Investigating spatiotemporal patterns of the COVID-19 in São Paulo State, Brazil

Enner Alcântara,1 José Mantovani,1 Luiz Rotta,2 Edward Park,3 Thanan Rodrigues,4 Fernando Campos Carvalho,1 Carlos Roberto Souza Filho5

1Institute of Science and Technology, São Paulo State University, São José dos Campos, Brazil; 2Faculty of Sciences and Technology, São Paulo State University, Presidente Prudente, São Paulo, Brazil; 3National Institute of Education, Nanyang Technological University, Singapore; 4Federal Institute of Education, Science and Technology of Pará State, Castanhal, Brazil; 5Institute of Geosciences, University of Campinas, Campinas, São Paulo, Brazil

Abstract

As of 16 May 2020, the number of confirmed cases and deaths in Brazil due to COVID-19 hit 233,142 and 15,633, respectively, making the country one of the most affected by the pandemic. The State of São Paulo (SSP) hosts the largest number of confirmed cases in Brazil, with over 60,000 cases to date. Here we investigate the spatial distribution and spreading patterns of COVID-19 in the SSP by mapping the spatial autocorrelation and the clustering patterns of the virus in relation to the population density and the number of hospital beds. Clustering analysis indicated that São Paulo City is a significant hotspot for both the confirmed cases and deaths, whereas other cities across the state were less affected.

Bivariate Moran’s I showed a low relationship between the number of deaths and population density, whereas the number of hospital beds was less related, implying that the fatality depends substantially on the actual patients’ conditions. Multivariate Local Geary showed a positive relationship between the number of deaths and population density, with two cities near São Paulo City being negatively related; the relationship between the number of deaths and hospital beds availability in the São Paulo Metropolitan Area was basically positive. Social isolation measures throughout the State of São Paulo have been gradually increasing since early March, an action that helped to slow down the emergence of the new confirmed cases, highlighting the importance of the safe-distancing measures in mitigating the local transmission within and between cities in the state.

Introduction

In December 2019, health units in Wuhan City, Hubei Province, China, reported the occurrence of patients with pneumonia, of unknown cause, possibly epidemiologically associated with a seafood and live animals’ market. On December 31 of that year, the Chinese Center for Disease Control and Prevention (China CDC) sent a team to that province to accompany health authorities in conducting an epidemiological and etiological investigation that identified the cause as due to a new coronavirus named SARS-COV-2 (Zhu et al., 2020). Simultaneously, the World Health Organization (WHO) was informed resulting in the establishment of an Emergency Committee following International Health Regulations (IHR). The first meeting on the issue in question was convened by the WHO on 23 January 2020 discussing the possibility of the occurrence constituting a Public Health Emergency of International Importance (PHEII) (WHO, 2020a).

Given the rapid escalation on the number of cases in China, as well as in other countries, a subsequent meeting held on 30 January 2020, consolidated the PHEII status (WHO, 2020a). On 11 February 2020, with reference to the type of virus and the year it first occurred, WHO assigned the name of the new disease Coronavirus disease - 2019 (COVID-19) (WHO, 2020b). Until the end of February, the number of confirmed cases reached 80,000 in China, with a total of about 2,900 deaths (Li et al., 2020). In other countries the number of infected people during the same period reached 6,000 and close to 90 deaths (Middelburg and Rosendaal, 2020). The rapid spread of COVID-19 has demonstrated that the
international community was not prepared for this biological disaster (Djalante et al., 2020).

In Brazil, the monitoring of cases began in January 2020. On 22 January, the Emergency Operation Center ((Centro de Operações em Emergências, COE) of the Ministry of Health, coordinated by the Health Surveillance Secretariat, was activated to objectively plan and organize activities related to the control of local contamination, as well as international monitoring. A contingency plan was activated on 27 January, and on 3 February the then epidemic (now pandemic) was declared a Public Health Emergency of National Importance (Brazilian Ministry of Health, 2020). The first confirmed cases of COVID-19 in Brazil occurred at the end of February 2020 and were male individuals residing in the city of São Paulo, who had just returned from a trip to Italy. The cases were thus referred to as imported and evidence of local transmission of the virus was not found until the beginning of March 2020. The difficulty in containing the spread of the COVID-19 is that, similar to outbreaks caused by other coronaviruses which cause common cold, transmission can occur from asymptomatic infected cases (Munster et al., 2020). Several studies have been carried out from the initial cases (Zhou et al., 2020; Lam et al., 2020; Lu et al., 2020; Wang et al., 2020) investigating several aspects associated with COVID-19, from development of the outbreak to clinical characteristics.

The effective reproduction number (R) measures an infectious disease’s capacity to spread and it signifies the average number of people that one infected person will pass an infection to. According to the Brazilian Society of Infectious Diseases, in a statement dated 12 March, 2020, the contagion capacity of SARS-COV-2 is 2.74; in comparative terms, the measles rate is 15. Although the aggressive testing of cases nationwide and strong measures imposed by the government, such as social distancing, the number of both confirmed cases and deaths in Brazil grew exponentially throughout the country reaching 233,142 and 15,633, respectively (São Paulo State, 2020b), corresponding to a case fatality rate of 6.7% as of May 17th 2020. Given this situation in Brazil, it is urgent to understand how COVID-19 is spreading in different regions of the country. In this paper, we investigated the spatial distribution patterns and dynamics of the COVID-19 in the State of São Paulo, the most affected state by the pandemic in Brazil.

Materials and methods

Study area

The whole São Paulo State (SSP) was the study site (Figure 1). With approximately 46 million citizens (>20% of the country’s population)
inhabitants), SSP hosts the largest population of all Brazil’s 26 states (Instituto Brasileiro de Geografia e Estatística, IBGE, 2019). It is also the wealthiest state and responsible for 34% of the Brazilian Gross Domestic Product (GDP) and has the second-highest Human Development Index and GDP per capita in Brazil. To date (May 16, 2020), among 233,142 confirmed cases in all regions of Brazil, most of them (41,830) are in SSP (Brazilian Ministry of Health, 2020). The city of São Paulo, the state capital, hosts 26,273 of all confirmed cases and has reported 2,135 deaths at this date (São Paulo State, 2020b).

**Data sources**

The data about COVID-19 (per 100,000 inhabitants), population density and social isolation index in SSP from 1 March to 8 May 2020 was obtained from the São Paulo State Health Secretariat (São Paulo State, 2020b). The social isolation index, which uses geolocation data provided by the nation’s main telecom firms and shows the percentage of the population abiding by available social isolation advice. The data about hospital beds were obtained from the Department for Information on the Brazilian Public Health Care System (Departamento de Informática do Sistema Único de Saúde do Brasil, DATASUS, 2020). The shape file of the federal highways was downloaded from the Brazilian Institute of Geography and Statistics (Instituto Brasileiro de Geografia e Estatística, IBGE, 2020). The data were processed using GeoDa software (Anselin et al., 2006).

**COVID-19 spatial associations**

Spatial autocorrelation statistics measure the degree to which a sub-region is similar to or different from its neighbouring sub-regions for a particular indicator. Spatial autocorrelation may be evaluated globally as well as locally. Global measures summarize spatial autocorrelation, while local measures evaluate localized spatial autocorrelation within a study site. Moran’s I statistic is a measure of global spatial autocorrelation or clustering (Moran, 1950) and can be calculated as:

\[
I = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} (y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} (y_i - \bar{y})^2} \quad (\text{Eq. 1})
\]

where \(n\) is the number of spatial units indexed by \(i\) and \(j\); \(y_i\) and \(y_j\) the variables at points \(i\) and \(j\) (with \(i\neq j\)); \(\bar{y}\) the mean of \(y\); and \(w_{ij}\) an element of the weight matrix \((n \times m)\).

The Moran’s I range from -1 to 1. If \(I > 0\), there is a positive spatial correlation and if \(I < 0\), a negative spatial correlation exists. To test the statistical significance the inference was based on 999 permutations. In GeoDa the randomization is used in the context of permutation to compute the pseudo-significance level (Anselin, 1995). We computed Moran’s I and both \(z\)-score and pseudo-\(p\)-value to evaluate the statistical significance. In general, Moran’s I near +1 indicates clustering, while values near -1 indicates dispersion. The null hypothesis states that there is no spatial clustering of the values associated with the geographical units in the study site. The null hypothesis can be rejected when the pseudo \(p\)-value is small, and the \(z\)-score is large enough that it falls outside of the desired confidence level.

However, spatial correlation patterns may differ locally within the study site; therefore, measures of global spatial autocorrelation do not reflect the local spatial correlations within geographic units once there is a spatial heterogeneity. To overcome this, we used Anselin’s Local Indicators of Spatial Association (LISA), a measure to describe the heterogeneity of spatial association across different geographic units (Anselin, 1995) defined as:

\[
I_i = \frac{1}{n} \sum_{j=1}^{n} W_{ij} (y_i - \bar{y})(y_j - \bar{y}) \quad (\text{Eq. 2})
\]

where \(j\) is within distance of \(i\) and \(I\), refers to the value of local Moran’s I at points \(i\). The parameters \(y_i\), \(y_j\), \(\bar{y}\) and \(w_{ij}\) have the same meaning as in Eq. 1.

The spatial autocorrelations between COVID-19 and the number of deaths and hospital beds availability and population density were tested using the bivariate Moran’s I statistics. The global Moran’s I describe the overall spatial relationship across all geographic units for the entire study site and can be calculated as:

\[
I_B = \frac{N \sum_{i=1}^{N} \sum_{j=1}^{N} W_{ij} z_i z_j}{(N-1) \sum_{i=1}^{N} \sum_{j=1}^{N} W_{ij}} \quad (\text{Eq. 3})
\]

where \(I_B\) is the bivariate global spatial correlation index; \(N\) the total number of geographic units; \(W_{ij}\) the spatial weight matrix; \(z_i^*\) the standardized number of COVID-19 cases or deaths; and \(z_j^*\) standardization of hospital bed availability, human development index by cities and population density. The bivariate Moran’s I captures the relationship between the value for one variable at a location and the average of the neighbouring values for another variable, therefore the result is not a correlation between two variables. To measure the extent to which neighbours in multi-attribute space are also neighbours in geographically we used the Multivariate Local Geary (MLG) according to Anselin. The MLG measures the extent to which the average distance in attribute space between the values at a location and the values at its neighbouring locations are smaller or larger than what they would be under spatial randomness. Small distance corresponds to a positive spatial autocorrelation and larger corresponds to a negative spatial autocorrelation (Anselin, 2018). The MLG can be calculated as:

\[
C_i = \sum_{h=1}^{m} w_{ij} (x_{hi} - x_{hj})^2 \quad (\text{Eq. 4})
\]

where \(x_i\) is an observation on the variable of interest at location \(i\); \(w_{ij}\) the elements of the spatial weight matrix; and \(m\) variables indexed by \(h\).

At least seven cities in SSP instituted measures to prevent the interchange between local residents and tourists, only allowing movement of locals. This restriction only allowed people working in essential services, such as medical or transportation, to enter the São Paulo Metropolitan Area (SPMA). The other communities along the coast maintained the policy of social isolation, limiting hotel services and prohibiting loitering on the beaches. The state authority let it be known that the entire state health system would collapse if the number of cases on coast and inland continues to grow in the same way as in the state capital. Due to the highly heterogeneous distribution of COVID-19 cases observed in SSP, the Moran’s I method was applied to map the level of clustering of the cases around the cities that were socio-economically intimately connected.
Results

COVID-19 confirmed cases spatial distribution

The first Brazilian COVID-19 confirmed case was found in the city of São Paulo and the case was reported on 26 February 2020. Two weeks later, on 11 March, the WHO declared that the COVID-19 a pandemic. Figure 2 shows the spread throughout the SSP of COVID-19 confirmed cases from 1 March to 8 May 2020. From one case the number had jumped to 722 in just 24 days. Although the SSP Government declared partial lockdown on 24 March 2020, the number of confirmed cases had reached 24,273 by 8 May 2020.

The highest numbers of confirmed cases were registered in the SPMA and the lowest for the littoral cities and other communities. Regarding the São Paulo coast, most cases came from cities and towns shielded from tourists, who typically frequent the beaches on holidays and weekends.

COVID-19 spatial autocorrelation

The calculated Moran’s I values were positive (0.394 and 0.211; P-value = 0.05) for both confirmed cases and deaths (per 100,000 inhabitants) indicating a positive spatial autocorrelation. This indicates that there are similarities between the values of confirmed cases and deaths with regard to the spatial location (Figure 3). In Figure 3a and c, the LISA maps of statistical significance show the regions having passed the local Moran’s significance test for both confirmed cases and deaths in SSP. Based on that, Figure 3b highlights the main clusters of the state’s confirmed cases, where the High-High class (red) indicates regions with high numbers of confirmed cases relative to the average, surrounded by regions that also present high numbers. This class is represented by SPMA, the most populous region of the state, which incidentally also is the place where the first cases were reported. In contrast, the Low-Low class (blue) refers to a spatial association group with a low number of confirmed cases, i.e., below the average, surrounded by regions that also have low values. Compared to the state capital, these regions are represented by small towns and the places where the case reporting took a longer time (see the spatial distribution in

Figure 2. Spatial distribution of COVID-19 confirmed cases in São Paulo State.
Figure 2). The Low-High class (purple) in the figure are regions with low numbers of confirmed cases surrounded by regions with high numbers of confirmed cases. Although there are confirmed cases in the western region of the state, the mortality records are concentrated in the SPMA, where the number of confirmed cases also is high (Figure 3d).

**Relationship between COVID-19 deaths and population density and hospital beds**

The bivariate Moran’s I result show that a positive relationship between deaths (per 100,000 inhabitants) and population density, i.e. 0.283 (P-value = 0.05), indicating a low relationship. On the other hand, the number of hospital beds yielded I = -0.010 (P-value = 0.05), implying that there is no relation with the number of deaths (Figure 4).

The relationship between the number of deaths versus population density (Figure 4b) and deaths versus hospital beds (Figure 4d) shows a High-High cluster (mainly in the SPMA), surrounded by a Low-Low spatial association. The High-High means that this city has a high number of deaths and high population density, as well as a large number of hospital beds available (Figure 5). The Low-Low means the opposite. Cities classified as High-Low and Low-High indicate, respectively, that they had an increase of deaths with low population density and a small number of hospital beds, or a smaller death number with a high number of hospital beds. In this particular cluster, the city is in the transition between spatial regimes (e.g., moving from a low contamination scenario to a high contamination one or vice versa), i.e. these areas are experiencing higher death rates; alternatively, the hospital infrastructure and the social isolation is working.

**Multi-attribute space analysis**

The MLG analysis between deaths and population density presented 519 cities/towns without significance in a total of 645; 67 of which were significant with a pseudo p-value of 0.05, 31 with a p-value of 0.01 and 28 with a P-value of 0.001 (Figure 6a). As seen in Figure 6b, a total of 118 cities/towns presented positive (smaller distance at a location and the values at its neighbouring locations) and 9 negatives (larger distances at a location and the values at its neighbouring locations). On the other hand, the MLG between deaths and hospital beds availability presented 536 cities that were not statistically significant, 57 of which with a pseudo p-value of 0.05, 25 with a P-value of 0.01 and 27 with a p-value of 0.001 (Figure 6c). As seen in Figure 6d, a total of 82 places presented positive and 28 negatives.
Evolution of social isolation

The maps of social isolation in the cities/towns computed with data collected on 5 and 12 March 2020 show that the majority reached an isolation fraction of ≤40%. This pattern only varied slightly until 17 March 2020 when isolation started to increase towards 50% for some regions (Figure 7). On 15 April 2020 and 8 May 2020, these maps showed that cities/towns reduced their isolation status from 50-60% to 40-50%. According to São Paulo (2020) since the partial lockdown of the state started on 24 March 2020, social isolation varied from 54% (24 March), achieving a minimum of 47% (9 April) and a maximum of 59% (at several dates) with an average of 54%.

Discussion

In this work, we investigated the spatial distribution and spread patterns of the COVID-19 in the SSP by mapping the spatial autocorrelation and the clustering patterns in relation to the population density and the number of hospital beds. The results on spatial distribution patterns of the COVID-19 cases in the SSP showed a clear aggregation pattern among its cities and towns: a high rate of surrounding contamination with similar levels of infection was seen as well as the opposite: cities and towns with low contamination surrounded by areas characterized by low values. The highest contaminated cluster was in the SPMA and the lowest in the countryside areas; the same pattern was observed for the mortality as seen on the maps.

The lack of spatial autocorrelation between deaths and hospital beds highlights that quantity per se is not directly related to the chances of a patient’s survival. Hospital beds not equipped with ventilators and experienced health care teams may lead to increased death of critical patients. Requia et al. (2020) mapped equipments available to the Brazilian health care system highlighting that the presence of ventilators can decrease the fatality rate and therefore need to be considered as a key parameter to analyze the fatality rate due to the COVID-19.

The social isolation maps showed that isolation increased in cities with the highest populations soon after the partial lockdown but decreased only a few days later. Nakada et al. (2020) showed that the air quality improved during the partial lockdown due to the reduction of cars in the SSP, which corroborates with our results. The same pattern was observed for cities and towns in the countryside.

Figure 4. Statistical significance of according to LISA analysis of the relationship between deaths due to COVID-19 and population density and between these deaths and availability of hospital beds in São Paulo State. The left side of the figure denotes statistical significance for the relationship between deaths and population density and between deaths and availability of hospital beds, and the right side shows the clustering tendency with respect to these relationships. The presentation is based data collected until 8 May 2020.
Figure 5. Deaths by COVID-19, population density and hospital beds in the State of São Paulo.

Figure 6. Statistical significance according to Local Geary analysis of the relationship between deaths due to COVID-19 and population density and between these deaths and availability of hospital beds in São Paulo State. The left side of the figure denotes statistical significance for the relationship between deaths and population density and between deaths and availability of hospital beds, and the right side shows the clustering tendency with respect to these relationships. The presentation is applied to data collected until 8 May 2020.
The SSP presented a positive relationship (MLG) between mortality and population density, with nine cities being negatively related (Cotia, Itapeverica da Serra, Mauá, Maracai, Tarumã, Regente Feijó, Rancharia, Iepê and Ilhabela). On the other hand, the MLG relationship between the number of deaths and hospital beds availability in SPMA was basically positive. However, the population density per se is not a key parameter to understand the COVID-19 mortality in SSP, as the number of hospital beds seems to influence in the number of deaths. Other parameters such as social distancing, personal hygiene and mask use, can help to understand the spatiotemporal COVID-19 variability.

Figure 5 reveals that São Paulo City, the most populated city in the state, had the highest rate of contamination and number of deaths. In contrast, the countryside might be protected from the virus if the state government regulates movement within the state. The importance of the lockdown in curbing the transmission of COVID-19 in SSP has already been shown, as well as in other countries (Achter et al., 2020). It becomes even more important given that the first case in most of the cities and towns were imported (Candido et al., 2020). Social isolation is essential in the initial stage of the spread of the virus, limiting contamination to a certain area (Kraemer et al., 2020). Therefore, it is timely that we develop sustainable safe-distancing measures to mitigate the local transmission within and between cities.

Conclusions

The high number of confirmed cases and deaths as early as of 16 May 2020 (and continuously) makes Brazil rank among the most affected countries by the COVID-19. The SSP contains the largest number of confirmed cases in Brazil, with over 892,000 cases to date (13 September 2020). The analyses carried out indicates that São Paulo City is a significant hotspot for both confirmed infection by COVID-19 and related mortality, while other cities and towns across the state could be considered cold spots. Bivariate Moran’s I showed a low relationship between the number of deaths and population density, while the number of hospital beds showed no such relationship implying that the fatality depends substantially on the actual patient’s condition. MLG showed a positive relationship between the number of deaths and population density, with two cities near São Paulo City being negatively related; the relationship between the number of deaths and hospital beds availability in SPMA was basically positive. The population density per se is not a key parameter to understand the COVID-19 deaths in SSP; the availability of hospital beds seems to be had a bearing on the number of deaths. Social isolation measures throughout the state has been gradually increasing since early March that has helped to slow down the emergence of new confirmed cases highlighting the importance of safe-distancing measures in mitigating local transmission within and between areas in the state.

References


