



Article

Can Socioeconomic, Health, and Safety Data Explain the Spread of COVID-19 Outbreak on Brazilian Federative Units?

Diego Galvan ^{1,2,3,*} , Luciane Effting ⁴ , Hágata Cremasco ⁴ and Carlos Adam Conte-Junior ^{1,2,3}

¹ COVID-19 Research Group, Center for Food Analysis (NAL), Technological Development Support Laboratory (LADETEC), Cidade Universitária, Rio de Janeiro RJ 21941-598, Brazil; conte@iq.ufrj.br

² Laboratory of Advanced Analysis in Biochemistry and Molecular Biology (LAABBM), Department of Biochemistry, Federal University of Rio de Janeiro (UFRJ), Cidade Universitária, Rio de Janeiro RJ 21941-909, Brazil

³ Nanotechnology Network, Carlos Chagas Filho Research Support Foundation of the State of Rio de Janeiro (FAPERJ), Rio de Janeiro RJ 20020-000, Brazil

⁴ Chemistry Department, State University of Londrina (UEL), Londrina PR 86057-970, Brazil; luciane.effting@uel.br (L.E.); hagata@uel.br (H.C.)

* Correspondence: diegogalvann@gmail.com or diegogalvann@iq.ufrj.br

Received: 19 October 2020; Accepted: 14 November 2020; Published: 30 November 2020



Abstract: Infinite factors can influence the spread of COVID-19. Evaluating factors related to the spread of the disease is essential to point out measures that take effect. In this study, the influence of 14 variables was assessed together by Artificial Neural Networks (ANN) of the type Self-Organizing Maps (SOM), to verify the relationship between numbers of cases and deaths from COVID-19 in Brazilian states for 110 days. The SOM analysis showed that the variables that presented a more significant relationship with the numbers of cases and deaths by COVID-19 were influenza vaccine applied, Intensive Care Unit (ICU), ventilators, physicians, nurses, and the Human Development Index (HDI). In general, Brazilian states with the highest rates of influenza vaccine applied, ICU beds, ventilators, physicians, and nurses, per 100,000 inhabitants, had the lowest number of cases and deaths from COVID-19, while the states with the lowest rates were most affected by the disease. According to the SOM analysis, other variables such as Personal Protective Equipment (PPE), tests, drugs, and Federal funds, did not have as significant effect as expected.

Keywords: coronavirus disease; artificial neural networks; SARS-CoV-2; ventilator; index development index; developing country

1. Introduction

In December 2019, cases of acute pneumonia in the city of Wuhan, China, called attention due to the speed of contagion [1]. The virus that caused this pneumonia in affected individuals is the SARS-CoV-2, the causative agent of the novel coronavirus (COVID-19). The number of cases quickly evolved into a pandemic, decreed by the World Health Organization (WHO) on 11 March 2020. Currently, a few months after the first cases, this disease has been plaguing the world [2].

At the moment, the pandemic epicenter is in Latin America, with Brazil, Argentina, Colombia, Peru, Mexico, and Chile being the most representative countries, concentrating more than 10 million cases. Brazil already has more than 5 million cases and 150,000 deaths, ranking third in the world in the number of cases and deaths by COVID-19, behind only to the United States and India [3].

Intense efforts by the scientific community are focused on finding drugs to treat [4] and develop a vaccine for the disease [5]. Other research tries to understand how the disease has spread around the world. Some factors can influence the spread of the virus, including the disregard of preventive measures adopted by the government by population [6], climatic [7–9], comorbidities [10,11], hospital structure, prepared professionals, personal protective equipment, and financial resources [4].

In order to monitor the pandemic, it is necessary to collect information on the most significant possible number of factors that can influence the disease's dissemination behavior and, with this, establish a relationship that allows to affirm which measures delay the spread of COVID-19. Statistical tools have been used to study the behavior of epidemics for years; this area treats, analyzes, and obtains information from large datasets that cannot be analyzed by traditional systems [12].

Artificial Neural Networks (ANNs) are tools that have been gaining scientific relevance to perform pattern recognition, classification, or prediction tasks. ANN is a non-linear computational model attempting to simulate human brain structure and decision-making [13]. Its architecture is inspired by biological neural networks and consists of simple processing units that store empirical knowledge through a learning process [14].

Among the most common types of neural networks are the Feed-Forward Neural Network (FFNN), the Convolutional Neural Network (CNN), and the Recursive Neural Network (RNN). FFNNs are the most common type of ANN in practical applications, consisting of one or more hidden layers of perceptrons (neurons) that require supervised training. The input data of the desired sample datasets and the output results are sent to the network several times until the error in the output is minimized [13].

Traditionally, several types of neural networks exist with various techniques such as Autoregression (AR) [15,16], Moving Average (MA) [17,18], Exponential Smoothing (ES) [19], Hybrid Methods (HM) [20–22], and Autoregressive Integrated Moving Average (ARIMA) [23]. They have been used to predict the dependent variable in a time series [15–23]. Among these techniques, unsupervised neural networks of the type Self-Organizing Maps model has mostly outperformed others in precision and accuracy [14].

Self-Organizing Maps (SOM) or Kohonen Map is a method for the analysis of multivariate data used for pattern recognition and classification, which may be taken as a non-linear generalization of principal component analysis [24]. The SOM algorithm consists of input nodes and a grid of computational connected nodes (neurons), which compete among themselves for activation as the one that most closely resembles the input vector. If the input data exhibit some similarity across the input classes, the neurons will organize themselves, showing similarity patterns in a grid [13,25].

SOM has already been used successfully for data mining in several areas of knowledge [24,26–30], like food science [26,28,29], fuels [27], monitoring chemical reactions [24,30], and other applications [25,31,32]. Recently, SOM was used to verify the spatial relationship of the COVID-19 spread in countries and states in Mexico [31]. Most published papers have the main theme of the modelling or prediction of the spread of COVID-19 [2,9,33–36]. In our literature searches, we did not evidence reports that simultaneously consider the historical data on the number of cases/deaths from COVID-19 with possible external factors that affect the spread of the virus through a spatial analysis by pattern recognition.

In a previous study, we demonstrated that the spread of COVID-19 varies according to Brazilian regions, states, and cities. However, it was not possible to allege the reason for this behavior. In this study, 14 possible factors that can affect the spread of COVID-19 were analyzed jointly by SOM. This analysis will verify which variables may be essential and will point out possible variables that had the most significant effect on each Brazilian state.

2. Theoretical Foundations

The Kohonen algorithm was proposed by Teuvo Kohonen in 1982 [14]. In SOM, there is no given output target, the objective of the algorithm is to find a set of neurons to represent the cluster, but with

topological restrictions [31]. The learning process begins with the random initialization of the synaptic weights vector of each neuron. In sequence, three key steps are developed for the formation of the feature map: Competition, cooperation, and synaptic adaptation [24].

The following is a brief description of the Kohonen algorithm, as described previously by Haykin [14]. The function is chosen to represent the topological neighborhood in the following equation.

$$h_{j,i} = \exp(-d_{j,i}^2/2\sigma^2)$$

where σ is the effective radius of the topological neighborhood, and $d_{j,i}$ is the lateral distance between the winning neuron i and the excited neuron j , defined through Euclidean distance. Over the training epochs, there is a reduction in the size of the neighborhood due to an exponential decay, described by the equation.

$$\sigma(n) = \sigma_0 \exp(-n/\tau_1) \quad n = 0, 1, 2, \dots$$

where σ_0 is the effective radius in the initialization of the algorithm, τ_1 is the time constant, with $\tau_1 = 1000/\log \sigma_0$ being recommended, and n is the number of training epochs.

During the adaptive process, the synaptic weight vector (\mathbf{w}_j) of the j neuron in the grid must be modified concerning the input vector \mathbf{x} . The modification process is a modification of the Hebb learning postulate, described by the equation.

$$\mathbf{w}_j(n+1) = \mathbf{w}_j(n) + (\eta)h_{j,i(x)}(\mathbf{x} - \mathbf{w}_j(n))$$

where $\eta(n)$ is the learning rate, which is variable and decreases during the training epochs, n . The learning rate decrease may be modelled by an exponential decay, as described in the equation. In this equation, η_0 is the initial learning rate, and τ_2 is another time constant; the recommended values are, respectively, 0.1 and 1000:

$$\eta(n) = \eta_0 \exp(-n/\tau_2)$$

3. Methodology

3.1. Dataset

The dataset used in the study by neural networks was obtained from websites of Brazilian government agencies or institutions, available for download. Before the SOM analysis, the variables that would constitute the study were selected. All variables related to socioeconomic, health, and safety factors that could explain the dissemination of COVID-19 in Brazil were included in our research. They were made available and updated daily by institutional sites of the Brazilian federal government.

Initially, data for all variables were converted and expressed as rates for each 100,000 inhabitants. For the conversion, we used data from the Brazilian population estimated in 2019 by the Instituto Brasileiro de Geografia e Estatística – IBGE (Brazilian Institute of Geography and Statistics); the public agency responsible for conducting censuses and organizing information related to the country's geosciences and social, demographic, and economic statistics [37]; available in <https://www.ibge.gov.br/>.

The numbers of cases and accumulated deaths by COVID-19 were analyzed for 16 epidemiological weeks. According to the international convention, each week starts on Sunday and ends on Saturday. The information represents the numbers recorded from 26 February to 13 June, 2020 (110 days), a period that corresponds from the 9th to the 24th epidemiological week of the year, first COVID-19 case record in Brazil. The data were obtained from the website of the Ministério da Saúde—MS (Ministry of Health) of Brazil, the government sector responsible for the administration and maintenance of public health in the country, updated daily with information about the COVID-19 pandemic in the country [38]; available in <https://covid.saude.gov.br/>.

The Federal Government of Brazil transfers financial resources, supplies, and equipment for structuring health services in the country to combat or treat patients affected by COVID-19. The MS frequently updates a panel of inputs considered essential, distributed to each Brazilian

state: Influenza vaccination—H1N1 and H3N2 (distributed and applied), drugs (chloroquine and oseltamivir), COVID-19 tests (rapid and reverse-transcriptase polymerase chain reaction—RT-PCR), hand sanitizer per liter, and Personal Protective Equipment—PPE (surgical masks, N95 mask, gowns, gloves, glasses, face shield, caps, and sneakers) [39]. These data were obtained on June 13, 2020, at <https://covid-insumos.saude.gov.br/paineis/insumos/painel.php>.

The destination of Federal funds transferred to each federative unit (Brazilian states) is monitored and presented on the Federal Government website [40], available in <https://www.tesourotransparente.gov.br/visualizacao/painel-de-monitoramentos-dos-gastos-com-covid-19>. The data used in this study represent the sum of all transfers made until 13 June 2020. For better visualization, the values were converted from the real (R\$) to the US dollar (US\$) by the quotation of R\$ 5.05 equals to US\$ 1.00, referring to 12 June 2020, provided by the Banco Central do Brasil—BCB (Central Bank of Brazil) [41], available in <https://www.bcb.gov.br/>.

Health data from public and private institutions, such as the number of physicians, nurses, ventilators, and Intensive Care Unit (ICU) beds for COVID-19 in each state, were obtained from the IBGE [42]; available in <https://mapasinterativos.ibge.gov.br/covid/saude/>.

Finally, the data referring to the Human Development Index (HDI) of each state for the year 2015 (last census registered in the country) were obtained from the Atlas do Desenvolvimento Humano do Brasil—ADHB (Human Development Atlas of Brazil), a public institution that provides information about human development in Brazil [43]; available in <http://www.atlasbrasil.org.br/2013/pt/download/base/>.

3.2. Data Analysis by Artificial Neural Network

The Kohonen Map routine developed was carried out in Matlab software (MathWorks, Natick, MA, USA) according to the algorithm previously described in Haykin [14] and Cremasco et al. [26]. The datasets were evaluated using three different approaches. In the first step, two SOM analyses were conducted to verify the distribution of the numbers of cases and deaths from COVID-19 in Brazilian regions and states. In the second step, the distribution of 14 variables for the Brazilian states was verified: ICU beds, ventilators, physicians, nurses, PPE, hand sanitizer, rapid test, PCR test, vaccines distributed, vaccines applied, chloroquine tablets, oseltamivir capsules, HDI, and federal funds distribution.

The SOM setup was a hexagonal topology of 8×8 and 4×4 for 27 Brazilian states and five regions, respectively, for the number of cases and deaths from COVID-19 (**first step**); 8×8 for 27 Brazilian states with 14 variables (**second step**) with 7000 training epochs to ensure convergence of the average quantization error. The initial neighborhood relationship was 3.5, decreasing to 0.07, with an initial learning rate of 0.1, decaying exponentially with the training epochs to 9.11×10^{-5} . It was used a computer Intel® Core™ i7-4790 CPU© 3.60 GHz, 32 GB RAM, and 250 GB HDD.

In order to obtain a better representation, the values collected after the SOM analysis were transposed through the color scale to the Brazilian cartographic map for each variable. This procedure was adopted to facilitate the interpretation of unfamiliar people with the Kohonen map. The original SOM output weight maps for the 14 variables generated are displayed in the Supplementary Material in Figures S1–S14.

3.3. Spearman's Correlation Test

The Spearman's rank correlation coefficient is adopted to determine the correlation between variables. It analyzes how well the association between two variables can be defined using a monotonic function. The correlation matrix was performed using the package “corrplot” in the Software R Core Team (Vienna, Austria) [44].

4. Tests and Results

Figure 1 shows the number of cases and deaths accumulated by Brazilian states and regions per 100,000 inhabitants. All states belonging to the South (S) and Central-West (CW) regions of the country

had the lowest rates of cases and deaths by COVID-19 recorded, an average of 180 cases and 4 deaths per 100,000 inhabitants. Most of the Brazilian states with the highest rates of cases and deaths belong to the North (N) region, an average of 954 cases and 43 deaths per 100,000 inhabitants, mainly represented by Acre (AC), Amapá (AP), Amazonas (AM), Pará (PA), and Roraima (RR).

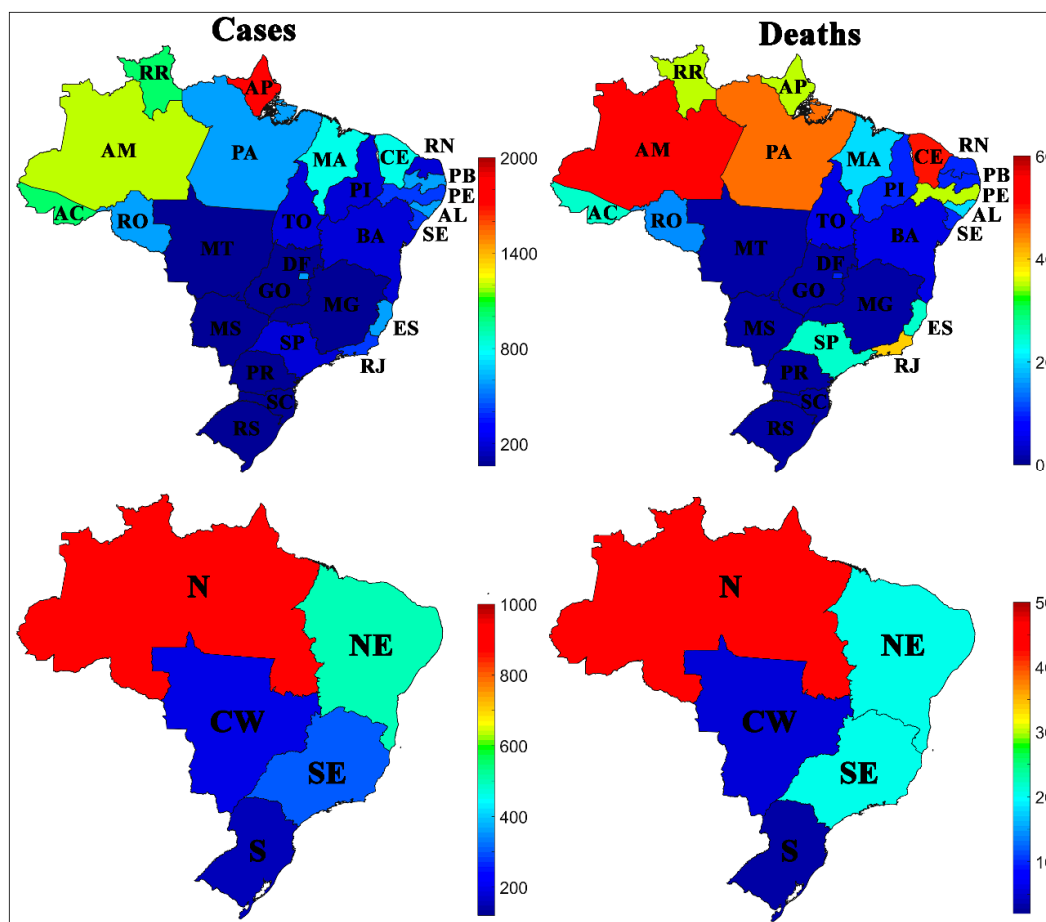


Figure 1. Geographical distribution of cases and deaths by COVID-19 in Brazilian states (above) and regions (below) per 100,000 inhabitants.

Some states in the Brazilian Northeast (NE) and Southeast (SE) also had high rates of cases and deaths when compared to the other states in the South (S) and Central-West (CW) regions. The Northeast (NE) is the second region with the highest rate of cases and deaths, an average of 527 cases and 24 deaths per 100,000 inhabitants, the states in this region that were most aggravated by the disease were Ceará (CE), Maranhão (MA), and Pernambuco (PE). The Southeast (SE) is the third region with the highest rate of cases and deaths, with an average of 338 cases and 23 deaths by COVID-19, with Espírito Santo (ES), Rio de Janeiro (RJ), and São Paulo (SP) being the most representative states in the region.

The neural network demonstrated that the spread of COVID-19 in Brazil has heterogeneous behavior; see Figure 1. It is essential to understand this behavior. The North of the country was more affected by COVID-19 than the South. The Northeast and Southeast differ in case rates but have similar death rates. In this sense, possible factors were evaluated that may explain the reason for this behavior and point out which measures are more relevant in the fight against COVID-19 in each Brazilian state.

After the training phase of the SOM network, we generated the topological map of the 14 variables evaluated together, which represents the distribution of each federative unit in Brazil, according to the winning neuron; see Figure 2. In the topological map, each federative unit is associated with a respective winning neuron, that is, the one that best represents in the analysis. The SOM network classifies the input data as clusters that can be formed by one or more neurons. The definition of

clusters is characterized by the presence of empty neurons between the groups. Nearby clusters share some similarity, that is, the greater the Euclidean distance, the greater the difference in behavior.

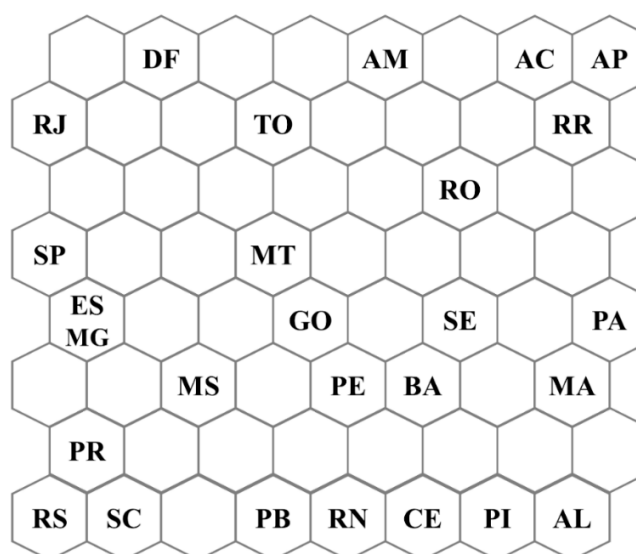


Figure 2. Distribution of Brazilian federative units according to the winning neuron. Where: Acre—AC; Alagoas—AL; Amapá—AP; Amazonas—AM; Bahia—BA; Ceará—CE; Distrito Federal—DF; Espírito Santo—ES; Goiás—GO; Maranhão—MA; Mato Grosso—MT; Mato Grosso do Sul—MS; Minas Gerais—MG; Pará—PA; Paraíba—PB; Paraná—PR; Pernambuco—PE; Piauí—PI; Roraima—RR; Rondônia—RO; Rio de Janeiro—RJ; Rio Grande do Norte—RN; Rio Grande do Sul—RS; Santa Catarina—SC; São Paulo—SP; Sergipe—SE; Tocantins—TO.

We identified the formation of some clusters by evaluating the topological map. Among some of the clusters formed, we highlight the one formed in the lower-left corner of the map in Figure 2, represented by the states of the Southern region of the country, which contain the federative units: Paraná (PR), Rio Grande do Sul (RS), and Santa Catarina (SC). This cluster contains the federal units and the region of the country with the lowest rates of deaths and cases by COVID-19, as shown in Figure 1. Another cluster formed that we can highlight is composed of some states in the North region of the country, represented by the states of Acre (AC), Amapá (AP), and Roraima (RR), presented in the upper-right corner of the map. This cluster represents the region and some of the federative units with the highest rates of deaths and cases due to COVID-19 in the country. This same analysis can be done for the other formed clusters.

To verify the applicability of SOM analysis, we compared the results obtained with another unsupervised method. Thus, we evaluated the ability to group the data obtained using the Hierarchical Cluster Analysis (HCA) method. The results obtained with the HCA were very similar to those obtained with the SOM algorithm. The dataset presents different patterns according to the region in which the Brazilian states are located. In other words, most states in the North and Northeast regions formed a cluster, while most federative units in the Central-West, South, and Southeast formed another cluster. The graphic output generated by the HCA method is shown in Figure S15 of the Supplementary Material.

In general, the topological map of the SOM network made it possible to state that the socioeconomic, health and safety data demonstrate that the spread of COVID-19 in the country varies according to the Brazilian federative units. However, only the topological map does not allow to state which of the 14 variables evaluated may be the main responsible for explaining the spread of COVID-19 in the country. Thus, we used weight maps in the following discussions. In order to obtain a better representation, the values collected from the weight maps were transposed through the color scale to the Brazilian cartographic map for each variable. However, when doing the transposition procedure,

we lose the neighborhood relationship. In Figures S1–S14 of the Supplementary Material, the original outputs of the weight maps for each variable in the SOM network are shown.

The 14 variables were evaluated jointly by ANN. However, for better results visualization and discussion, the weight maps were divided into four blocks and grouped according to similarities: (i) Available hospital infrastructure; (ii) inputs and tests available; (iii) drugs available; and (iv) financial resources,

The first set of variables is shown in Figure 3, composed of the rates of ICU beds, ventilators, physicians, and nurses per 100,000 inhabitants. In general, it is observed that the states that belong to the South, Southeast, and Central-West regions of Brazil had higher rates than the North and Northeast. There is evidence that these variables influence the rates of cases and deaths in each Brazilian state, and it can be said that the states with the lowest rates of ICU, respirators, physicians, and nurses, have the highest rates of cases and deaths from COVID-19.

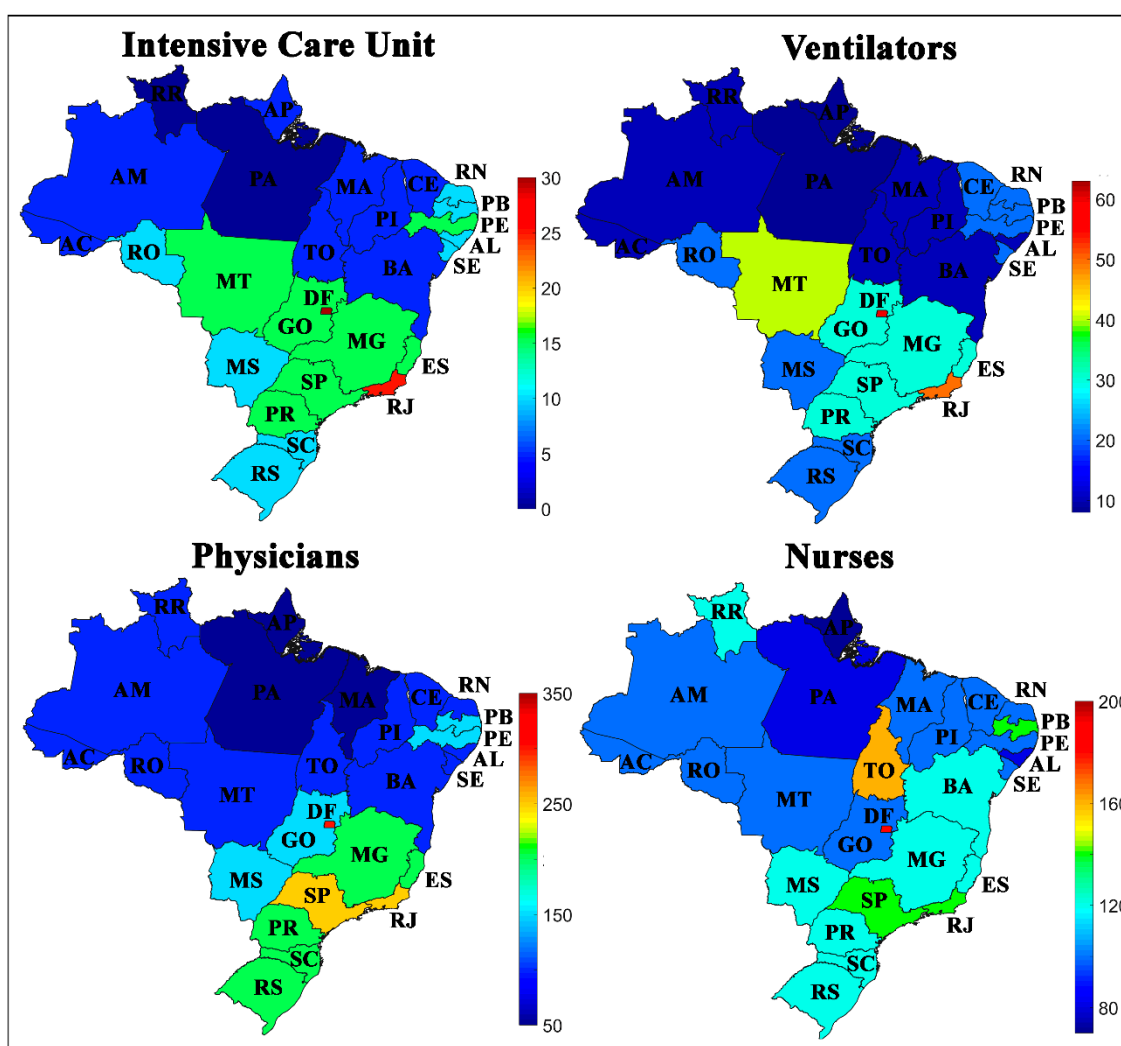


Figure 3. Geographical distribution of Intensive Care Unit (ICU) bed, ventilators, physicians and nurses in Brazilian states per 100,000 inhabitants.

It is also evident that these were not the only relevant factors. The state of Rio de Janeiro (RJ), for example, has high rates of ICU beds, respirators, physicians, and nurses and despite this, it has a higher death rate than other states in the North region, such as Acre (AC), Amapá (AP), and Roraima (RR).

The second set of variables presented in Figure 4 represents the inputs destined to each state by the Federal Government to combat and control the spread of COVID-19, composed by PPE, hand sanitizer,

rapid test, and PCR test per 100,000 inhabitants. In general, the PPE and hand sanitizer rates distributed for each state did not have a direct relationship with the rates of cases or deaths by COVID-19 according to the SOM analysis.

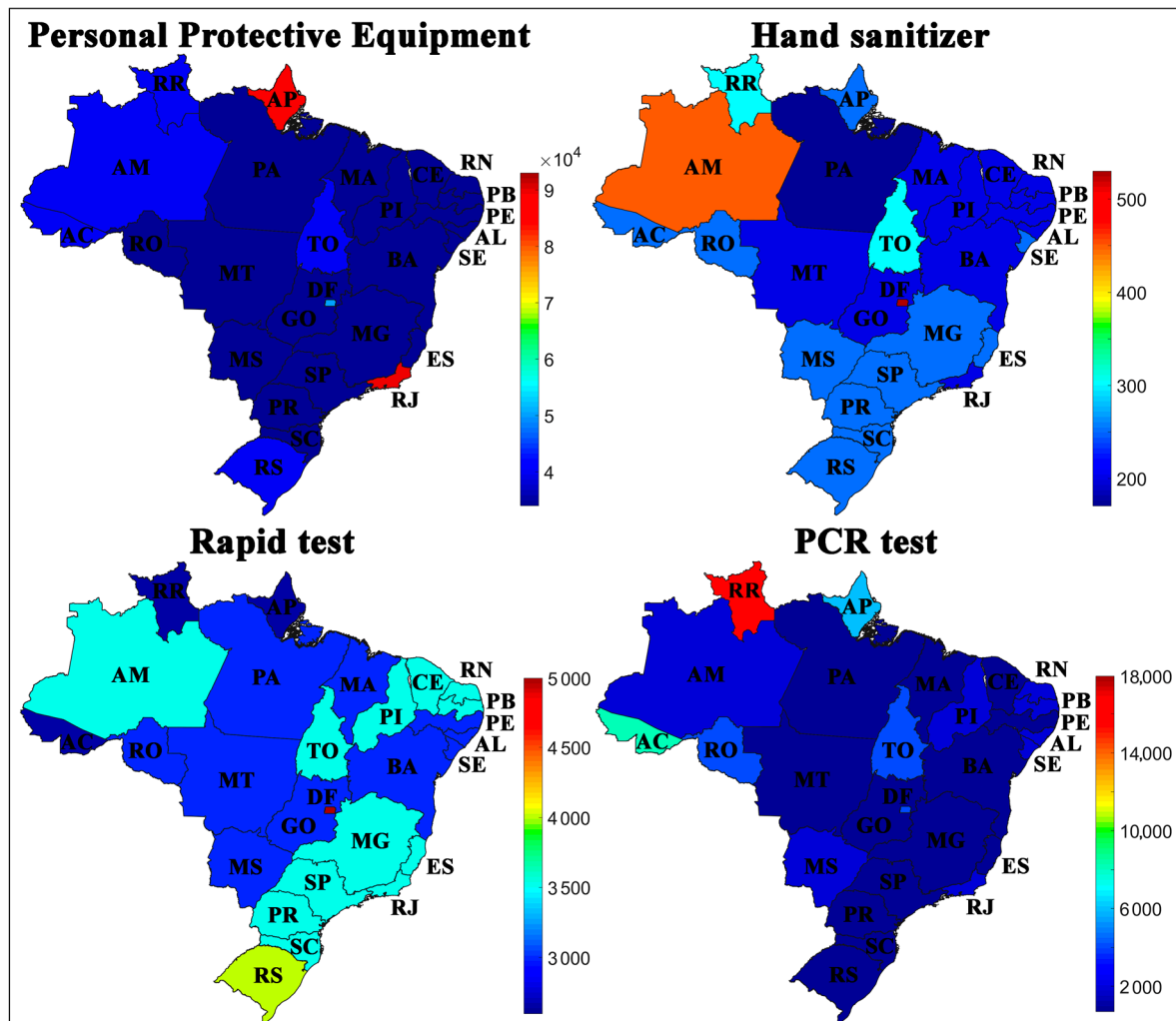


Figure 4. Geographical distribution of Personal Protective Equipment (PPE), hand sanitizer, rapid tests, and PCR tests available to Brazilian states per 100,000 inhabitants.

The states of Amapá (AP) and Rio de Janeiro (RJ), for example, had the highest rates of PPE distributed, and Amazonas (AM) had the highest rates of liters of hand sanitizer distributed per 100,000 inhabitants; however, these three states present higher rates of cases and deaths when compared to other Brazilian states.

It is important to note that the data refer to the distributed quantity of inputs, and it is not possible to estimate which portion was used by hospitals or the population, or even say if they were used correctly. The cultural factor also interferes in these variables, mainly concerning the education and awareness of the local population about the use of these inputs [45].

The disease tests in the country have been carried out homogeneously, as shown in Figure 4. This homogeneity is considered an important aspect and demonstrates that differences do not influence the data collected regarding the number of cases in the form of confirmation of the disease, which could question the reliability of the data. The states that had the highest test rates available were the Distrito Federal (DF), Rio Grande do Sul (RS), and Roraima (RR), which belong to the Central-West, South, and North regions, respectively, i.e., the most and least affected regions by COVID-19 in the country.

Rapid tests are cheaper, less reliable, and can result in false positives or negatives, while the PCR test is reliable but expensive and time-consuming [46]. It is necessary to adopt a balance between rapid tests and PCR; for example, if PCR test indices are higher, it will not be possible to follow the spread in a short period, while rapid tests, when misused, can generate mistaken information.

Previous measures as vaccines and drugs against diseases with similar symptoms, such as influenza H1N1 and H3N2, can facilitate the diagnosis of COVID-19, since they are relevant information during the patient’s anamnesis. These procedures allow faster and more efficient identification of the disease, which makes the adopted treatments more assertive and precocious [47]. Another measure that focuses on significant discussions in the scientific community is the use of drugs against the coronavirus [48]. Among several drugs, the Brazilian Federal Government has invested and passed on chloroquine tablets to the states.

The third set of variables in Figure 5 consists of influenza vaccines distributed and applied, and drugs distributed by states per 100,000 inhabitants. The states of the North region, such as Amazonas (AM), Roraima (RR), Pará (PA), and Acre (AC), had a more significant disparity than the other states in the rates of vaccines applied and distributed. At first, this behavior could be related to health factors, shown in Figure 3. However, this disparity would also be observed in the Northeast region. Other intrinsic factors may include geographic and logistical factors that make distribution and access to the population a challenge to receive the vaccine.

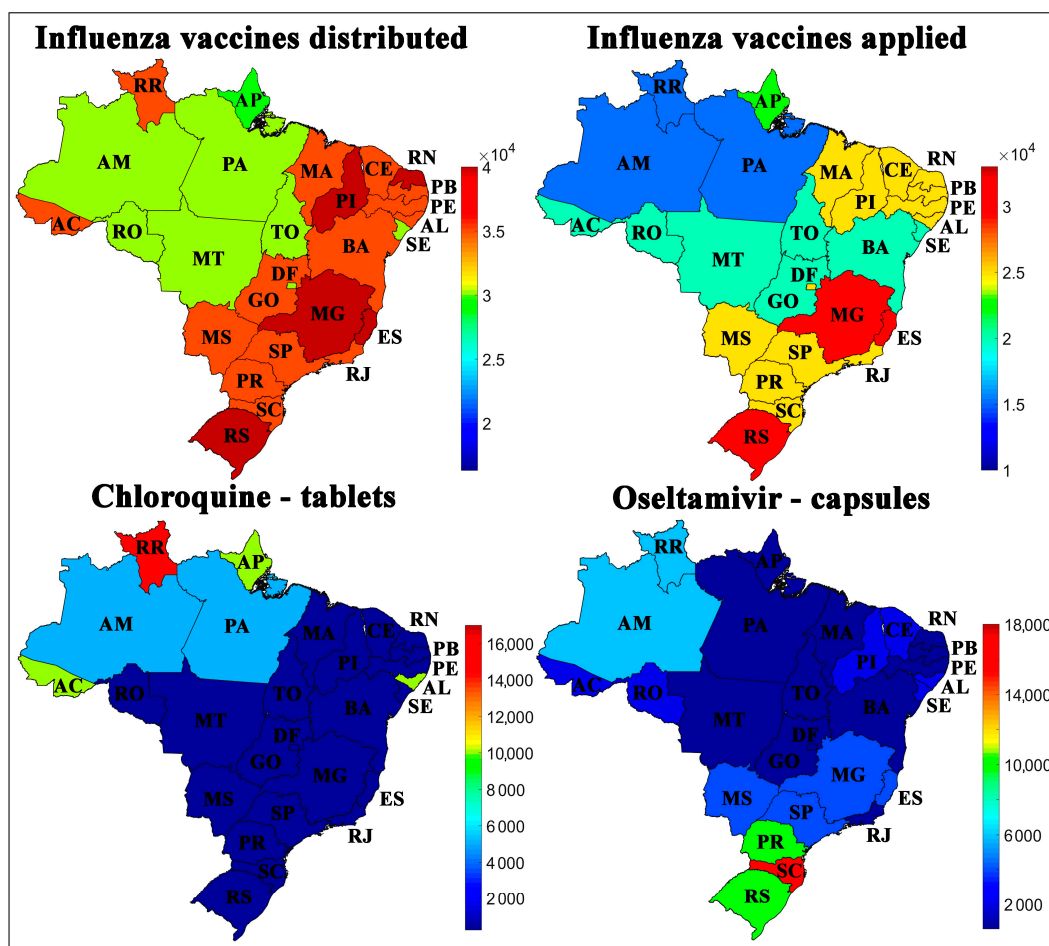


Figure 5. Geographical distribution of influenza vaccines distributed, influenza vaccines applied, chloroquine tablets, and oseltamivir capsules available in Brazilian states per 100,000 inhabitants.

As far as we know, there is no drug with proven efficiency against SARS-CoV-2; some drugs have been evaluated to combat or alleviate the disease’s symptoms [48]. Recent studies demonstrate that

dexamethasone welcomes the preliminary treatment of critically ill patients with COVID-19 [49]. Recently, the Food and Drug Administration (FDA) revoked emergency use authorization for chloroquine and hydroxychloroquine in patients with COVID-19 [50].

Figure 5 shows the chloroquine rates distributed by the Federal Government per 100,000 inhabitants. Chloroquine tablets were more widely distributed in the states of the North region, which has higher rates of deaths and cases of COVID-19 in Brazil. The possibility of greater distribution of chloroquine in these states may have occurred as an attempt to treat the disease symptoms when there were no restrictions on its use.

The oseltamivir rates distributed by the Federal Government per 100,000 inhabitants is shown in Figure 5. This drug is commonly used to treat symptoms caused by the influenza virus and is administered when there is still no confirmed diagnosis of COVID-19 [51]. Oseltamivir was more widely distributed in the Southern and Northern states of Brazil; however, it is not possible to establish a direct relationship with the rates of cases and deaths by COVID-19 using SOM analysis.

The last set of variables evaluated comprises the HDI and the distribution of federal funds destined to each state. According to Figure 6, the destination of Federal funds was more significant for some states in the North, the region which has the highest rates of cases and deaths due to COVID-19. More resources destined for this region can be guided by the scenario of cases and deaths registered.

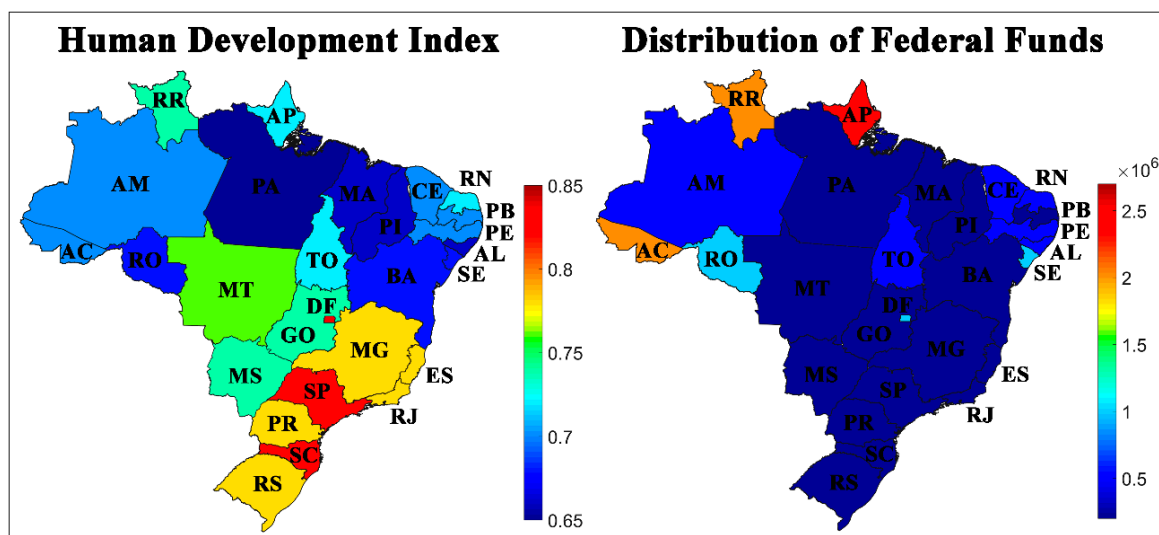


Figure 6. Geographical distribution by Index Development Index (HDI) and Federal funds destined to combat COVID-19 in Brazilian states per 100,000 inhabitants.

The highest HDI values are in the Central-West, South, and Southeast regions, with the Federal District (DF), Santa Catarina (SC), and São Paulo (SP) being the highest values in the country. It is also worth mentioning that the most affected states by COVID-19 have some of the lowest HDIs in the country, showing problems that some of these states have in fundamental areas of the population service.

Figure 7 shows the Spearman's correlation test adopted to determine the correlation among variables. The test demonstrated that the influenza vaccine applied, ICU, ventilators, physicians, nurses, and HDI were the most correlated significant variables, all positive correlations, while chloroquine had a negative correlation with all these variables. The HDI is directly related to essential aspects of the population, such as family income, education level, health, among other factors; therefore, the correlation with the HDI was already expected, since it takes into account important aspects of health for its determination. Other variables evaluated showed minor correlations, while PPE did not correlate with any other variable.

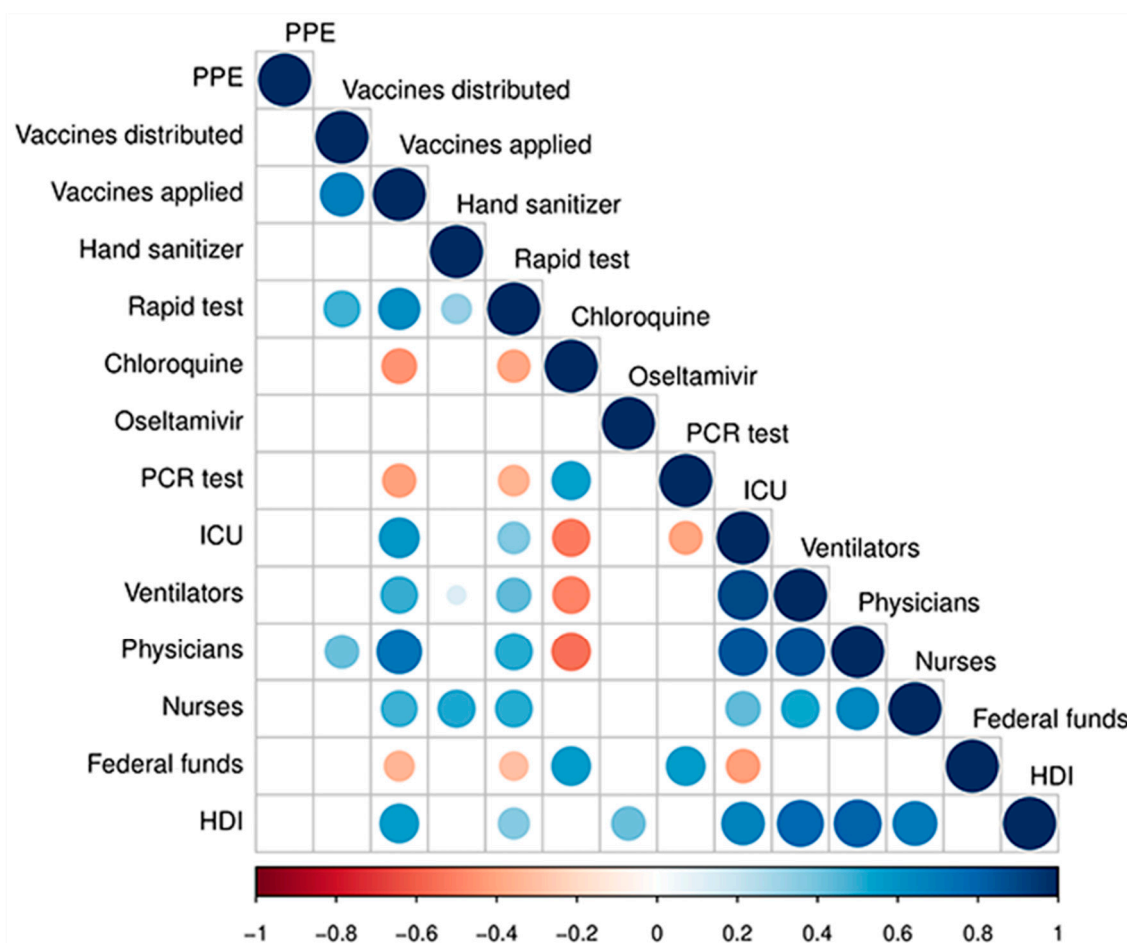


Figure 7. Spearman correlation coefficients for the 14 evaluated variables. Correlations with p -value > 0.05 are considered insignificant.

5. Discussion

Considering the importance of understanding the spread of the virus, many studies have been presented in recent months related to different aspects of the pandemic, with the application of computational intelligence tools to model and predict the spread of the disease [34].

In our study, we gave the data a differentiated approach from the reported works. It is not common to find simultaneous studies with historical data on the number of COVID-19 cases, related to external factors, that affect the spread of the virus through spatial analysis [31]. For this, we used unsupervised pattern recognition to analyze some variables, among many others, that may be related to the disease spread.

Brazil is one of the most affected countries by the pandemic and by the tremendous socioeconomic, territorial, climatic differences, among the country's federative units. We understand that it may be interesting to try to establish spread patterns of the virus in the country. Thus, once the relationship of these factors with the spread and lethality of the disease has been identified, we hope that our study can assist in the analysis of the data that have been generated by institutions and, consequently, assist in directing decision-making on practical combat actions.

It is important to note that our analysis does not aim to point out or compare which were the best actions adopted to contain the spread of COVID-19, so we approached each variable according to the significance indicated by the SOM analysis.

Based on SOM's clustering skills, we were able to spatially group similar federative units in terms of the number of cases and deaths by COVID-19 with data on the distribution of financial resources, equipment, health professionals and HDI, represented by a color scale. Thus, SOM's ability to cluster,

made it possible to cluster together with the federal units that are behaving similarly and, therefore, can benefit from similar strategies to deal with the virus spread.

The SOM analysis allowed us to raise some hypotheses about the spread of the virus in Brazil. In general, the analysis indicated that the spread of the disease has a direct relationship with quite heterogeneous socioeconomic, health, and safety variations in the national territory; see topological map in Figure 2.

In Figure 2, we can see that the spread of the disease in Brazilian states has been differentiated and regionalized. Interestingly, the network separated the most and least regions affected by COVID-19. We found that there was a separation pattern, in which all the Southern federative units are located at the bottom left of the map (PR, RS, and SC), while the Southeast states (SP, RJ, MG, and ES) and the Central-West (MS, MT, DF, and GO) are in the upper left corner of the map. In the North and Northeast regions, the federative units were classified on the right side of the map.

For better visualization, we invite readers to check the original output weight maps of the variables in the Supplementary Material. Weight maps represent the overlap of the topological map (Figure 2) and allow us to evaluate the behavior of each variable for the segmentation of the federative units. Analyzing the weight maps for the 14 variables extracted from the SOM network, it was possible to show which were the main responsible for differentiating the states, regarding the number of cases and deaths by COVID-19 in Brazil by socioeconomic, health, and safety data.

According to the 14 weight maps, the HDI was one of the essential variables for the distribution obtained (Figure 6 and Figure S13). The analysis indicated that the federal units with the lowest HDI were considerably more affected by the pandemic and that they had greater difficulty in combating the spread of the virus. Meanwhile, the best-prepared Brazilian federations with higher HDI values were more capable of deal with the pandemic.

Possibly, the HDI was the main measure responsible for the separation due to its high correlation with the other variables. It is noteworthy that this index relates to health (life expectancy), education (adult literacy index and levels of education), and income (GDP—gross domestic product—per capita). Thus, it is possible to observe the correlation of the HDI in Figure 7, with other variables, such as ICU (Figure S1), ventilators (Figure S2), doctors (Figure S3), and nurses (Figure S4).

Another important variable that allowed us to assess the behavior of the spread of COVID-19 in Brazil was the influenza vaccine rates of doses distributed and applied, highly correlated according to the Spearman correlation test (Figure 7). According to Figure 6, Figures S9 and S10, we show that the federative units with the highest rates of vaccines applied had lower rates of cases and deaths from COVID-19, as pointed out by the SOM analysis. Although this procedure is not effective against SARS-CoV-2, this measure may have facilitated clinical diagnosis and made the treatment of affected patients by COVID-19 faster and more accurate, and may have reflected in the number of cases and deaths in these regions and federative units.

Other measures adopted in the country, such as the use of drugs as chloroquine and oseltamivir, did not allow to evaluate the spread of COVID-19, as shown in Figure 5 and the weight maps of the variables in Figures S11 and S12. The SOM analysis indicated that the federative units in the North and Northeast regions had higher rates of chloroquine tablets distribution but more cases and deaths from the disease were registered. On the other hand, oseltamivir capsules had higher distribution rates in the Southern and Northern states, representing the regions more and less affected by the disease, which means that this variable, the spread of COVID-19 was similar.

We evidence that PPE (Figure S5) and hand sanitizer (Figure S6) rates distributed for each state did not have a direct relationship with the rates of cases or deaths by COVID-19 when considering the SOM analysis (Figure 4). It is known that these measures directly influence the control of the COVID-19 dissemination [45]; in this sense, we show that the SOM analysis allowed us to verify that although these items have been distributed to federative units, the population has not followed the measures adopted or made use of these items properly. Figure 7 reinforces this hypothesis since the

PPE did not correlate with any other variable. Thereby, government agencies must further encourage the use of available items.

Among the other less expressive variables evaluated, we show that the rapid test (Figure S8), PCR test (Figure S9), and distribution of Federal funds (Figure S14) did not present an obvious behavior that would allow explaining the spread of COVID-19 using the numbers of cases and deaths in the country, according to the SOM analysis indicated.

6. Conclusions

Unsupervised Artificial Neural Networks of the Self-Organizing Map type have demonstrated that the spread of the coronavirus has heterogeneous behavior and varies among Brazilian regions and states. According to the analysis, among the 14 variables evaluated, the factors responsible for the highest relationship with the numbers of cases and deaths by COVID-19 in Brazilian states were: Rates of influenza vaccine applied, ICU beds, ventilators, physicians, nurses, and HDI, positively correlated according to the Spearman's correlation test. In general, the lowest rates of cases and deaths by COVID-19 were recorded in the Brazilian states with the highest rates of influenza vaccine applied, ICU beds, respirators, physicians, and nurses, per 100,000 inhabitants, which consequently has some of the highest HDI in the country.

As a future work, we intend to integrate the SOM algorithm to analyze the spatial and temporal aspects of the COVID-19 spread in a unified way to obtain a complete view and solution to the problem. In addition, to analyze the possibility of applying SOM in other diseases that have affected Brazil, caused by dengue virus, Chikungunya virus, Zika virus, and yellow fever virus.

Supplementary Materials: The following are available online at <http://www.mdpi.com/1660-4601/17/23/8921/s1>, Figure S1: Weight maps overlaid by topological maps for variable Intensive Care Unit (ICU) by Brazilian federative unit. Figure S2: Weight maps overlaid by topological maps for variable Ventilators by Brazilian federative unit. Figure S3: Weight maps overlaid by topological maps for variable Physicians by Brazilian federative unit. Figure S4: Weight maps overlaid by topological maps for variable Nurses by Brazilian federative unit. Figure S5: Weight maps overlaid by topological maps for variable Personal Protective Equipment by Brazilian federative unit. Figure S6: Weight maps overlaid by topological maps for variable Hand sanitizer by Brazilian federative unit. Figure S7: Weight maps overlaid by topological maps for variable Rapid test by Brazilian federative unit. Figure S8: Weight maps overlaid by topological maps for variable PCR test by Brazilian federative unit. Figure S9: Weight maps overlaid by topological maps for variable Influenza vaccines distributed by Brazilian federative unit. Figure S10: Weight maps overlaid by topological maps for variable Influenza vaccines applied by Brazilian federative unit. Figure S11: Weight maps overlaid by topological maps for variable Chloroquine tablets by Brazilian federative unit. Figure S12: Weight maps overlaid by topological maps for variable Oseltamivir capsules by Brazilian federative unit. Figure S13: Weight maps overlaid by topological maps for variable Human Development Index (HDI) by Brazilian federative unit. Figure S14: Weight maps overlaid by topological maps for variable Distribution of Federal funds by Brazilian federative unit. Figure S15: Dendrogram of an HCA for Brazilian Federative Units.

Author Contributions: D.G.: Conceptualization, formal analysis, investigation, methodology, software, validation, visualization, writing—review and editing. L.E.: Conceptualization, formal analysis, investigation, methodology, software, validation, visualization, writing—review and editing. H.C.: Conceptualization, formal analysis, investigation, methodology, software, validation, visualization, writing—review and editing. C.A.C.-J.: Methodology, writing—review and editing, supervision, project administration, funding acquisition. All authors have read and agreed to the published version of the manuscript.

Funding: The authors are thankful for the financial support provided by the Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ), D.G.: [grant number E-26/200.062/2020, E-26/010.000.984/2019, and E-26/010.000148/2020]. <http://www.faperj.br/>. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflicts of Interest: The authors have declared that no competing interests exist.

References

1. Wells, C.R.; Sah, P.; Moghadas, S.M.; Pandey, A.; Shoukat, A.; Wang, Y.; Wang, Z.; Meyers, L.A.; Singer, B.H.; Galvani, A.P. Impact of international travel and border control measures on the global spread of the novel 2019 coronavirus outbreak. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 7504–7509. [[CrossRef](#)] [[PubMed](#)]

2. Gatto, M.; Bertuzzo, E.; Mari, L.; Miccoli, S.; Carraro, L.; Casagrandi, R.; Rinaldo, A. Spread and dynamics of the COVID-19 epidemic in Italy: Effects of emergency containment measures. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 10484–10491. [[CrossRef](#)] [[PubMed](#)]
3. WHO. WHO Coronavirus Disease (COVID-19) Dashboard. Available online: <https://covid19.who.int/> (accessed on 22 June 2020).
4. Sanders, J.M.; Monogue, M.L.; Jodlowski, T.Z.; Cutrell, J.B. Pharmacologic Treatments for Coronavirus Disease 2019 (COVID-19): A Review. *JAMA J. Am. Med. Assoc.* **2020**, *323*, 1824–1836. [[CrossRef](#)] [[PubMed](#)]
5. AminJafari, A.; Ghasemi, S. The possible of immunotherapy for COVID-19: A systematic review. *Int. Immunopharmacol.* **2020**, *83*. [[CrossRef](#)] [[PubMed](#)]
6. Zhang, R.; Li, Y.; Zhang, A.L.; Wang, Y.; Molina, M.J. Identifying airborne transmission as the dominant route for the spread of COVID-19. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 14857–14863. [[CrossRef](#)]
7. Xu, H.; Yan, C.; Fu, Q.; Xiao, K.; Yu, Y.; Han, D.; Wang, W.; Cheng, J. Possible environmental effects on the spread of COVID-19 in China. *Sci. Total Environ.* **2020**, *731*. [[CrossRef](#)]
8. Rosario, D.K.A.; Mutz, Y.S.; Bernardes, P.C.; Conte-Junior, C.A. Relationship between COVID-19 and weather: Case study in a tropical country. *Int. J. Hyg. Environ. Health* **2020**, 113587. [[CrossRef](#)]
9. Lin, S.; Fu, Y.; Jia, X.; Ding, S.; Wu, Y.; Huang, Z. Discovering Correlations between the COVID-19 Epidemic Spread and Climate. *Int. J. Environ. Res. Public Health* **2020**, *17*, 7958. [[CrossRef](#)]
10. Al-Quteimat, O.M.; Amer, A.M. The Impact of the COVID-19 Pandemic on Cancer Patients. *Am. J. Clin. Oncol.* **2020**, *43*, 452–455. [[CrossRef](#)]
11. Schofield, J.; Leelarathna, L.; Thabit, H. COVID-19: Impact of and on Diabetes. *Diabetes Ther.* **2020**, *11*, 1429–1435. [[CrossRef](#)]
12. Butt, C.; Gill, J.; Chun, D.; Babu, B.A. Deep learning system to screen coronavirus disease 2019 pneumonia. *Appl. Intell.* **2020**, 1–7. [[CrossRef](#)]
13. De Carvalho Rocha, W.F.; Do Prado, C.B.; Blonder, N. Comparison of chemometric problems in food analysis using non-linear methods. *Molecules* **2020**, *25*, 3025. [[CrossRef](#)] [[PubMed](#)]
14. Haykin, S. *Neural Networks: A Comprehensive Foundation*; Prentice Hall: New York, NY, USA, 2001; ISBN 978-0-02-352761-6.
15. Jiang, C.; Jiang, M.; Xu, Q.; Huang, X. Expectile regression neural network model with applications. *Neurocomputing* **2017**, *247*, 73–86. [[CrossRef](#)]
16. Castañeda-Miranda, A.; Castaño, V.M. Smart frost control in greenhouses by neural networks models. *Comput. Electron. Agric.* **2017**, *137*, 102–114. [[CrossRef](#)]
17. Arora, S.; Taylor, J.W. Rule-based autoregressive moving average models for forecasting load on special days: A case study for France. *Eur. J. Oper. Res.* **2018**, *266*, 259–268. [[CrossRef](#)]
18. Hassan, M.M.; Huda, S.; Yearwood, J.; Jelinek, H.F.; Almogren, A. Multistage fusion approaches based on a generative model and multivariate exponentially weighted moving average for diagnosis of cardiovascular autonomic nerve dysfunction. *Inf. Fusion* **2018**, *41*, 105–118. [[CrossRef](#)]
19. Barrow, D.; Kourentzes, N.; Sandberg, R.; Niklewski, J. Automatic robust estimation for exponential smoothing: Perspectives from statistics and machine learning. *Expert Syst. Appl.* **2020**, *160*, 113637. [[CrossRef](#)]
20. Amo Baffour, A.; Feng, J.; Taylor, E.K. A hybrid artificial neural network-GJR modeling approach to forecasting currency exchange rate volatility. *Neurocomputing* **2019**, *365*, 285–301. [[CrossRef](#)]
21. Castañeda-Miranda, A.; Castaño-Meneses, V.M. Smart frost measurement for anti-disaster intelligent control in greenhouses via embedding IoT and hybrid AI methods. *Meas. J. Int. Meas. Confed.* **2020**, *164*. [[CrossRef](#)]
22. Pradeepkumar, D.; Ravi, V. Soft computing hybrids for FOREX rate prediction: A comprehensive review. *Comput. Oper. Res.* **2018**, *99*, 262–284. [[CrossRef](#)]
23. Castañeda-Miranda, A.; Castaño-Meneses, V.M. Internet of things for smart farming and frost intelligent control in greenhouses. *Comput. Electron. Agric.* **2020**, *176*, 105614. [[CrossRef](#)]
24. Galvan, D.; Cremasco, H.; Gomes Mantovani, A.C.; Bona, E.; Killner, M.; Borsato, D. Kinetic study of the transesterification reaction by artificial neural networks and parametric particle swarm optimization. *Fuel* **2020**, *267*, 1–9. [[CrossRef](#)]
25. Clark, S.; Sisson, S.A.; Sharma, A. Tools for enhancing the application of self-organizing maps in water resources research and engineering. *Adv. Water Resour.* **2020**, *143*, 103676. [[CrossRef](#)]

26. Cremasco, H.; Borsato, D.; Angilelli, K.G.; Galão, O.F.; Bona, E.; Valle, M.E. Application of self-organising maps towards segmentation of soybean samples by determination of inorganic compounds content. *J. Sci. Food Agric.* **2016**, *96*, 306–310. [CrossRef]
27. Kimura, M.; Savada, F.Y.; Tashima, D.L.M.; Romagnoli, É.S.; Chendynski, L.T.; Silva, L.R.C.; Borsato, D. Application of the self-organizing map in the classification of natural antioxidants in commercial biodiesel. *Biofuels* **2018**, 1–6. [CrossRef]
28. Cremasco, H.; Galvan, D.; Angilelli, K.G.; Borsato, D.; de Oliveira, A.G. Influence of film coefficient during multicomponent diffusion–KCL/NaCL in biosolid for static and agitated system using 3D computational simulation. *Food Sci. Technol.* **2019**, *39*, 173–181. [CrossRef]
29. Silva, L.R.C.; Angilelli, K.G.; Cremasco, H.; Romagnoli, É.S.; Galão, O.F.; Borsato, D.; Moraes, L.A.C.; Mandarino, J.M.G. Application of self-organising maps towards segmentation of soybean samples by determination of amino acids concentration. *Plant Physiol. Biochem.* **2016**, *106*, 264–268. [CrossRef]
30. Sanchez, J.L.; Pereira, S.B.G.; Tanamati, A.; Tanamati, A.A.C.; Bona, E. Monitoring industrial hydrogenation of soybean oil using self-organizing maps. *Emir. J. Food Agric.* **2019**, *31*, 779–787. [CrossRef]
31. Melin, P.; Monica, J.C.; Sanchez, D.; Castillo, O. Analysis of Spatial Spread Relationships of Coronavirus (COVID-19) Pandemic in the World using Self Organizing Maps. *Chaos Solitons Fractals* **2020**, *138*, 109917. [CrossRef]
32. Kebonye, N.M.; Eze, P.N.; Kingsley, J.; Gholizadeh, A.; Dajčl, J.; Drábek, O.; Němeček, K.; Borůvka, L. A combined self-organizing map artificial neural networks and conditional Gaussian simulation technique for mapping potentially toxic element hotspots in polluted mining soils. *J. Geochem. Explor.* **2020**, 106680. [CrossRef]
33. Moghadas, S.M.; Shoukat, A.; Fitzpatrick, M.C.; Wells, C.R.; Sah, P.; Pandey, A.; Sachs, J.D.; Wang, Z.; Meyers, L.A.; Singer, B.H.; et al. Projecting hospital utilization during the COVID-19 outbreaks in the United States. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 9122–9126. [CrossRef] [PubMed]
34. Haghshenas, S.S.; Pirouz, B.; Haghshenas, S.S.; Pirouz, B.; Piro, P.; Na, K.S.; Cho, S.E.; Geem, Z.W. Prioritizing and analyzing the role of climate and urban parameters in the confirmed cases of COVID-19 based on artificial intelligence applications. *Int. J. Environ. Res. Public Health* **2020**, *17*, 3730. [CrossRef]
35. Al-Anzi, B.S.; Alenizi, M.; Al Dallal, J.; Abookleesh, F.L.; Ullah, A. An overview of the world current and future assessment of novel COVID-19 trajectory, impact, and potential preventive strategies at healthcare settings. *Int. J. Environ. Res. Public Health* **2020**, *17*, 7016. [CrossRef] [PubMed]
36. Eltoukhy, A.E.E.; Shaban, I.A.; Chan, F.T.S.; Abdel-Aal, M.A.M. Data analytics for predicting covid-19 cases in top affected countries: Observations and recommendations. *Int. J. Environ. Res. Public Health* **2020**, *17*, 7080. [CrossRef] [PubMed]
37. IBGE. Estimated Brazilian Population for 2019. Available online: <https://covid.saude.gov.br/> (accessed on 13 June 2020).
38. MS Coronavirus Pandemic Data in Brazil. Available online: <https://covid.saude.gov.br/> (accessed on 13 June 2020).
39. Federal Government Federal Government of Brazil Transfers Financial Resources. Available online: <https://covid-insumos.saude.gov.br/paineis/insumos/painel.php> (accessed on 13 June 2020).
40. Federal Government Covid-19 Spending Monitoring Panel. Available online: <https://www.tesourotransparente.gov.br/visualizacao/painel-de-monitoramentos-dos-gastos-com-covid-19> (accessed on 13 June 2020).
41. BCB Dollar Exchange Rate. Available online: <https://www.bcb.gov.br/> (accessed on 13 June 2020).
42. IBGE. Interactive Health Maps in Brazil. Available online: <https://mapasinterativos.ibge.gov.br/covid/saude/> (accessed on 13 June 2020).
43. ADHB. Brazilian Human Development Database. Available online: <http://www.atlasbrasil.org.br/2013/pt/download/base/> (accessed on 13 June 2020).
44. Taiyun, W.; Simko, V. R Package “Corrplot”: Visualization of a Correlation Matrix; CRAN, Germany; R Core Team: Vienna, Austria, 2017.
45. Ranney, M.L.; Griffeth, V.; Jha, A.K. Critical Supply Shortages—The Need for Ventilators and Personal Protective Equipment during the Covid-19 Pandemic. *N. Engl. J. Med.* **2020**, *382*, e41. [CrossRef] [PubMed]

46. Carter, L.J.; Garner, L.V.; Smoot, J.W.; Li, Y.; Zhou, Q.; Saveson, C.J.; Sasso, J.M.; Gregg, A.C.; Soares, D.J.; Beskid, T.R.; et al. Assay Techniques and Test Development for COVID-19 Diagnosis. *ACS Cent. Sci.* **2020**, *6*, 591–605. [[CrossRef](#)]
47. Liu, X.; Liu, C.; Liu, G.; Luo, W.; Xia, N. Theranostics COVID-19: Progress in diagnostics, therapy and vaccination. *Theranostics* **2020**, *10*, 7821–7835. [[CrossRef](#)]
48. Wu, R.; Wang, L.; Kuo, H.-C.D.; Shannar, A.; Peter, R.; Chou, P.J.; Li, S.; Hudlikar, R.; Liu, X.; Liu, Z.; et al. An Update on Current Therapeutic Drugs Treating COVID-19. *Curr. Pharmacol. Rep.* **2020**, *6*, 56–70. [[CrossRef](#)]
49. NDPH Low-Cost Dexamethasone Reduces Death by Up to One Third in Hospitalised Patients with Severe Respiratory Complications of COVID-19. Available online: <https://www.recoverytrial.net/news/low-cost-dexamethasone-reduces-death-by-up-to-one-third-in-hospitalised-patients-with-severe-respiratory-complications-of-covid-19> (accessed on 20 June 2020).
50. FDA Coronavirus (COVID-19) Update: FDA Revokes Emergency Use Authorization for Chloroquine and Hydroxychloroquine. Available online: <https://www.fda.gov/news-events/press-announcements/coronavirus-covid-19-update-fda-revokes-emergency-use-authorization-chloroquine-and> (accessed on 20 June 2020).
51. Li, I.W.; Hung, I.F.; To, K.K.; Chan, K.H.; Wong, S.S.Y.; Chan, J.F.; Cheng, V.C.; Tsang, O.T.; Lai, S.T.; Lau, Y.L.; et al. The natural viral load profile of patients with pandemic 2009 influenza A(H1N1) and the effect of oseltamivir treatment. *Chest* **2010**, *137*, 759–768. [[CrossRef](#)]

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).