI after different periods of exposure. The red line representes the expected cases and deaths and blue lines upper and lower confidence intervals 95%. The vertical black lines show the first quantile, median and third quantile of RMI SIMI-SP. The histograms of RMI SIMI-SP are repeated so that it is easier to the reader to compare cases with RMI.

Air pollution and COVID-19 cases

We assessed the association between COVID-19 and increment of 10 $\mu g \cdot m^{-3}$ of $PM_{2.5}$ and O_3. Figure 6 shows the relative risks of COVID-19 cases and deaths after 1-21 days of exposure. We found that $PM_{2.5}$ and O_3 have positive relative risks for cases and deaths with both distributions. Furthermore, O_3 increment cases after four and 13 days of exposure, and deaths after two, four, 19 and 20 days of exposure. While, $PM_{2.5}$ poses positive relative risks after two, 10 and 13 days of exposure for cases and 17 days of exposure for deaths. Nevertheless, there are some relative risks below 1, which would provide protective factors for O_3 and $PM_{2.5}$ after 17 days for cases where more research

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21250113; this version posted February 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

is needed. The order of magnitude of relative risks within lagged group is very similar with different models, as shown on equations 21-28. Therefore, we averaged the relative risks within groups resulting that an increment of 10 $\mu g \cdot m^{-3}$ of O_3 produce cases-relative risks of 1.066 (95% CI: 1.005 to 1.131) and 1.084 (95% CI: 1.007 to 1.168) after four and 13 days of exposure, respectively. In the case of deaths, the O_3-related relative risks are 1.067 (95% CI: 1.008 to 1.129), 1.050 (95% CI: 1.003 to 1.100), 1.070 (95% CI: 1.012 to 1.131) and 1.066 (95% CI: 1.001 to 1.135) after two, four, 19 and 20 days of exposure, respectively. Likewise, an increment of 10 $\mu g \cdot m^{-3}$ of $PM_{2.5}$ produce cases-relative risks 1.151 (95% CI: 1.048 to 1.264), 1.113 (95% CI: 1.002 to 1.236) and 1.157(95% CI: 1.012 to 1.323) after three, 10 and 14 days of exposure and the risk for death is 1.086 (95% CI: 1.008 to 1.170) after 17 days. Based on these results, air pollution significantly increases COVID-19 cases and deaths. The mean relative risks for cases are 1.140 for $PM_{2.5}$ and 1.075 for O_3, meaning that $PM_{2.5}$ increments 1.06 times more COVID-19 cases than O_3. In the case of deaths, the relative risk for O_3 is 1.063 and for $PM_{2.5}$ 1.086.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21250113; this version posted February 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

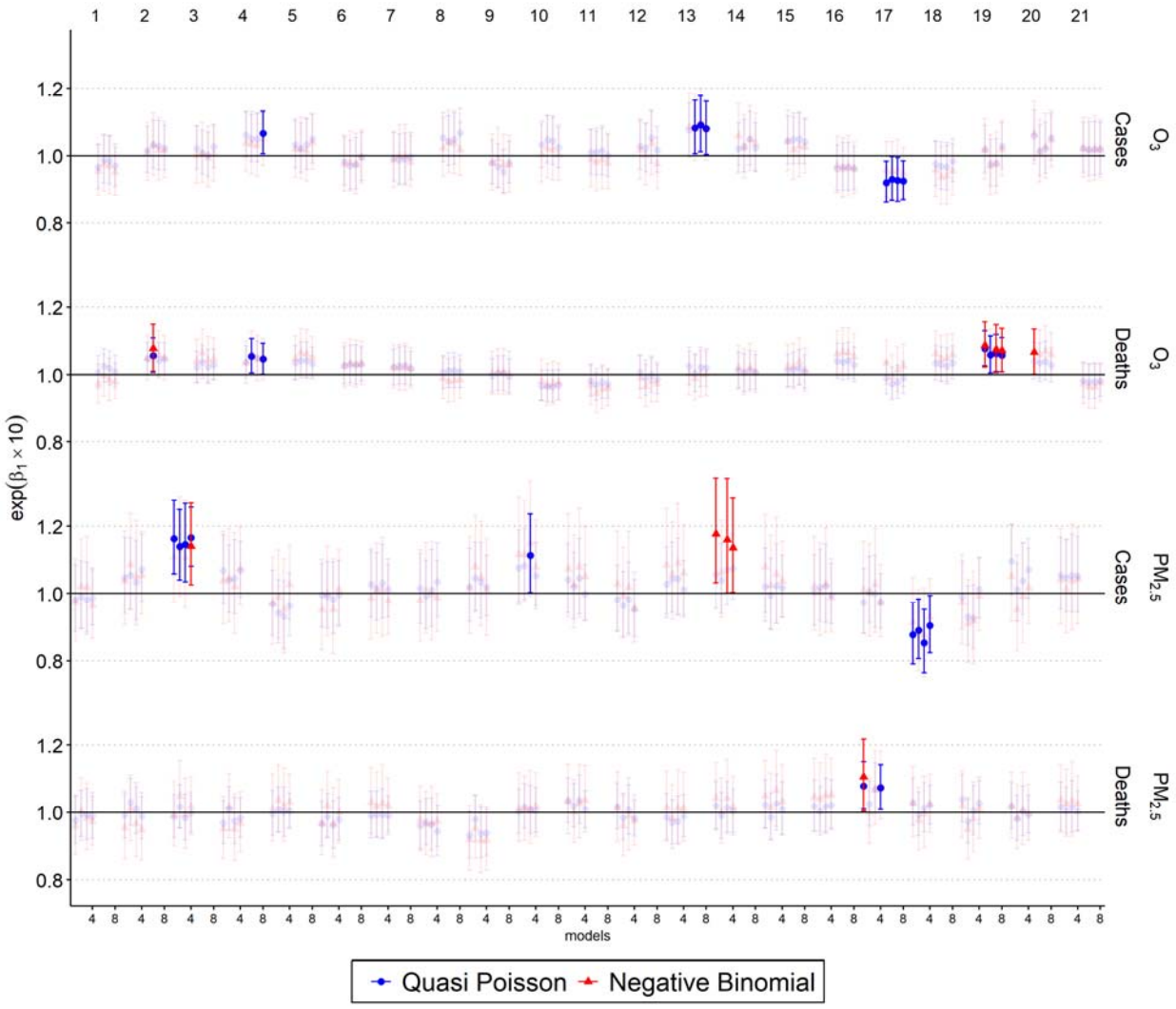

Figure 6. The relative risks of COVID-19 cases and deaths due to 10 μg·m-³ of $PM_{2.5}$ and O_3 after 1-21 days of exposure with single-lag models. All non-significant associations (p-value > 0.05) of β_1 were plotted transparent (alpha = 0.2) so that we can clearly see the significant associations.

Discussion

In this study, we used a semiparametric Generalized Additive Model (GAM) to explore possible associations between RMI, air pollutants, and COVID-19 cases and deaths in São Paulo, Brazil, for March 27, 2020, through December 30, 2020. We controlled for environmental factors such as air temperature, relative humidity, and air pollutant concentrations of $PM_{2.5}$ and O_3 with thin splines.

We found statistical associations between RMI and COVID-19 cases and a lower RMI (i.e., the increase of residents staying-at-home) increase COVID-19 cases. Likewise, increased RMI or less outdoor activities decreases COVID-19 cases. The median RMI after quarantine started was 45.28%, which represents most of the period of study. Under

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21250113; this version posted February 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

this RMI, we would expect 1,212 cases (95% CI: 1,189 to 1,235) and 44 deaths (95% CI: 40 to 47). São Paulo's COVID-19 median values are 1,214 and 46 for cases and deaths, which means that our predictions align with our observations. A Brazilian study covering several cities also found that increased mobility is associated with COVID-19 cases (Martins et al. 2020). We analyzed RMI values to provide policymakers with several options to mitigate the number of COVID-19 cases and deaths and support public health system. Then, if the mobility is reduced by increasing the RMI SIMI-SP index from 45.28% to 50.00%, we would expect 774 (95% CI: 751 to 797) cases and 23 (95% CI: 19 to 26) deaths, representing a reduction 438 cases and 21 deaths by only increasing the RMI 5%. Therefore, a policymaker can use this information and define RMI targets based on the capacity of their health system.

We evaluated the effect of moving average air pollution on COVID-19 cases and deaths and we found strong associations. The average of the significant relative risk over the 21 days of delayed exposure is 1.140 (95% CI: 1.021 to 1.274) for cases and 1.086 (95% CI: 1.008 to 1.170) for deaths due to an increment of 10 $\mu g \cdot m^{-3}$ of $PM_{2.5,}$ and 1.075 (95% CI: 1.006 to 1.150) for cases and 1.086 (95% CI: 1.008 to 1.170) for deaths due to an increment of 10 $\mu g \cdot m^{-3}$ of $O_{3.}$ A global study about the association between air pollution and death-risk for COVID-19 found that in South America the attributable fraction (AF) of COVID-19 mortality due $PM_{2.5}$, calculated as 1-1/RR, is approximately 15% in São Paulo (Pozzer et al. 2020), while the AF due the increment of $PM_{2.5}$ in Sao Paulo is 12.28. Another study in China showed that 10 $\mu g \cdot m^{-3}$ of O_3 results in relative risk of 1.047 (Zhu et al. 2020), which is slightly lower than our results of 1.075 presented in this study. Finally, Zhang et al (2021) found a country-average relative risk for $PM_{2.5}$ of 1.06 (95% CI: 1.03 to 1.08), while relative risk found in Northeast and Southwest China oscillate around 1.2, similar to our result 1.140. Therefore, our results are in agreement with other studies.

As this study is based on 292 days of data, future research should consider the potential effects of a more extended period to study the effect of air pollution on COVID-19 related cases and deaths. Recent social distancing and quarantines have been introduced on unprecedented scales, made necessary by the high transmissivity and severity of COVID-19, and the lack of effective vaccines or testing programs (Cohen and Kupferschmidt 2020). This strongly suggests that mobility indices can be used to study infectious disease transmission and assess the effectiveness of large-scale isolation and quarantine style management activities. Therefore, policymakers can use the new mobility dataset to enforce efforts to implement more effective social distancing and quarantine-based management strategies for COVID-19.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21250113; this version posted February 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Arguably, models like the one used here are more commonly applied to much larger datasets. However, we present findings and caveats, to provide early evidence on the transmission of COVID-19, and as part of efforts to highlight the potential value of the recently developed mobility indices.

Conclusion

Spatial distancing was proven to be a determining factor to control COVID-19 cases and deaths. RMI is also significantly associated with COVID-19 cases and deaths. Higher RMI values lower COVID-19 cases while lower values are related to a higher number of COVID-19 cases and deaths. Our predictions align with mean observations of COVID-19 cases. Air pollutant models revealed that an increment of 10 $\mu g \cdot m^{-3}$ of $PM_{2.5}$ and O_3 produces a relative risk of 1.140 (95% CI: 1.021 to 1.274) for cases and of 1.086 (95% CI: 1.008 to 1.170), and 1.075 (95% CI: 1.006 to 1.150) for cases and 1.063 (95% CI: 1.006 to 1.124) for deaths, respectively.

References

Abrutzky R, Ibarra S, Matus P, Romero-Lankao P, Pereyra V, Dawidowski L. 2013. Atmospheric pollution and mortality. A comparative study between two Latin American cities: Buenos Aires (Argentina) and Santiago (Chile). Int J Environ Heal 6; doi:10.1504/IJENVH.2013.056977.

Alves THE, de Souza TA, de Almeida Silva S, Ramos NA, de Oliveira SV. 2020. Underreporting of death by COVID-19 in Brazil's second most populous state. medRxiv.

Badr HS, Du H, Marshall M, Dong E, Squire MM, Gardner LM. 2020. Association between mobility patterns and COVID-19 transmission in the USA: a mathematical modelling study. Lancet Infect Dis 20: 1247–1254.

Bolaño-Ortiz TR, Camargo-Caicedo Y, Puliafito SE, Ruggeri MF, Bolaño-Diaz S, Pascual-Flores R, et al. 2020. Spread of SARS-CoV-2 through Latin America and the Caribbean region: a look from its economic conditions, climate and air pollution indicators. Environ Res 109938; doi:https://doi.org/10.1016/j.envres.2020.109938.

Buss LF, Prete CA, Abrahim CMM, Mendrone A, Salomon T, de Almeida-Neto C, et al. 2020. Three-quarters attack rate of SARS-CoV-2 in the Brazilian Amazon during a largely unmitigated epidemic. Science (80-); doi:10.1126/science.abe9728.

Carracedo-Mart\'\inez E, Taracido M, Tobias A, Saez M, Figueiras A. 2010. Case-crossover analysis of air pollution

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21250113; this version posted February 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

health effects: a systematic review of methodology and application. Environ Health Perspect 118: 1173–1182.

CETESB. 2021. Sistema de monitoreamento de qualidade do ar QUALAR. Available: https://qualar.cetesb.sp.gov.br/ [accessed 4 January 2021].

Cohen J, Kupferschmidt K. 2020. Countries test tactics in 'war'against COVID-19.

Dantas G, Siciliano B, França BB, da Silva CM, Arbilla G. 2020. The impact of COVID-19 partial lockdown on the air quality of the city of Rio de Janeiro, Brazil. Sci Total Environ 729:139085; doi:https://doi.org/10.1016/j.scitotenv.2020.139085.

Debone D, da Costa M V, Miraglia SGEK. 2020. 90 days of COVID-19 social distancing and its impacts on air quality and health in Sao Paulo, Brazil. Sustainability 12: 7440.

Dominici F, McDermott A, Hastie TJ. 2004. Improved semiparametric time series models of air pollution and mortality. J Am Stat Assoc 99: 938–948.

Dong E, Du H, Gardner L. 2020. An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect Dis 20: 533–534.

Dutheil F, Baker JS, Navel V. 2020. COVID-19 as a factor influencing air pollution? Environ Pollut (Barking, Essex 1987).

Freitas ED, Rozoff CM, Cotton WR, Dias PLS. 2007. Interactions of an urban heat island and sea-breeze circulations during winter over the metropolitan area of São Paulo, Brazil. Boundary-Layer Meteorol 122: 43–65.

Freitas, Ibarra-Espinosa S, Gavidia-Calderón M, Rehbein A, Abou Rafeem Sameh Martins J, Martins L, et al. 2020. Mobility restrictions and air quality under COVID-19 pandemic in São Paulo, Brazil. Preprints; doi:10.20944/preprints202004.0515.v1.

Gasparrini A. 2011. Distributed lag linear and non-linear models in {R}: the package {dlnm}. J Stat Softw 43: 1–20.

Gatalo O, Tseng K, Hamilton A, Lin G, Klein E. 2020. Associations between phone mobility data and COVID-19 cases.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21250113; this version posted February 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Lancet Infect Dis.

GESP. 2020. Decreta quarentena no Estado de São Paulo, no contexto da pandemia do CO VID-19. Available: http://dobuscadireta.imprensaoficial.com.br/default.aspx?DataPublicacao=20200323&Caderno=DOE-I&NumeroPagina=1.

Google. 2021. Community Mobility Reports. Available: https://www.google.com/covid19/mobility/ [accessed 4 January 2021].

Gupta R, Collier D, De Marco A, Ferreira I, Meng B, Datir R, et al. 2021. SARS-CoV-2 B. 1.1. 7 escape from mRNA vaccine-elicited neutralizing antibodies. Prepr (Version 1) available Res Sq [https//doi.org/1021203/rs3.rs-156101/v1].

Hastie TJ, Tibshirani RJ. 1990. *Generalized additive models*. CRC press.

IBGE. 2014. Instituto Brasileiro de Geografia e Estatística -- Población estimada 2014.

Jacob DJ. 1999. *Introduction to atmospheric chemistry*. Princeton University Press.

Kleiber C, Zeileis A. 2008. *Applied econometrics with R*. Springer Science & Business Media.

Kraemer MUG, Yang C-H, Gutierrez B, Wu C-H, Klein B, Pigott DM, et al. 2020. The effect of human mobility and control measures on the COVID-19 epidemic in China. Science (80-) 368: 493–497.

Kupferschmidt K. 2021. Fast-spreading U.K. virus variant raises alarms. Science (80-) 371:9–10; doi:10.1126/science.371.6524.9.

Lai C-C, Shih T-P, Ko W-C, Tang H-J, Hsueh P-R. 2020. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and corona virus disease-2019 (COVID-19): the epidemic and the challenges. Int J Antimicrob Agents 105924.

Lauer SA, Grantz KH, Bi Q, Jones FK, Zheng Q, Meredith HR, et al. 2020. The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. Ann Intern Med.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21250113; this version posted February 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Leitte AM, Petrescu C, Franck U, Richter M, Suciu O, Ionovici R, et al. 2009. Respiratory health, effects of ambient air pollution and its modification by air humidity in Drobeta-Turnu Severin, Romania. Sci Total Environ 407: 4004–4011.

Martins LD, da Silva I, Batista WV, de Fátima Andrade M, de Freitas ED, Martins JA. 2020. How socio-economic and atmospheric variables impact COVID-19 and influenza outbreaks in tropical and subtropical regions of Brazil. Environ Res 191: 110184.

Morawska L, Milton DK. 2020. It Is Time to Address Airborne Transmission of Coronavirus Disease 2019 (COVID-19). Clin Infect Dis; doi:10.1093/cid/ciaa939.

Morris DH, Yinda KC, Gamble A, Rossine FW, Huang Q, Bushmaker T, et al. 2020. Mechanistic theory predicts the effects of temperature and humidity on inactivation of SARS-CoV-2 and other enveloped viruses. bioRxiv; doi:10.1101/2020.10.16.341883.

Nakada LYK, Urban RC. 2020. COVID-19 pandemic: Impacts on the air quality during the partial lockdown in São Paulo state, Brazil. Sci Total Environ 730:139087; doi:https://doi.org/10.1016/j.scitotenv.2020.139087.

Peng RD, Dominici F. 2008. Statistical methods for environmental epidemiology with R. R a case study air Pollut Heal.

Pozzer A, Dominici F, Haines A, Witt C, Münzel T, Lelieveld J. 2020. Regional and global contributions of air pollution to risk of death from COVID-19. Cardiovasc Res; doi:10.1093/cvr/cvaa288.

R Core Team. 2020. R: A Language and Environment for Statistical Computing.

Rehbein A, Dutra LMM, Ambrizzi T, da Rocha RP, Reboita MS, da Silva GAM, et al. 2018. Severe weather events over southeastern Brazil during the 2016 dry season. Adv Meteorol 2018.

Saúde M de. 2021. Coronavírus Brasil. Available: https://covid.saude.gov.br/ [accessed 4 January 2021].

Tian H, Liu Y, Li Y, Wu C-H, Chen B, Kraemer MUG, et al. 2020. An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. Science (80-) 368:638–642; doi:10.1126/science.abb6105.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21250113; this version posted February 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Tosepu R, Gunawan J, Effendy DS, Lestari H, Bahar H, Asfian P, et al. 2020. Correlation between weather and Covid-19 pandemic in Jakarta, Indonesia. Sci Total Environ 138436.

Vera C, Higgins W, Amador J, Ambrizzi T, Garreaud R, Gochis D, et al. 2006. Toward a Unified View of the American Monsoon Systems. J Clim 19:4977–5000; doi:10.1175/JCLI3896.1.

Wang J, Tang K, Feng K, Lv W. 2020a. High temperature and high humidity reduce the transmission of COVID-19. Available SSRN 3551767.

Wang M, Jiang A, Gong L, Luo L, Guo W, Li C, et al. 2020b. Temperature significant change COVID-19 Transmission in 429 cities. medRxiv.

Wickham H. 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

Wood S. 2017. *Generalized Additive Models: An Introduction with R*. Chapman and Hall/CRC.

Wood SN. 2003. Thin plate regression splines. J R Stat Soc Ser B (Statistical Methodol 65: 95–114.

Wu X, Nethery RC, Sabath BM, Braun D, Dominici F. 2020. Exposure to air pollution and COVID-19 mortality in the United States. medRxiv.

Zeileis A, Kleiber C, Jackman S. 2008. Regression Models for Count Data in R. J Stat Software, Artic 27:1–25; doi:10.18637/jss.v027.i08.

Zhang X, Tang M, Guo F, Wei F, Yu Z, Gao K, et al. 2021. Associations between air pollution and COVID-19 epidemic during quarantine period in China. Environ Pollut 268:115897; doi:https://doi.org/10.1016/j.envpol.2020.115897.

Zhu Y, Xie J, Huang F, Cao L. 2020. Association between short-term exposure to air pollution and COVID-19 infection: Evidence from China. Sci Total Environ 727:138704; doi:https://doi.org/10.1016/j.scitotenv.2020.138704.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.08.21250113; this version posted February 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Acknowledgments: The authors would like to thank the São Paulo's Environmental Agency (CETESB), for making possible the use of air quality measurements through its QUALAR platform. The authors would like to thank Lena Goodwin for the support. **Funding:** This work was supported by FAPESP (Grants No. 2015/03804-9, 2016/18438-0 and 2016/10557-0) and "Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil" (CAPES) Finance Code 001; **Author contributions:** Conceptualization, S.I.E. and E.F.; Methodology, S.I.E, E,F., F.D., K.R., A.R.; Investigation, S.I.E, E.F., K.R., F.D. and A.R.; Writing – Original Draft, S.I.E., K.R. and E.D.; Writing – Review & Editing, S.I.E., K.R., E.F., F.D. and A.R.; Funding Acquisition, E.F.; Supervision, E.F. and F.D; **Competing interests:** Authors declare no competing interests; and **Data and materials availability:** All the data and code used in this study is available at the only repository https://gitlab.com/ibarraespinosa/covid191.