


RESEARCH ARTICLE | MAY 10 2023

Temporal relation between human mobility, climate, and COVID-19 disease

Carlos F. O. Mendes; Eduardo L. Brugnago; Marcus W. Beims ; ... et. al



Chaos 33, 053110 (2023)

<https://doi.org/10.1063/5.0138469>



CrossMark

Articles You May Be Interested In

Natural Radiation from Soil using Gamma-Ray Spectrometry

AIP Conference Proceedings (June 2009)

Effective gamma-ray doses due to natural radiation from soils of southeastern Brazil

AIP Conference Proceedings (August 2010)

Letters from the edge: Less conventional acoustical solutions

Proc. Mtgs. Acoust (June 2013)



Chaos

Special Topic: Nonlinear Model
Reduction From Equations and Data

Submit Today!

Temporal relation between human mobility, climate, and COVID-19 disease



Cite as: Chaos 33, 053110 (2023); doi: 10.1063/5.0138469

Submitted: 12 December 2022 · Accepted: 20 March 2023 ·

Published Online: 10 May 2023



View Online



Export Citation



CrossMark

Carlos F. O. Mendes,^{1,a)} Eduardo L. Brugnago,^{2,3,b)} Marcus W. Beims,^{3,c)} and Alice M. Grimm^{3,d)}

AFFILIATIONS

¹Escola Normal Superior, Universidade do Estado do Amazonas, 69050-010 Manaus, Amazonas, Brazil

²Instituto de Física, Universidade de São Paulo, 05508-090 São Paulo, SP, Brazil

³Departamento de Física, Universidade Federal do Paraná, 81531-980 Curitiba, Paraná, Brazil

^{a)}Electronic mail: cfabio.mendes@gmail.com

^{b)}Electronic mail: elb@if.usp.br

^{c)}Author to whom correspondence should be addressed: mbeims@fisica.ufpr.br

^{d)}Electronic mail: grimm@fisica.ufpr.br

ABSTRACT

Using the example of the city of São Paulo (Brazil), in this paper, we analyze the temporal relation between human mobility and meteorological variables with the number of infected individuals by the COVID-19 disease. For the temporal relation, we use the significant values of distance correlation t_0 (DC), which is a recently proposed quantity capable of detecting nonlinear correlations between time series. The analyzed period was from February 26, 2020 to June 28, 2020. Fewer movements in recreation and transit stations and the increase in the maximal temperature have strong correlations with the number of newly infected cases occurring 17 days after. Furthermore, more significant changes in grocery and pharmacy, parks, and recreation and sudden changes in the maximal pressure occurring 10 and 11 days before the disease begins are also correlated with it. Scanning the whole period of the data, not only the early stage of the disease, we observe that changes in human mobility also primarily affect the disease for 0–19 days after. In other words, our results demonstrate the crucial role of the municipal decree declaring an emergency in the city to influence the number of infected individuals.

Published under an exclusive license by AIP Publishing. <https://doi.org/10.1063/5.0138469>

Since 2019, the spreading of coronavirus SARS-CoV-2 worldwide has been a crucial issue in human life. Many processes can affect the dissemination of the virus, ranging from the distance between humans, vaccination, correct use of masks, intrinsic rate of infection of the virus, human health, and human mobility, which can be a consequence of the weather, climate, or government containment measures, for example. It is a nonlinear complex system in which many variables may become relevant. In the present work, we study the temporal relation between newly infected individuals by the coronavirus, human mobility, and weather in São Paulo (Brazil) from February 26, 2020 to June 28, 2020. For human mobility, we considered the categories grocery and pharmacy, parks, residential, retail and recreation, transit stations, and workplaces, and for meteorological variables, we have radiation, precipitation, wind, and maximum, minimum, and mean values of temperature, pressure, and humidity. To find the temporal relation between all variables, we use the distance correlation, one of the latest statistical measures that detect nonlinear

correlations. We show that while some meteorological variables can affect the number of newly infected cases around 17 days after, changes in human mobility are crucial to deciding what happens in the coming three weeks. The underlying mechanism that correlates newly infected individuals with viruses, human mobility, and weather should be independent of the time they occur and applicable to present circumstances.

I. INTRODUCTION

The relation between meteorological data and individuals' contamination by the coronavirus SARS-CoV-2 is an actual open issue of major relevance. Such a relation can be studied based on the direct effects of meteorological data on the coronavirus contamination of individuals or indirectly via meteorological effects on the social behavior, which then affects the contamination. While almost all recent studies focus on direct analysis, the indirect effect is certainly

relevant and whose importance should be quantified. In this context, it has been suggested that specific meteorological variables may play a role in the spread of the coronavirus in some countries,^{1–19} while others oppose strong seasonal effects.^{20,21} There are also proposals that several meteorological factors combined could describe the epidemic trend much better than single-factor models.²² On the other hand, big social events were shown to be analogous to the perturbations of information dissemination, increasing the predominance of the coronavirus infection.²³ Furthermore, social distancing in social networks²⁴ can also affect the spreading of the coronavirus.

It is known in statistics that correlation does not necessarily imply causation. The adequate term for the causality between times series is *temporal relation*.²⁵ A time-series X is temporally related to a time-series Y if the t_0 -student test indicates that X and Y have a statistically significant relationship within a given lag between the time series. In the present paper, we use the *distance correlation* (DC)²⁶ to analyze the temporal relation between meteorological quantities and individuals' cases with the coronavirus. The DC varies between -1 and 1 , meaning maximal anti-correlation and correlation, respectively. In opposition to the Pearson correlation (ρ), the DC has the advantage of detecting nonlinear correlations, and it will be zero if and only if the analyzed quantities are independent. The Pearson correlation can be zero, for example, even though the quantities are still correlated. Therefore, the DC vanishes only when data are completely uncorrelated.

DC is a robust statistical measure that extracts nonlinear properties from dynamical systems. Thus, DC is most welcome in the context of chaotic time series, and we mention some examples. The DC and Pearson's correlation between noisy and noiseless quadratic maps are studied in periodic and chaotic regimes.²⁷ In distinction to Pearson's correlations, DC can recognize the correct qualitative behavior of noise-induced escape time decays and the mixing of chaotic trajectories. Furthermore, the decay of the DC has been shown to be related to the Lyapunov exponent of discrete dynamical systems with one positive Lyapunov.²⁸ In the study of spatially extended systems, DC is used as a powerful quantifier to describe correlation and synchronization (chimera states).²⁹ In other contexts, DC has been shown to be a suitable statistical procedure for the analysis of disease spreading, such as COVID-19 (SARS-CoV-2), between different countries³⁰ and distinct large territories inside Brazil.³¹ Other utilities of DC in physical systems are still under analysis. It is used, for example, in astrophysical databases,³² in clinical data analysis,³³ functional connectivity in neuroimages,³⁴ in the detection of nonlinear model structure in the parameter estimation,³⁵ and in the climate system and the human heart.³⁶ Thus, the conclusions obtained using the computation of DC and the elimination of insignificant values through the test statistics, as described in the [Appendix](#), are in the context of chaotic time series. Another causal measure could be used as, for example, the PCMCI measure, which includes mutual information between large datasets.³⁶ However, for the purpose of the present work, the DC and Pearson correlations are robust and sufficient.

Here, we use the DC and ρ to analyze the temporal relation between meteorological quantities, human mobility, and individuals' cases with the coronavirus in the city of São Paulo (SP), Brazil. Its population is over 12×10^6 inhabitants so that our analysis is representative of large cities around the world. Meteorological data

were obtained from the stations *Mirante de Santana* and *Interlagos* from the INMET (*Instituto Nacional de Meteorologia*), the Brazilian Meteorological Service. While the Mirante de Santana station is localized in the north part of the city, the Interlagos station is an automatic station localized in the south part of the city. Together, both stations appropriately describe the weather in the city of SP. Human mobility was considered in the following categories: grocery and pharmacy (places like grocery markets, food warehouses, farmers markets, specialty food shops, drug stores, and pharmacies), parks (national parks, public beaches, marinas, dog parks, and public gardens), residential, retail and recreation (restaurants, cafes, shopping centers, theme parks, museums, libraries, and movie theaters), transit stations (public transport hubs such as subway, bus, and train stations), and workplaces. Data were obtained from the community mobility reports provided by Google.³⁷ Finally, data for COVID-19 were obtained from the FIOCRUZ (*Fundação Oswaldo Cruz*)³⁸ and the coronavirus panel of the *Ministry of Health*.³⁹

In short, our analysis demonstrates that social behavior mostly affects the earlier stage of the disease after 8–17 days. Specifically relevant are (i) fewer movements in recreation and transit stations and increasing values of the maximal temperature occurring 17 days before the early stage of the disease and (ii) larger changes in grocery and pharmacy, parks, and recreation and sudden change in the maximal pressure occurring 10–11 days before the disease.

II. COVID-19 EVOLUTION FOR THE CITY OF SP

[Figure 1\(a\)](#) displays the data for the new cases (NCs) of infected individuals used in our analyses as a function of time from February 26, 2020 to June 28, 2020. In [Fig. 1\(b\)](#), the two time windows with 7 and 14 days in the early stages of the disease are shown. While the 7-day window has a sudden increase of NC on April 1, the 14-day window includes some additional slow increases of NC for days between March 22 and 27, followed by a slight decrease until March 30. The NC values inside the 7(14)-day window are used to calculate the DC with data from the human mobility and meteorology inside another 7(14)-day window but shifted by a lag, as illustrated schematically in [Fig. 1\(c\)](#). For lag (0), the two windows overlap. The computational procedures to determine DC and ρ with their significances through the test statistics are presented in the [Appendix](#). The test statistics analysis guarantees statistically relevant values for the temporal relation mentioned in the Introduction. From the definition, DC is always positive, but in order to describe anti-correlations, we incorporate the signal of ρ in the DC, namely, $DC = \text{sgn}(\rho)DC$. Furthermore, while the mean value of the NC inside the 7-day window is 433.57 with relative variance 251.89, for the 14-day window, the mean is 216.78 with relative variance 163.12. For later purposes, we call to attention that when comparing the meteorological data with the 7-day and 14-day windows of the NC in [Fig. 1\(b\)](#), larger variations of the NC data occur in the former.

Results presented in this paper are a compendium of an extensive numerical analysis calculating the DC, ρ , $t_0(\text{DC})$, and $t_0(\rho)$ between quantities related to the spread of the COVID-19 with human mobility and meteorological data. For the disease, we used the number of daily new cases (NC) and total cases (TC) of individuals infected, the daily number of new deaths (ND) and the total number of deaths (TD). We used the meteorological variables:

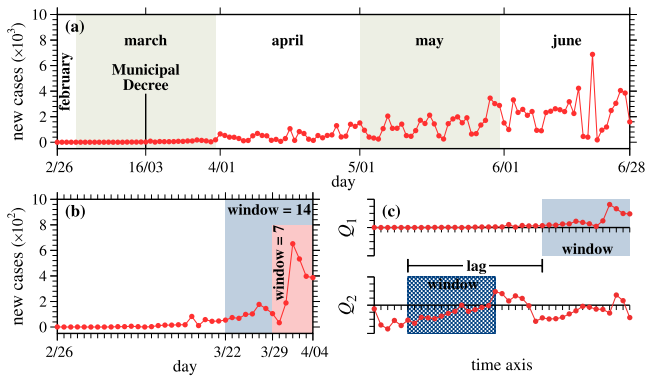


FIG. 1. (a) Number of new cases (NCs) in the city of SP starting on February 26 for 124 days. The vertical black straight line on March 16 demarcates the date of the municipal decree declaring emergency situation in the city of SP; (b) NC showing the two time windows at the early stage of the disease which are used for the analysis of DC in Sec. III; and (c) illustration of the lag between the meteorological or mobility data windows (hatched, Q_2) with the NC window (Q_1), used in Sec. IV.

radiation, precipitation, wind, and maximum, minimum, and mean values of temperature, pressure, and humidity. After analyzing all correlations, we concluded that NC comprises the main results of the analyzed COVID-19 data and that only significant values of DC between NC and T_{\max} and P_{\max} are of relevance. In addition, not all results are shown here for both windows of 7 and 14 days, just those with the most significant values. In general, the autocorrelation of NC data displays (not shown) large values for 7-day lags, a consequence that data from Sunday are included in Monday's data. That is why we use 7- and 14-day time windows to avoid apparent correlations irrelevant to the paper.

A relevant point is the ability to recognize if significant values of DC and ρ [indicated by $t_0(\text{DC})$ and $t_0(\rho)$, respectively] are related to increasing or decreasing cases of NC. We do not have a general answer to this point since the NC data oscillate too much. However, some hints can be furnished. Figure 1(b) shows the two time windows in the early stage of the disease. While the 7-day time window includes a sudden increase of NC on March 30, the 14-day time window includes this sudden increase but also some additional smaller values of NC for days between March 22 and 29. Roughly speaking, when comparing to the 14-day window, we can argue that the 7-day window has proportionally more cases with larger values of NC and also a proportionally larger sudden increase. Consequently, larger correlations obtained using 7-day windows when compared to those obtained using 14-day windows suggest a worsening of the disease situation.

III. THE EARLY STAGE OF THE DISEASE (SCANNING ONE WINDOW)

In this section, we discuss results regarding the correlation between the moving 7- and 14-day time windows for the human mobility, T_{\max} , P_{\max} , and the corresponding 7- and 14-day time windows for the NC in the early stage of

the disease, as shown in Fig. 1(b). For the numerical procedure, we use X for the COVID-19 data and Y for meteorological data and human mobility. For the case of 7-day time windows, this means $(X, Y) = \{(X_k, Y_{k-\text{lag}}) : k = 1, \dots, 7; \text{lag} = 0, -1, -2, \dots, -32\}$, with k referring to the dates inside the red window from Fig. 1(b), namely, 3/29, 3/30, \dots , 4/4. For the case of 14-day time windows, this means $(X, Y) = \{(X_k, Y_{k-\text{lag}}) : k = 1, \dots, 14; \text{lag} = 0, -1, -2, \dots, -25\}$, with k referring to the dates inside the blue window from Fig. 1(b), namely, 3/22, 3/23, \dots , 4/4.

Here, by scanning only one window, the meteorological data and human mobility windows intersect for lags from 0 (full intersection) to -6 (-13) (1-day intersection) for 7(14)-day time window.

A. Influence of human mobility

Figure 2 displays the community mobility as a function for the considered time interval in different locations: Fig. 2(a) for grocery and pharmacy, Fig. 2(b) parks, Fig. 2(c) residential, Fig. 2(g) retail and recreation, Fig. 2(h) transit stations, and Fig. 2(i) for workplaces. The vertical pink dashed lines mark the day of the municipal decree number 59 283 declaring an emergency in the city of SP. Changes for each day are compared to a baseline value for that day of the week: the baseline is the median value for the corresponding day of the week, during the five weeks from January 3 to February 6, 2020. During the period of the determination of the baseline, there was one holiday on January 25. Since this is at the weekend, we do not expect to have significant changes in the baseline. Changes in mobility are relative changes compared to the baseline values. For example, the value of 30% in the parks on Friday, March 6, corresponds to a relative change, which occurred on this day compared to Friday, February 28. Significant changes in mobility are observed in all locations (except for grocery, pharmacy, and residential) around March 16, the municipal decree day. For the grocery and pharmacy, a relevant mobility change was observed on March 21. The mobility for residential essentially increases smoothly for the whole period.

Figures 2(d)–2(f) and 2(j)–2(l) show the corresponding values of $t_0(\text{DC})$ in dark blue filled circles and $t_0(\rho)$ in orange filled circles, between the 14-day time window of the human mobility and the 14-day time window from NC, shown in Fig. 1(b), as a function of the time lag between both windows. The light blue (yellow) background colors mark the positive (negative) significant values of DC. While positive values reflect correlations between the time series, negative values belong to anti-correlations between the signals. In all cases, $t_0(\text{DC}) > t_0(\rho)$, which reveals the relevance of nonlinear effects in the time series. However, significant values of both correlators occur essentially at the same lags (some exceptions occur with a one-day delay). For each lag, there is a corresponding 14-day window demarcated with a light blue or yellow rectangle in Figs. 2(a)–2(c). For example, in Fig. 2(d), we have 11 points at lags $(0, -4 \rightarrow -13)$, which are significant for both correlators. The lags at $(-4 \rightarrow -13)$, which belong to the time interval from March 9 to 31, are marked by the yellow rectangle in the positive part of the vertical axis in Fig. 2(a). Thus, an anti-correlation between the series is observed, not meaning that the decrease in human mobility in grocery and pharmacy is responsible for the increase of NC. It is merely a consequence that the data from one series decrease while from the other

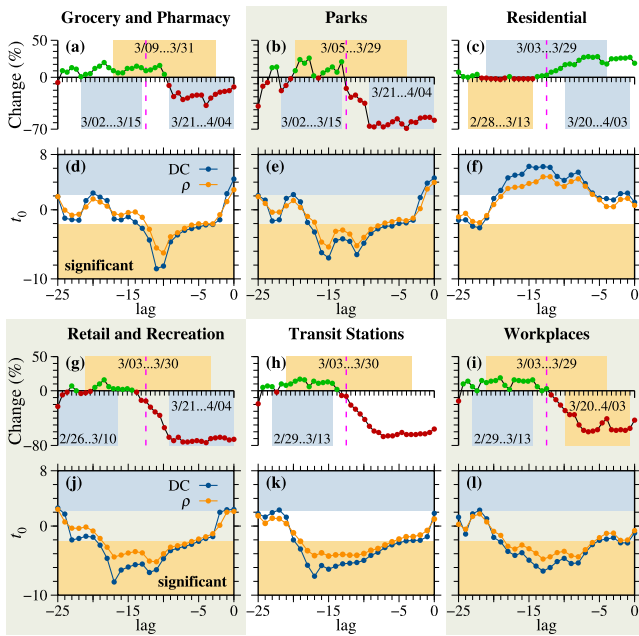


FIG. 2. Human mobility in the city of SP, starting at February 26 for five weeks in different locations (a) grocery and pharmacy, (b) parks, (c) residential, (g) retail and recreation, (h) transit stations, and (i) workplaces. Positive (negative) mobility is plotted by green (red) filled circles. Panels (d)–(f) and (j)–(l) display the statistically significant tests, $t_0(\text{DC})$ and $t_0(\rho)$, between the 14-day time window from the human mobility and the 14-day time window from NC. Vertical dashed lines in panels (a)–(c) and (g)–(i) mark the day of the emergency decree.

increase in the time interval. For the lag (0), the corresponding time interval is from March 21 to April 4 and is marked by the light blue rectangle in the negative part of the vertical axis in Fig. 2(a). Such identification with the light blue and yellow rectangles allows us to analyze in all figures the periods of human mobility, which have significant values of DC and ρ . All light blue and yellow background rectangles in the upper part of Figs. 2(a)–2(c) and Figs. 2(g)–2(i) encompass days for which big changes from positive to negative human mobility occur. The second observation is that there are lags with pronounced significant values of DC and ρ for most cases. Examples are the lags (–10, –11) in Fig. 2(d), lags (–11, –15) in Fig. 2(e), and, finally, lag (–17) in Figs. 2(j) and 2(k), related to recreation and transit stations. Since (–17) is the largest lag with the most pronounced significant values, we can associate the *primordial* human mobility in retail, recreation and transit stations as the most relevant for the earlier stage of the disease. Following this way of thinking, the next pronounced lag occurs at (–15) for parks in Fig. 2(e). This lag is also relevant for residential [Fig. 2(f)] but with previous positive mobility values in this location associated with NC reduction. The corresponding light blue and yellow windows are shown in Figs. 2(b) and 2(c), respectively. Thus, parks and residential human mobility appear to be the next in the relevant order. The additional pronounced lags that appear around (–12, –11) for retail, recreation and parks could be interpreted as a “second wave” effect from these locations. Workplaces and residential movements

are almost equally relevant for lags between (–8 → –17), but with opposite signals of the correlations. Furthermore, we may argue that our analysis’s most relevant quantity is the light blue and yellow area below $t_0(\text{DC})$ and $t_0(\rho)$ curves in Figs. 2(d)–2(f) and 2(j)–2(l). From this point of view, all analyzed locations have roughly the same relevance for lags between (–8 → –17), and we conclude that social behavior mostly affects the earlier stage of the disease for these lags.

B. Influence of maximum temperature and maximum pressure

Figure 3 displays meteorological data in the city of SP for 39 days in the early stage of the disease, starting on February 26 and finishing on April 4. While in Figs. 3(c)–3(d) we present results for the 7-day time window, Figs. 3(g)–3(h) show results for the 14-day time window. Figures 3(a) and 3(b) show data from the Interlagos station for P_{max} and T_{max} , respectively. Figures 3(c) and 3(d) display $t_0(\text{DC})$ and $t_0(\rho)$ as a function of the lag, calculated for DC and ρ , respectively, between the 7-day time windows for P_{max} and T_{max} , and the NC values inside the 7-day time window from Fig. 1(b). Significant values of DC are seen in Figs. 3(c) and 3(d) at lags (–6, –7, –10, –16, –17, –29) and (–4, –5, –6, –10, –11, –12, –17, –22, –23, –24, –28), respectively. On the other hand, significant values for ρ in Figs. 3(c) and 3(d) occur at lags (–6, –7, –10, –29) and (–4, –5, –11, –17, –22, –23), respectively. The corresponding windows for significant DC values are shown as light blue and yellow background rectangles in Figs. 3(a) and 3(b). The largest significant values of DC and ρ occur for lag (–10) in Fig. 3(c) and lag (–4, –11, –17) in Fig. 3(d). The lag (–10) from Fig. 3(c) corresponds to the light blue rectangle in Fig. 3(a) for days 19–26 of March, for which a pronounced minimum followed by a sudden increase of P_{max} is seen. For the lag (–4, –5, –6) in Fig. 3(d), the corresponding light blue rectangle in Fig. 3(b) occurs for days 23–31 of March, for which a relative increase of T_{max} is observed. The next two lags with significant values in Fig. 3(d) occur for (–11) with negative correlation, and for (–17) with positive correlation, and corresponds to a sudden decrease (yellow rectangle from March 17 to 25) and a sudden increase (light blue rectangle from March 11 to 18) of T_{max} , respectively. Thus, *apparently, a sudden increase (decrease) of T_{max} induce significant correlations (anti-correlations) between this meteorological quantity and NC.* The same conclusion can be made regarding P_{max} . Furthermore, let us compare this with human mobility using just some pronounced peaks of $t_0(\text{DC})$. The minimum in Fig. 3(d) at lag (–11) is related to the rectangle from 17 to 25 of March (sudden decrease of T_{max}) that corresponds to the period for which almost all relevant changes occurred in human mobility, as seen in Fig. 2.

Figure 3(e) displays the T_{max} for the Interlagos station, and Fig. 3(f) exhibits the T_{max} for the Mirante de Santana station. Figures 3(g) and 3(h) show, respectively, to Figs. 3(e) and 3(f), the corresponding $t_0(\text{DC})$ and $t_0(\rho)$, as a function of the lags, calculated for DC and ρ between the 14-day time window of T_{max} and NC from Fig. 1(b). The yellow and blue backgrounds and rectangles have the same meaning as in Fig. 2. Significant values for DC occur at lags (–1, → –5, –8, → –12, –15, → –25) in Fig. 3(g). The corresponding yellow and blue rectangles are shown in Fig. 3(e). The pronounced significant value at lag (–17) is related to an increasing

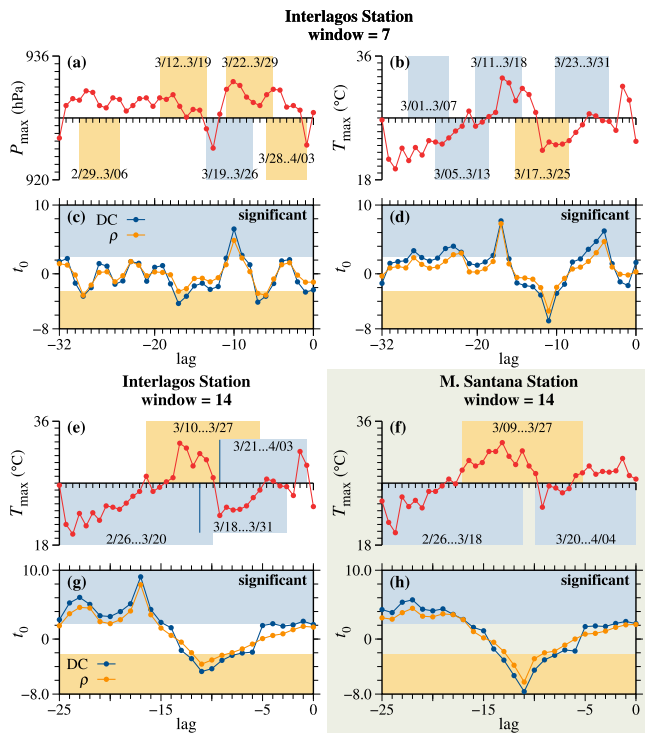


FIG. 3. Shown are data for (a) and (b) P_{\max} and T_{\max} , respectively, for the Interlagos station, (e) T_{\max} for the Interlagos station, and (f) T_{\max} for the Mirante de Santana station. The time interval corresponds to Fig. 1(b) in the early stage of the disease. While panels (c) and (d) display $t_0(\text{DC})$ and $t_0(\rho)$ using a 7-day time window, panels (g) and (h) present results for a 14-day time window [see Fig. 1(b)].

T_{\max} in the interval 4–18 of March, as seen inside the light blue rectangle in Fig. 3(e). In Fig. 3(h), significant values for DC occur at lags $(-1, -2, -8 \rightarrow -13, -17 \rightarrow -25)$. The pronounced negative significant value occurs at lag (-11) , which corresponds to decreasing T_{\max} from 10 to 24 of March, inside the yellow rectangle in Fig. 3(f). These results confirm what has been seen for the 7-day time window, that increasing (decreasing) values of T_{\max} induce significant correlations (anti-correlations) between the maximal temperature and NC.

Now, we compare Figs. 3(d) and 3(g) for which the data for the T_{\max} are the same and solely the time window is changed from 7 to 14, respectively. We observe that the pronounced significant value at lag (-4) in Fig. 3(d) for the 7-day time window is barely significant in the 14-day time window of Fig. 3(e). Something similar occurs, with less intensity, however, at lag (-11) in Figs. 3(d) and 3(g). The peak in the latter is almost not significant anymore, which means that significant values of DC at lags $(-4, -11)$ are more relevant to the 7-day window. Furthermore, the prominent significant values of DC and ρ at lag (-17) seem to be independent of the size of the time window used. Thus, the increasing values of T_{\max} around March 11–18 seem to have strong positive (and size-window stable) correlations, independent of the used window size.

C. Human mobility vs temperature and pressure

Here, we compare results obtained in Secs. III A and III B. For this, we focus on the most significant values of DC at specific lags. One example is the lag (-17) for which prominent $t_0(\text{DC})$ values are seen in Figs. 2(j) and 2(k) and in Figs. 3(d) and 3(g). In Fig. 3(b), this corresponds to the time window from March 11 to 18 and in Fig. 3(e) to the time window from March 6 to 20, which are related to a continuous increase of T_{\max} . In Figs. 2(g) and 2(h), this corresponds to the period of relevant changes in recreation and transit stations. Thus, increasing T_{\max} , lesser mobility in recreation and transit stations have all the strongest positive correlations with the increase of NC in the early stages of the disease. Another example occurs for lags around $(-10, -11)$ in Figs. 3(c) and 3(d) and Figs. 2(d), 2(e), and 2(j). For human mobility, this corresponds to larger changes in grocery and pharmacy, parks, and recreation and is related to a sudden increase in P_{\max} [see the light blue rectangle from March 19 to 26 in Fig. 3(a)].

IV. THE PERIOD UNTIL JUNE 28 (FIXED LAG)

In this section, instead of keeping the NC window fixed at the early stage of the disease [Fig. 1(b)] and changing the lag, we keep fixed the lag and scan from February 26 to June 28. In other words, in the schematic representation of Fig. 1(c), while the window in Q_2 contains the human mobility or meteorological data, the window in Q_1 contains NC data. Thus, we determine $t_0(\text{DC})$ between data inside the left and right windows in Fig. 1(c), keeping fixed the lag and scanning the whole times' axis. Here, we use the 14-day time window exclusively, and the lags vary from -30 to 0 .

For the numerical procedure, we use X for the COVID-19 data and Y for meteorological data or human mobility, meaning $(X, Y) = \{(X_{k,i}, Y_{k-\text{lag},i}) : k = 1, \dots, 14; i = 1, \dots, I\}$, for a fixed lag and k referring to the dates inside the corresponding time window. For better comprehension, we mention two examples: for lag (0) , $k = 1, \dots, 14$ corresponds to dates from February 26 to June 28 and $i = 1, \dots, I (= 50)$ represents the time scanning along the whole analyzed period. For lag (-5) , $k = 1, \dots, 14$, and while the meteorological and mobility windows start on February 26, the NC window starts on March 1 (here $I = 44$).

In this case, meteorological data and human mobility windows intersect for lags from 0 (full intersection) to -13 (1-day intersection).

A. Human mobility

In Fig. 4, $t_0(\text{DC})$ is plotted in colors (see color palette) between human mobility and NC for lags in the interval $[-30, 0]$. While significant values of DC increase from light blue to dark blue for correlations and from yellow to red for anti-correlation, non-significant values of DC decrease from black to gray. For reference, the day of the emergency decree in the city of SP, namely, March 16, is shown by the pink horizontal line.

It is worth mentioning some statistical properties (not shown) of the NC data when scanning the 14-day time windows from February 26 to June 15. While the mean values of NC increase when times go by, the relative variance $\sigma^2(\text{NC})$ has a sudden increase on the 6th scanning day, which remains almost constant (plateau) until the 21st

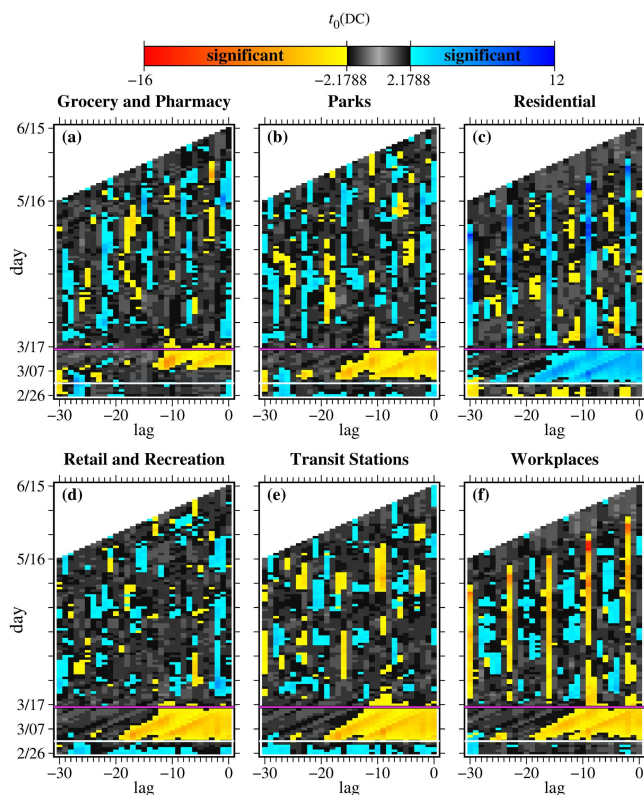


FIG. 4. Values (see color bar) of $t_0(\text{DC})$ between human mobility and NC for the whole considered period and different lags between the 14-day windows. The pink horizontal line marks the day of the emergency decree in the city of SP, namely, March 16. The white dashed (continuous) horizontal line marks the day March 2 for which the 14-day time window of the human mobility does not include the decree day. Panels (a) grocery and pharmacy, (b) parks, (c) residential, (d) retail and recreation, (e) transit stations, and (f) workplaces.

scanning day. This plateau will become relevant later on. A second sudden increase of the variance of NC occurs by the 22nd scanning day.

To make the presentation easy, we separate the discussion of Fig. 4 into three periods: (1) *Before* the data of the human mobility that includes the emergency decree day. This means that the window starts for days below the white horizontal line; (2) *During* the days for which the data of the human mobility includes the decree day. Window starts between pink and white horizontal lines; and (3) *After* data of human mobility, which include the decree day. Above the pink horizontal line.

1. Period of human mobility data before including the health emergency decree (February 26–March 2)

For lags around $(-1, 0)$ and days from February 26 to March 10, only residential and workplaces do not have significant values in DC, and all other human mobilities display positive correlations. Subsequent significant values of DC occur at lags around $(-3, -2)$,

but now for residential, transit stations, and workplaces. For lag (-2) , the scanned NC 14-day windows are located from February 28 to March 12, and for lag (-3) , they are located from February 29 to March 13. The general tendency is that for increasing lags, the mean number of NC inside the 14-day windows increases. More specifically, from lag (-1) to lags (-2) , data show that the number of NC increases by around 400% (not shown). This means that human mobilities in residential, transit stations, and workplaces are somehow correlated with the abrupt increase of NC in these 2 days.

Successive significant correlations along the horizontal are observed from lags (-21) to (-6) for parks, recreation, and transit stations, which match with the plateau of $\sigma^2(\text{NC})$ mentioned above, suggesting that the larger relative variances of the NC data are correlated with these human mobilities. The scanned NC 14-day windows for these lags range from March 2–15 to March 23 to April 5. Positive correlations are essentially the increasing values of NC compared to the increasing mean mobilities in parks, recreation, and transit stations. No significant values relative to pharmacy and workplaces are seen inside the plateau mentioned above. Some anti-correlations are observed for residential at lags (-21) to (-10) (inside the plateau), a clear consequence of the increasing NC values and decreasing Residential mobility. Finally, correlations are also observed for lags around (-28) and (-25) for pharmacy, parks, recreations, and transit stations. The NC windows range from March 8–21 to 11–24. Furthermore, lags from (-29) to (-30) display significant values for pharmacy and residential with anti-correlations and transit stations and workplaces with positive correlations.

2. Period of human mobility data that includes the health emergency decree day (March 3–16)

Here, the period for which Fig. 4 contains the largest concentrated areas with significant values of DC for all human mobility, going essentially from lags $(-19 \rightarrow 0)$ and starting on March 3. The exception occurs for grocery and pharmacy, for which the area is a bit smaller, occurring for lags $(-13 \rightarrow 0)$ and starting later, around March 7. The delayed correlation for pharmacy is a consequence of the later reaction of this human mobility to the emergency decree, as observed in Fig. 2(a). Only residential presents a positive correlation with the NC, and all other human mobilities are anti-correlated to the increasing values of NC as time goes on. These results demonstrate the crucial role of the emergency decree in affecting the NC.

3. Period after the human mobility data includes the health emergency decree day ($> \text{March 16}$)

In general, for later days that do not include the emergency decree, Fig. 4 demonstrates (above the pink line) the existence of many cases of (anti-)correlations for specific dates and lags, which we will not discuss in full detail here. However, some vertical stripes with significant values of DC are visible. To be more specific, stripes with positive correlations in Figs. 4(c) and stripes with negative correlations in 4(f). These stripes appear in intervals of 7 days, at lags $(-2, -9, -16, -23, -30)$. The most relevant stripes occur at the lag (-2) , while the stripes at the other lags are a consequence of

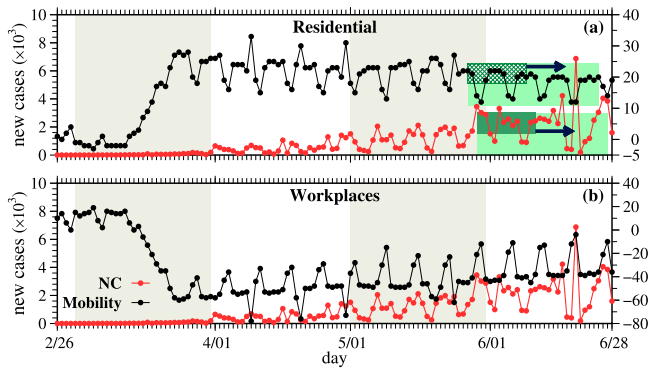


FIG. 5. (a) Residential and (b) workplace mobility changes and NC for the whole analyzed period. Light green boxes indicate the period for which the stripe at lag (-2) has significant values of $t_0(\text{DC})$. In other words, the dark green boxes (hatched for the mobilities) with fixed lag (-2) are the 14-day analyzed time windows with the strongest significant values of $t_0(\text{DC})$ inside the period determined by the light green boxes. This period corresponds to the whole stripe observed in Figs. 4(c) and 4(f) for lag (-2) . The arrows indicate that the dark green boxes move inside the light green boxes.

weekly periodicity in the data. At the lag (-2) , the largest significant values of DC occur around May 29–30, being positive (dark blue) for residential and negative (dark red) for workplaces. These most significant values of $t_0(\text{DC})$ are essentially repeated at the lags $(-9, -16, -30)$ for May 20, 10, and 2, respectively, meaning that something relevant occurred inside the 14-day time windows of residential workplaces around May 29–30.

For better visualization, the whole period of the data for the corresponding mobility changes is plotted in Fig. 5(a) for residential and Fig. 5(b) for workplaces, together with NC. The green boxes with fixed lag (-2) are the 14-day analyzed time windows with the largest significant values of DC, namely, May 29, mentioned above.

B. Influence of maximal temperature

Figure 6 displays the $t_0(\text{DC})$ between the 14-day time window of T_{max} and NC for lags varying in the interval $(-30, 0)$ for stations Interlagos in Fig. 6(a) and Mirante de Santana in Fig. 6(b). While significant positive values of DC increase from light to dark blue, negative values increase from yellow to red. Not-significant values of DC decrease from black to gray. We observe that regions with significant positive values of DC in Fig. 6 (large blue region) occur for lags from $(-26 \rightarrow 0)$ and days from February 26 to March 7. In Figs. 3(e) and 3(f), this corresponds to the increasing mean values of T_{max} inside the 14-day time window from February 26 to March 7. Furthermore, significant negative values of DC are seen at lag $(-12 \rightarrow -9)$ and $(-6 \rightarrow 0)$ for the days from March 7 to 16 and are a consequence of the decreasing mean values of T_{max} in Figs. 3(e) and 3(f). Additional positive significant values are observed for lags $(-6 \rightarrow 0)$ and days around March 17–22. This corresponds to increasing mean values of T_{max} inside the 14-day time window in Figs. 3(e) and 3(f). These cases essentially say that increasing/decreasing values of T_{max} lead to significant correlations to the increasing values of NC as time goes on. In general,

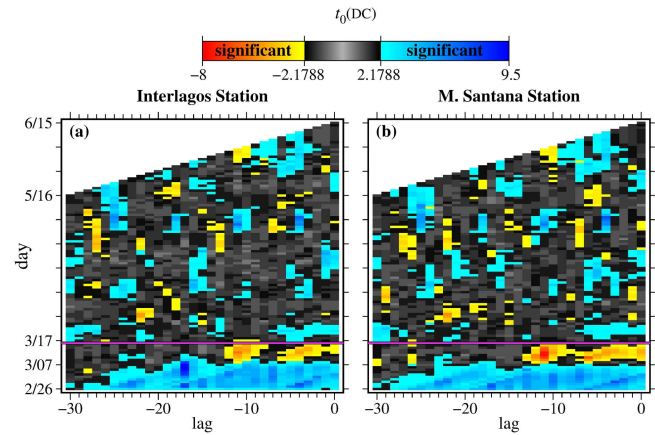


FIG. 6. For the period February 26 to June 15, this figure displays $t_0(\text{DC})$ in colors between T_{max} and NC for lags varying in the interval $(-30, 0)$ for (a) Interlagos and (b) Mirante de Santana station.

many other significant (anti-) correlation values are observed in Figs. 6(a) and 6(b) and occur approximately for the same days and lags, independent of the chosen station.

V. CONCLUSIONS

This paper uses the distance correlation (DC) and the Pearson correlation (ρ) to study the influence of meteorological and human mobility data on the number of new cases (NC) of COVID-19 infected individuals in the city of São Paulo, Brazil. For this analysis, we use the period between 26th February and 28th June of 2020, related to the early stage of the disease in this city. Meteorological data from the Interlagos and Mirante de Santana stations are used in this study. Results are presented only for the most relevant results, obtained after doing an extensive numerical analysis calculating Student's t-distribution $t_0(\text{DC})$ (significant values of DC), and $t_0(\rho)$ (significant values of ρ), between quantities like NC, TC, ND, and TD, related to the spread of the COVID-19, with human mobility data and meteorological variables like radiation, precipitation, wind, and maxima, minima, and mean values of the temperature, pressure, and humidity. We focused our presentation on the temporal relation of DC and ρ between meteorological and human mobility data and NC considering two distinct situations. The first one, presented in Sec. III, analyzes the effects in the early stages of the disease in the city of São Paulo. In the second situation, presented in Sec. IV, we scanned the whole analyzed period keeping a constant lag between the windows. Data are analyzed inside windows of 7 and 14 days.

In Sec. III, results are shown in Fig. 2 for the human mobility and in Fig. 3 for meteorological data T_{max} and P_{max} . Combining human mobility and T_{max} and P_{max} , we found two relevant situations: (a) increasing (decreasing) values of T_{max} induce significant correlations (anti-correlations) between the maximal temperature and NC; (b) most prominent significant values of DC and ρ occur at lag (-17) , meaning that fewer movements in recreation and transit stations, and increasing values T_{max} , have strong correlations with the early stage of the disease; (c) at lags $(-10, -11)$, we observe that

larger changes in grocery and pharmacy, parks, and recreation, and sudden change in P_{\max} are very correlated. Social behavior mostly affects the earlier stage of the disease after 8 to 17 days. In the case of Sec. IV, we use a 14-day time window, keep the lag fixed and scan the whole analyzed period of 50 days, from February 26 to June 28. Relevant contributions are observed essentially for lags between (0) and (−19) in the early stage of the disease. Days March 8–18 are affected by the human mobility changes, and days February 26 to March 15 are affected by P_{\max} and T_{\max} . Finally, the emergency decree for human mobility and increasing/decreasing values of T_{\max} have shown to present the most significant correlations with the NC.

ACKNOWLEDGMENTS

A.M.G. and M.W.B. thank National Council of Research (CNPq, Brazil) for financial support (Grant Nos. 316104/2021-3 and 310294/2022-3, respectively). E.L.B. is grateful for the financial support from the São Paulo Research Foundation (FAPESP) under Grant Nos. 2021/12232-0 and 2018/03211-6.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Carlos F. O. Mendes: Conceptualization (equal); Formal analysis (lead); Investigation (lead); Methodology (equal); Software (equal); Validation (equal); Writing – review & editing (supporting). **Eduardo L. Brugnago:** Conceptualization (equal); Formal analysis (lead); Investigation (lead); Methodology (equal); Software (equal); Validation (equal); Writing – review & editing (equal). **Marcus W. Beims:** Conceptualization (equal); Investigation (lead); Methodology (equal); Supervision (lead); Validation (equal); Writing – original draft (lead); Writing – review & editing (lead). **Alice M. Grimm:** Conceptualization (equal); Data curation (equal); Investigation (lead); Methodology (equal); Writing – review & editing (supporting).

DATA AVAILABILITY

The data that support the findings of this study are available within the article.

APPENDIX: CORRELATIONS AND SIGNIFICANT VALUES

In this appendix, we show computational methodology to determine the statistical measures of the distance correlation²⁶ and the Pearson correlation⁴⁰ and also the procedure for the significance test.

Distance correlation (DC). Consider a joint random sample $(X, Y) = \{(X_k, Y_k) : k = 1, \dots, N\}$ with $X \in \mathbb{R}^p$ and $N \geq 2$, the matrix, with $i, j = 1, \dots, N$, is defined by

$$A_{ij} = a_{ij} - \bar{a}_i - \bar{a}_j + \bar{a}_{..}, \tag{A1}$$

where $a_{ij} = |X_i - X_j|_p$ is the Euclidean distance; $\bar{a}_i = \frac{1}{N} \sum_{j=1}^N a_{ij}$ and $\bar{a}_j = \frac{1}{N} \sum_{i=1}^N a_{ij}$ are the arithmetic mean of the rows and columns, respectively; and general mean is defined by $\bar{a}_{..} = \frac{1}{N^2} \sum_{i,j=1}^N a_{ij}$. Similarly, with $Y \in \mathbb{R}^q$, matrix B_{ij} is obtained.

The empirical distance covariance for a joint sample (X, Y) is defined by

$$\sigma_N(X, Y) = \frac{1}{N} \left(\sum_{i,j=1}^N A_{ij} B_{ij} \right)^{1/2}. \tag{A2}$$

The empirical distance variance for a sample X is

$$\sigma_N(X) = \frac{1}{N} \left(\sum_{i,j=1}^N A_{ij}^2 \right)^{1/2} \tag{A3}$$

and for Y ,

$$\sigma_N(Y) = \frac{1}{N} \left(\sum_{i,j=1}^N B_{ij}^2 \right)^{1/2}. \tag{A4}$$

The distance correlation coefficient is given by

$$DC_N(X, Y) = \frac{\sigma_N(X, Y)}{\sqrt{\sigma_N(X)}\sqrt{\sigma_N(Y)}}, \tag{A5}$$

with DC varying in the interval [0, 1], but we consider between −1 and +1 with $DC = \text{sgn}(\rho)DC$. For simplicity, let us consider the notation $DC_N(X, Y) = DC$.

Pearson correlation (ρ). Assuming we have a sample $(X, Y) = \{(X_k, Y_k) : k = 1, \dots, N\}$ with $N \geq 2$, the statistical measure can be calculated using the expressions

$$S_{XX} \equiv \sum_{k=1}^N (X_k - \bar{X})^2, \tag{A6}$$

$$S_{YY} \equiv \sum_{k=1}^N (Y_k - \bar{Y})^2, \tag{A7}$$

and

$$S_{XY} \equiv \sum_{k=1}^N (X_k - \bar{X})(Y_k - \bar{Y}). \tag{A8}$$

The Pearson correlation coefficient can be obtained by replacing (A6)–(A8) in the expression

$$\rho_N(X, Y) = \frac{S_{XY}}{\sqrt{S_{XX}S_{YY}}}, \tag{A9}$$

resulting in

$$\rho_N(X, Y) = \frac{\sum_{k=1}^N (X_k - \bar{X})(Y_k - \bar{Y})}{\sqrt{\sum_{k=1}^N (X_k - \bar{X})^2 \cdot \sum_{k=1}^N (Y_k - \bar{Y})^2}}. \tag{A10}$$

Here, $\bar{X} = \frac{1}{N} \sum_{k=1}^N X_k$ and $\bar{Y} = \frac{1}{N} \sum_{k=1}^N Y_k$ are the arithmetic means of the samples.^{32,41} Its values belong to the range [−1, +1]. For simplicity, let us consider the notation $\rho_N(X, Y) = \rho$.

Student's t-test (t_0). Now follows the procedure for the significance test:

- Choose a critical value for the significance, $\alpha = 0.05$.
- As the sample (X, Y) contains N pairs of data, we consult a table of Student's t-distribution^{42,43} to obtain the value of $t(df)$ for the chosen α value, where $df = N - 2$ are the degrees of freedom. As shown in Sec. II, we use 7- and 14-day time windows. Therefore, we have, respectively,

$$df = 7 - 2 = 5 \Rightarrow t(5) = 2.5706,$$

$$df = 14 - 2 = 12 \Rightarrow t(12) = 2.1788.$$

- Calculate t_0 for DC or ρ by the relations^{44,45}

$$t_0(DC) = DC \sqrt{\frac{N-2}{1-DC^2}} \quad \text{or} \quad t_0(\rho) = \rho \sqrt{\frac{N-2}{1-\rho^2}}. \quad (\text{A11})$$

- If $t_0 > t(df)$ or $t_0 < -t(df)$, the coefficients are considered significant. Thus, we consider $t_0 > 2.5706$ or $t_0 < -2.5706$ and $t_0 > 2.1788$ or $t_0 < -2.1788$ for the corresponding 7-day and 14-day time windows, respectively.

REFERENCES

- J. Wang, K. Tang, K. Feng, X. Lin, W. Lv, K. Chen, and F. Wang (2020), see <https://doi.org/10.2139/ssrn.3551767> for "High temperature and high humidity reduce the transmission of COVID-19."
- R. Alvarez-Ramirez and M. Meraz, "Role of meteorological temperature and relative humidity in the January-February 2020 propagation of 2019-nCoV in Wuhan, China," medRxiv (2020). <https://doi.org/10.1101/2020.03.19.20039164>.
- M. Bannister-Tyrrell, A. Meyer, C. Faverjon, and A. Cameron, "Preliminary evidence that higher temperatures are associated with lower incidence of COVID-19, for cases reported globally up to 29th February 2020," medRxiv (2020). <https://doi.org/10.1101/2020.03.18.20036731>.
- Q. Bukhari and Y. Jameel (2020), see <http://dx.doi.org/10.2139/ssrn.3556998> for "Will coronavirus pandemic diminish by summer?"
- Y. Ma, Y. Zhao, J. Liu, X. He, B. Wang, S. Fu, J. Yan, J. Niu, J. Zhou, and B. Luo, "Effects of temperature variation and humidity on the mortality of COVID-19 in Wuhan," medRxiv (2020). <https://doi.org/10.1101/2020.03.15.20036426>.
- M. Sajadi, P. Habibzadeh, A. Vintzileos, S. Shokouhi, F. Miralles-Wilhelm, and A. Amoroso (2020), see <http://dx.doi.org/10.2139/ssrn.3550308> for "Temperature, humidity and latitude analysis to predict potential spread and seasonality for COVID-19."
- J. Wang, A. Jiang, L. Gong, L. Luo, W. Guo, C. Li, J. Zheng, C. Li, B. Yang, J. Zeng, Y. Chen, K. Zheng, and H. Li, "Temperature significant change COVID-19 transmission in 429 cities," medRxiv (2020). <https://doi.org/10.1101/2020.02.22.20025791>.
- M. Araujo and B. Naimi, "Spread of SARS-CoV-2 coronavirus likely to be constrained by climate," medRxiv (2020). <https://doi.org/10.1101/2020.03.12.20034728>.
- T. Carleton and K. Meng, "Causal empirical estimates suggest COVID-19 transmission rates are highly seasonal," medRxiv (2020). <https://doi.org/10.1101/2020.03.26.20044420>.
- R. Ficetola and D. Rubolini, "Climate affects global patterns of COVID-19 early outbreak dynamics," medRxiv (2020). <https://doi.org/10.1101/2020.03.23.20040501>.
- N. Islam, Q. Bukhari, Y. Jameel, S. Shabnam, A. Erzurumluoglu, M. A. Siddique, J. M. Massaro, and R. B. D'Agostino, "Covid-19 and climatic factors: A global analysis," *Environ. Res.* **193**, 110355 (2021).
- J. M. Loché Fernández-Ahúja and J. L. Fernández Martínez, "Effects of climate variables on the COVID-19 outbreak in Spain," *Int. J. Hygiene Environ. Health* **234**, 113723 (2021).
- A. B. Chelani and S. Gautam, "The influence of meteorological variables and lockdowns on COVID-19 cases in urban agglomerations of Indian cities," *Stochast. Environ. Res. Risk Assess.* **36**, 2949 (2022).
- S. Shetty, A. Gawade, S. Deolekar, V. Patil, R. Pandharkar, and U. Salunkhe, "Impact of key meteorological parameters on the spread of COVID-19 in Mumbai: Correlation and regression analysis," medRxiv (2022). <https://doi.org/10.1101/2022.02.22.22271376>.
- S. A. D. Souza Júnior, P. V. C. D. Freitas, Y. D. V. Hamberger, L. W. Bisol, and F. G. D. M. Souza, "Low temperature, high relative humidity and higher precipitation are associated with a higher number of deaths from COVID-19," *Res. Soc. Dev.* **11**, e49111427616 (2022).
- A. Briz-Redón and A. Serrano-Aroca, "The effect of climate on the spread of the COVID-19 pandemic: A review of findings, and statistical and modelling techniques," *Prog. Phys. Geogr. Earth Environ.* **44**, 591–604 (2020).
- S. Ogunjo, O. Olaniyan, C. F. Olusegun, F. Kayode, D. Okoh, and G. Jenkins, "The role of meteorological variables and aerosols in the transmission of COVID-19 during Harmattan season," *GeoHealth* **6**, e2021GH000521 (2022).
- O. Damette, C. Mathonnat, and S. Goutte, "Meteorological factors against COVID-19 and the role of human mobility," *PLoS One* **16**, e0252405 (2021).
- H. Wei, S. Liu, Y. Liu, B. Liu, and X. Gong, "The impact of meteorological factors on COVID-19 of California and its lag effect," *Meteorol. Appl.* **29**, e2045 (2022).
- K. O'Reilly, M. Auzenbergs, Y. Jafari, Y. Liu, S. Flasche, and R. Lowe, see <https://cmimid.github.io/topics/covid19/current-patterns-transmission/role-of-climate.html> for "Effective transmission across the globe: The role of climate in COVID-19 mitigation strategies" (2020).
- T. Jamil, I. Alam, T. Gojobori, and C. Duarte, see <https://doi.org/10.1101/2020.03.29.20046706> for "No evidence for temperature-dependence of the COVID-19" (2020).
- B. Chen, H. Liang, X. Yuan, Y. Hu, M. Xu, Y. Zhao, B. Zhang, F. Tian, and X. Zhu, "Roles of meteorological conditions in COVID-19 transmission on a worldwide scale," medRxiv (2020). <https://doi.org/10.1101/2020.03.16.20037168>.
- B. Yang, K. Shang, M. Small, and C. Naipeng, see <https://doi.org/10.1360/nso/20220051> for "Information overload: How hot topics distract from news—COVID-19 spread in the USA" (2022).
- M. Small and D. Cavanagh, "Modelling strong control measures for epidemic propagation with networks—A COVID-19 case study," *IEEE Access* **8**, 109719 (2022).
- C. Granger and P. Newbold, *Forecasting Economic Time Series* (Academic Press, New York, 1977), p. 225.
- G. J. Székely, M. L. Rizzo, and N. K. Bakirov, "Measuring and testing dependence by correlation of distances," *Ann. Stat.* **35**, 2769–2794 (2007).
- C. F. O. Mendes and M. W. Beims, "Distance correlation detecting Lyapunov instabilities, noise-induced escape times and mixing," *Phys. A* **512**, 721–730 (2018).
- C. F. O. Mendes, R. M. da Silva, and M. W. Beims, "Decay of the distance autocorrelation and Lyapunov exponents," *Phys. Rev. E* **99**, 062206 (2019).
- C. F. O. Mendes, L. Menon, R. M. da Silva, and M. W. Beims, "Detecting correlations in desynchronized chaotic chimera states," *Int. J. Bifurcation Chaos* **32**, 2230012 (2022).
- C. Manchein, E. L. Brugnago, R. M. da Silva, C. F. O. Mendes, and M. W. Beims, "Strong correlations between power-law growth of COVID-19 in four continents and the inefficiency of soft quarantine strategies," *Chaos* **30**, 041102 (2020).
- R. M. da Silva, C. F. O. Mendes, and C. Manchein, "Scrutinizing the heterogeneous spreading of COVID-19 outbreak in large territorial countries," *Phys. Biol.* **18**, 025002 (2021).
- E. Martínez-Gómez, M. T. Richards, and D. S. P. Richards, "Distance correlation methods for discovering associations in large astrophysical databases," *Astrophys. J.* **781**, 39 (2014).
- A. Bhattacharjee, "Distance correlation coefficient: An application in Bayesian approach in clinical data analysis," *J. Mod. Appl. Stat. Methods* **13**, 354–366 (2014).
- L. Geerligs, Cam-CAN and R. N. Henson, "Functional connectivity and structural covariance between regions of interest can be measured more accurately using multivariate distance correlation," *NeuroImage* **135**, 16–31 (2016).

- ³⁵J. R. Ayala Solares and H.-L. Wei, “Nonlinear model structure detection and parameter estimation using a novel bagging method based on distance correlation metric,” *Nonlinear Dyn.* **82**, 201–215 (2015).
- ³⁶J. Runge, P. Nowack, M. Kretschmer, S. Flaxman, and D. Sejdinovic, “Detecting and quantifying causal associations in large nonlinear time series datasets,” *Sci. Adv.* **5**, 1 (2019).
- ³⁷See <https://www.google.com/covid19/mobility/> for “COVID-19 community mobility reports” (2021).
- ³⁸See <https://bigdata-api.fiocruz.br/relatorios/> for “FIOCRUZ (2021).
- ³⁹M. da Saúde, see <https://covid.saude.gov.br/> for “Painel Coronavirus” (2021).
- ⁴⁰K. Pearson, “Note on regression and inheritance in the case of two parents,” *Proc. R. Soc. London* **58**, 240–242 (1895).
- ⁴¹P. F. Dunn, *Measurement, Data Analysis, and Sensor*, 2nd ed. (CRC Press, 2010).
- ⁴²T.-W. Hsu, M. Usman, Y.-J. Lan, and Y.-P. Lee, “A study of extreme value analysis on typhoon wave,” *Coast. Eng. Proc.* **1**, waves.38 (2014).
- ⁴³S. L. Jackson, *Research Methods and Statistics: A Critical Thinking Approach* (Wadsworth, 2006).
- ⁴⁴A. D. Hutson, “A robust Pearson correlation test for a general point null using a surrogate bootstrap distribution,” *PLoS One* **14**, 1–14 (2019).
- ⁴⁵E. I. Obilor and E. C. Amadi, “Test for significance of Pearson’s correlation coefficient,” *Int. J. Innovative Math. Stat. Energy Policies* **2**, 11–23 (2018); available at <https://seahipaj.org/journals/engineering-technology-and-environment/ijimsep/vol-6-issue-1/>.