Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Temperature significantly changes COVID-19 transmission in (sub)tropical cities of Brazil

David N. Prata,⁎, Waldecy Rodrigues, Paulo H. Bermejo

HIGHLIGHTS

• The dose-response relationships suggest that the relationship between the annual average of temperature compensation and COVID-19 confirmed cases was approximately linear in the range of less than 25.8°C, which became flat above 25.8°C.

• When the average temperature was below 25.8°C, each 1°C rise was associated with a −4.8951% (t = −2.29, p = 0.0226) decrease in the number of daily cumulative confirmed cases of COVID-19.

• There is no evidence supporting that case counts of COVID-19 could decline when the weather becomes warmer, in temperatures is above 25.8°C.

• The polynomial (cubic) regression model can give insights to other researchers for testing new factors and revealing new determinants capable of fitting the trend regression to a maximum of R-squared in COVID-19 cases.

Abstract

The coronavirus disease 2019 (COVID-19) outbreak has become a severe public health issue. The novelty of the virus prompts a search for understanding of how ecological factors affect the transmission and survival of the virus. Several studies have robustly identified a relationship between temperature and the number of cases. However, there is no specific study for a tropical climate such as Brazil. This work aims to determine the relationship of temperature to COVID-19 infection for the state capital cities of Brazil.

Cumulative data with the daily number of confirmed cases was collected from February 27 to April 1, 2020, for all 27 state capital cities of Brazil affected by COVID-19. A generalized additive model (GAM) was applied to explore the linear and nonlinear relationship between annual average temperature compensation and confirmed cases. Also, a polynomial linear regression model was proposed to represent the behavior of the growth curve of COVID-19 in the capital cities of Brazil.

The GAM dose-response curve suggested a negative linear relationship between temperatures and daily cumulative confirmed cases of COVID-19 in the range from 16.8 °C to 27.4 °C. Each 1 °C rise of temperature was...
2.1. Study area

COVID-19 is a respiratory epidemic caused by the coronavirus family (2019-nCoV). The disease may cause the rapid death of those affected, depending on their general health conditions and socio-demographic profile. A relevant question about the spread of 2019-nCoV is whether climatic or demographic characteristics enable a more significant expansion of the virus. Bukhari and Jameel (2020) reported that 2019-nCoV spread rapidly to many countries, and a world pandemic crisis was subsequently declared by the World Health Organization (WHO). The authors observed that a similar variant of 2019-nCoV, the influenza virus, has been affected by climate.

Comparable to COVID-19, the SARS-associated coronavirus (SARS-CoV) may cause severe acute respiratory syndrome (SARS) (Xu et al., 2020). It has also been suggested that SARS-CoV-2 and other closely related coronaviruses, such as influenza and Ebola, have significant relations to environmental factors (Yip et al., 2007; Thai et al., 2015; Ng and Cowling, 2014; Lowen and Steel, 2014; Bi et al., 2007; Barreca and Shimshack, 2012; Moriyama and Ichinohe, 2019; Casanova et al., 2010). Some studies (Chan et al., 2011; Van Doremalen et al., 2013; Hastie and Tibshirani, 1990) have claimed that the survival time of some coronaviruses depends on temperature increases or decreases; therefore, temperature could affect the virus transmission risk. Thus far, it is reasonable to hypothesize that differences in annual average temperatures could significantly influence the transmission of the virus.

The prevalence of studies on the relationship between virus transmission and temperature have been conducted for non-tropical countries (30° N and above) in temperatures ranging from −20 °C to a maximum of 20 °C. Brazil is an expansive tropical country, with most of its territory located between the Tropic of Cancer in the north (about 23° 26′ N) to the Tropic of Capricorn in the south (about 23° 26′ S). Brazil's geographic range makes it possible to compare specific sub-regions called hot and torrid with temperate climate zones to examine the spread of 2019-nCoV in different climates and territories.

This study aims to investigate the role of Brazilian tropical weather in the transmission of coronavirus by exploring the relationship between annual average temperatures and confirmed COVID-19 cases for the state capital cities.

2.2. Data collection

The study population is the daily number of cumulative confirmed cases of COVID-19 in the 27 state capital cities, as officially reported by the Ministry of Health of Brazil from February 27 to April 1, 2020. This work focused on the capital cities because of the land cover of Brazilian territory and the few cases to date in the interior of the Brazilian states. Meteorological data were collected from the National Institute of Meteorology authority in Brazil. Meteorological data were limited and did not include all cities in Brazil. Demographic density and estimated population were collected from the Brazilian Institute of Geography and Statistics (IBGE), the official provider of geographic and statistical information of Brazil.

2.3. Statistical analysis

A descriptive analysis was performed, with numerical variables described using means, standard deviations, and distributions. A generalized additive model (GAM) was used to calculate the relationships between the temperature data and the number of cumulative total confirmed cases (lgN), respectively, to fit equations and splines. GAM fits generalized additive models (Liu et al., 2020; Wu et al., 2018) for parametric and nonparametric regression and smoothing. GAM can be useful to explore linear and nonlinear weather effects and health outcomes (Zhu and Xie, 2020). The model defined is semiparametric and additive, as follows:

\[
\log(y_i) = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + s(m_{1i}) + s(d_{di}) + s(ep_i) + e_i
\]

The model attempts to represent the polynomial behavior of the growth curve of the cumulative confirmed cases of the state capital cities of Brazil. The log-transformed daily cumulative COVID-19 counts in capital city \( i \) on day \( t \), \( \beta_0 \) is the intercept, \( s(\cdot) \) denotes a spline function with a maximum of two degrees of freedom to avoid overfitting (Liu et al., 2020; Wang et al., 2018), \( \beta \) is the parameter of \( x \), and \( x \) is the linear variable \( countdays \) in capital city \( i \) on day \( t \). The variable \( countdays \) is the counting days since the first outbreak in city \( i \). The annual average of temperatures compensation \( s(m_{1i}) \), the demographic density \( s(d_{di}) \), and the estimated population \( s(ep_i) \) were controlled for confounding effects.

Two tests for measuring the robustness of the sensitivity of the model were applied. First, the São Paulo capital city was removed from the data for two reasons: (1) it is by far the largest state capital city of Brazil, with almost the double the population of the second-most populated state capital, and (2) the total confirmed COVID-19 cases for São Paulo is more than three times the quantity of cases than the second-highest city. In the sensitivity analysis, the second test considered the logN of confirmed cases of COVID-19 per habitant for two reasons. First, the test accounted for the proportion of the population in the confirmed cases, and second, because a significantly negative correlation between temperature and population was found. A hypothesis, therefore, arises whether the significantly negative correlation between
the COVID-19 confirmed cases and temperature is due to the greater size of the population in the cities with low temperatures. The GAM model was built in SAS, with two-sided tests, and $p < 0.05$ was considered statistically significant.

3. Results

3.1. Descriptive analysis

Table 1 summarizes the descriptive statistics for the daily cumulative confirmed cases of COVID-19, meteorological, and demographic variables. This study included 586 cases during the observation period (February 27 to April 1, 2020), and the average daily number was 94.04. The mean value of the annual average of temperature compensation, 2019 demographic density, 2019 estimated population, and the counting days since the first outbreak in each city (in inhabitants (hab) per square kilometer), were 23.82 °C, 2804.47 hab/km², 2,224,190.33 hab, 11.86 days, respectively.

Table 2 shows the Spearman’s and Pearson correlation coefficients among the total confirmed cases of COVID-19 by city. Total cases had significantly negative Spearman’s and Pearson correlations with temperature ($r_s = -0.40037$, $p = 0.0385$; $r = -0.38318$, $p = 0.0485$), and positive correlation with demographic density ($r_s = 0.59010$, $p = 0.0012$; $r = 0.50991$, $p = 0.0066$), estimated population ($r_s = 0.80086$, $p < 0.0001$; $r = 0.96234$, $p < 0.0001$), and the count days since the first outbreak in each city ($r_s = 0.63382$, $p = 0.0004$; $r = 0.63689$, $p = 0.0004$).

3.2. Dose-response relationship and sensitivity analysis

The dose-response relationship in Fig. 2 is the combination of linear and nonparametric trends of GAM, suggesting a significantly

| Variable     | N   | Mean      | Std dev | Minimum | Maximum |
|--------------|-----|-----------|---------|---------|---------|
| Est_population (hab) | 586 | 2,224,190.33 | 2,866,579.43 | 299,127.00 | 12,252,023.00 |
| Countdays (N)   | 586 | 11.8634812  | 7.1400015  | 1.0000000 | 34.000000 |
| Demdensity (Km²/Hab) | 586 | 2804.47 | 2597.34 | 12.0000000 | 7786.00 |
| Casesdays (N)   | 586 | 94,0409556  | 292,2163378 | 1.0000000 | 3202.00 |
| Temperature (°C) | 586 | 23.8215017  | 2.85349866 | 16.8000000 | 27.4000000 |

Fig. 1. Map of Brazil country, with tropical and subtropical climates.

Table 2

| Variable     | Tcasesdays | Temper | Demden | Est_pop | Cdays |
|--------------|------------|--------|--------|---------|-------|
| Total casesdays (N) | 1.000 | L/NP | L/NP | L/NP | L/NP |
| Temperature (°C)  | -0.40037* | 1.000 | / | L/NP | L/NP |
| Demdensity (km²/hab) | 0.59010* | -0.26084 | 1.000 | L/NP | L/NP |
| Est_population (hab) | 0.96234* | -0.41044 | 0.62088 | 1.000 | L/NP |
| Countdays (N)     | 0.63689* | -0.66330 | 0.54290 | 0.69846 | 1.000 |

L – significant linear correlation. NP – Significant Non-Parametric Correlation.
PS.: Pearson correlation appears only if there is no Spearman’s correlation.
* $p < 0.05$. 

---

D.N. Prata et al. / Science of the Total Environment 729 (2020) 138862
negative relationship between temperature and COVID-19 confirmed cases, even accounting for the population per habitant and excluding São Paulo city from the data. Specifically, the relationship was linear in the range from 16.8 °C to 25.8 °C and then became flat above 25.8 °C, indicating a single threshold of the temperature effect on COVID-19 at 25.8 °C.

Although the decline of the curve stopped at 25.8 °C, Table 3 shows that for each 1 °C rise in temperature, there is a decrease in the cumulative daily number of COVID-19 confirmed cases. The percentage of this decrease depends on factoring in the population per habitant and removing the São Paulo city data.

The dose-response relationship was robust to the findings of the sensitivity analysis.

3.3. Further detailing the threshold

A piecewise combination was conducted of linear and nonparametric trends of GAM to quantify the effect below and above the threshold. Below the threshold of 25.8 °C, the cumulative confirmed cases decreased by $-5.9035 \ (t = -4.22, p \leq 0.0001)$ for every 1 °C rise in temperature. Because of the small quantity of cumulative confirmed cases

![Additive Component for Icases](image1)

*Fig. 2. Dose-response relationship for the effects of temperature on COVID-19 confirmed cases. The x axis is the annual average of temperature compensation. The y axis indicates the contribution of the smoother to the fitted values.*
for temperatures above the threshold, it was not possible to measure piecewise results above 25.8 °C.

3.4. The linear model results for temperature

Although the combination of linear and nonparametric trends of GAM prone to optimize the linear model, results revealed that all the variables of the model also had linear statistical significance. Linear temperature, estimate \(-0.05761\), t-value \(-5.62\), and \(p < 0.0001\). Linear density, estimate \(0.00009524\), t-value \(7.67\), and \(p < 0.0001\). Linear estimated population, estimate \(7.752899 \times 10^{-8}\), t-value \(6.06\), \(p < 0.0001\).

Thus, the linear model can be defined as follow:

\[
\log(y_{it}) = \beta_0 + \beta_1 x_{2it}^3 + \beta_2 x_{it}^2 + \beta_3 x_{it} + \beta_4 (mt_{it}) + \beta_5 (dd_{it}) + \beta_6 (dd_{it}) + \epsilon_{it}
\]

The parameters are shown in Table 4. The model is an attempt to represent the behavior of the growth curve of COVID-19 for the state capital cities of Brazil. Fig. 3 shows the growth curve of COVID-19 confirmed cases and predicted values. The individual R-squared of the factors was also calculated. The adjusted R-squared of the model was 0.81357, R-squared 0.81580, and the predicted R-squared 0.81053. The R-squared of countdays was 0.69311, the adjusted R-squared 0.69152, and the predicted R-squared 0.68924. The R-squared of temperature was 0.095873, the adjusted R-squared 0.094325, and the predicted R-squared was 0.09251. The R-squared of demographic density was 0.17693, the adjusted R-squared 0.17410, and the predicted R-squared was 0.16797. The R-squared of the estimated population was 0.20670, the adjusted R-squared was 0.20534, and the predicted R-squared was 0.19865. Table 5 summarizes the R-squared results. Findings suggest that these factors have a reasonable contribution to explaining the behavior of the COVID-19, at least for this linear model.

4. Discussion

The strong motivation for this work was the lack of a study of COVID-19 in tropical climate countries, to the authors’ knowledge. To achieve this purpose, we explored linear and nonlinear relationships between the annual average of temperatures compensation and COVID-19 confirmed cases by using a generalized additive model. Also, a generalized linear model was built to better understand the behavior of the growth curve of COVID-19 and the role of each factor to explain trends.

The linear model predicted R-square was a reasonable 0.81053, indicating that the model explains approximately 81% of COVID-19 confirmed cases. The sum of the individually predicted R-squared of the component variables of the model was 1.14111, which suggests that the component variables of the model all support the prediction of the model and are not merely the addition of variables that increase the R-squared. The temperature had a better p-value in Spearman’s correlation than the Pearson correlation. This result suggests that temperature may be explored in nonlinear models (Wu et al., 2018).

In our data, the first counting days of the outbreak had few cases reported. This may be because the outbreak was just beginning, and there was a natural lack of COVID-19 tests, and also because of health infrastructure deficiencies when the first virus contact occurred. Managing data, such as creating a mean value of the case numbers for the three

---

**Table 3**

The effects of a 1 °C increase in temperature on COVID-19 confirmed cases.

| Annual average temperature compensation from 16.8 °C to 27 °C | Percentage change (%) | t-value | p     |
|-------------------------------------------------------------|------------------------|--------|-------|
| LeCasesDays                                                  | \(-5.1781\)            | \(-4.890\) | \(-0.0001\) |
| LeCasesDayPhab                                              | \(-3.9373\)            | \(-3.44\)  | \(0.0006\)  |
| LeCasesDayPhab-SP                                           | \(-5.0224\)            | \(-4.54\)  | \(-0.0001\) |

LeCasesDays = Log of daily cumulative confirmed cases of COVID-19.
LeCasesDayPhab = Log of daily cumulative confirmed cases of COVID-19 per habitant.
LeCasesDayPhab-SP = Log of daily cumulative confirmed cases of COVID-19 per habitant removing the state capital city of São Paulo.
or four days following the first diagnosed case, appears to be a sound strategy for data analysis, as used in other works (Wang et al., 2020). The counting days also could help to abstract data, allowing for a time-less analysis by placing all data in the same time of the counting days.

Previous studies (Bukhari and Jameel, 2020) have shown significant relationships between cooler and warmer-humid regions. Thus, warmer-humidity is another interesting factor to investigate in tropical climate zones.

Many studies (Zhu and Xie, 2020; Wang et al., 2020; Yongjiiana et al., 2020; Núñez-Delgado, 2020; Liu et al., 2020; Yip et al., 2007; Thai et al., 2015; Ng and Cowling, 2014; Chu et al., 1982; Lowen and Steel, 2014; Bi et al., 2007; Barreca and Shimshack, 2012; Moriyama and Ichinohe, 2019; Casanova et al., 2010; Wang et al., 2018; Xu et al., 2020) observed temperature conditions beneficial to the coronavirus. Research also revealed that an elevated temperature was harmful to the virus (Bi et al., 2007; Casanova et al., 2010; Chan et al., 2011; Van Doremalen et al., 2013). However, in this study, we could not evince a negative effect on COVID-19 infection for higher temperatures above 25 °C. A likely reason may be the lack of quantitative data to explore, or perhaps that COVID-19 could, in fact, fit these higher temperatures. Further studies need to be conducted to discover new findings and determinants.

To our knowledge, these study findings show that COVID-19 may not vanish by itself because the weather becomes warmer. We stress that the governance of healthcare public policies cannot wait for higher temperatures to defeat COVID-19. After all, the efficient adoption of social distance policies by the Brazilian governments was an improvement in the prevention and obstruction of the viral infection. Notably, social distancing may have had a direct impact on these research results because of the change in the natural behavior of the virus.

5. Conclusion

This study intends to contribute to community research with a polynomial linear regression model (cubic) that attempts to represent the behavior of the COVID-19 growth curve. A polynomial linear model fitted well for the state capital cities of Brazil. The results were in agreement with published studies (Zhu and Xie, 2020; Wang et al., 2020; Le et al., 2020; Bukhari and Jameel, 2020; Yongjiiana et al., 2020; Núñez-Delgado, 2020; Liu et al., 2020) of COVID-19, that is, at lower temperatures, confirmed cases rise. In this study, for the tropical temperatures of Brazil, the variation in annual average temperature ranges from 16.8 °C to 27.4 °C. Results indicated that temperatures had a negative linear relationship with the number of confirmed cases. The curve became flat with a threshold of 25.8 °C. There is no evidence supporting that the curve declined for temperatures above 25.8 °C. The authors believe that by using this model, it is possible to test new factors and reveal new determinants capable of fitting the trend regression to a maximum of R-squared. The work also has the purpose of supporting governance for healthcare policymakers.

![Fig. 3. The growth curve of the 27 state capital cities of Brazil was generated by the polynomial linear regression model of this study. The x axis is the collected date of COVID-19 confirmed cases. The y axis indicates the COVID-19 actual and predicted cases.](image-url)
CRedit authorship contribution statement

David N. Prata: Conceptualization, Methodology, Software, Writing - original draft. Waldecy Rodrigues: Data curation, Writing - original draft. Paulo H. Bermejo: Visualization, Investigation, Writing - review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that might appear to influence the work reported in this paper.

Acknowledgment

The authors acknowledge the financial support of the Ministry of Health of Brazil.

References

Alvarenga, C., Hering, J., Onias, G., da Silva, L., Monteiro, S., et al., 2020. Microclimate and the transmission of severe acute respiratory syndrome (SARS) coronavirus at the Amoy Gardens outbreak. J. Clin. Virol. 121, 104222. https://doi.org/10.1016/j.jcv.2020.104222.

Barreca, A.I., Shimshack, J.P., 2012. Absolute humidity, temperature, and in...