

## The correspondence between the structure of the terrestrial mobility network and the spreading of COVID-19 in Brazil

A correspondência entre a estrutura da rede de mobilidade terrestre e a propagação da COVID-19 no Brasil

La correspondencia entre la estructura de la red de movilidad terrestre y la propagación de la COVID-19 en Brasil

Vander Luis de Souza Freitas <sup>1</sup>  
Thais Cláudia Roma de Oliveira Konstantyner <sup>2</sup>  
Jeferson Feitosa Mendes <sup>2</sup>  
Cátia Souza do Nascimento Sepetauskas <sup>3</sup>  
Leonardo Bacelar Lima Santos <sup>3</sup>

doi: 10.1590/0102-311X00184820

### Abstract

The inter-cities mobility network is of great importance in understanding outbreaks, especially in Brazil, a continental-dimension country. We adopt the data from the Brazilian Ministry of Health and the terrestrial flow of people between cities from the Brazilian Institute of Geography and Statistics database in two scales: cities from Brazil, without the North region, and from the São Paulo State. **Grounded** on the complex networks approach, and considering that the mobility network serves as a proxy for the SARS-CoV-2 spreading, the nodes and edges represent cities and flows, respectively. Network centrality measures such as strength and degree are ranked and compared to the list of cities, ordered according to the day that they confirmed the first case of COVID-19. The strength measure captures the cities with a higher vulnerability of receiving new cases. Besides, it follows the interiorization process of SARS-CoV-2 in the São Paulo State when the network flows are above specific thresholds. Some countryside cities such as Feira de Santana (Bahia State), Ribeirão Preto (São Paulo State), and Caruaru (Pernambuco State) have strength comparable to states' capitals. Our analysis offers additional tools for understanding and decision support to inter-cities mobility interventions regarding the SARS-CoV-2 and other epidemics.

COVID-19; Public Health Surveillance; Epidemics

### Correspondence

V. L. S. Freitas  
Universidade Federal de Ouro Preto.  
Rua Quatro 786, sala 19, Ouro Preto, MG 35400-000, Brasil.  
vander.freitas@ufop.edu.br

<sup>1</sup> Universidade Federal de Ouro Preto, Ouro Preto, Brasil.

<sup>2</sup> Universidade Federal de São Paulo, São Paulo, Brasil.

<sup>3</sup> Centro Nacional de Monitoramento e Alertas de Desastres Naturais, São José dos Campos, Brasil.



This article is published in Open Access under the Creative Commons Attribution license, which allows use, distribution, and reproduction in any medium, without restrictions, as long as the original work is correctly cited.

## Introduction

The world is currently facing a global public health emergency due to the COVID-19 pandemic, declared on March 11th, 2020 by the World Health Organization (WHO) <sup>1</sup>. As of June 4th, 2020, more than 6.7 million cases have been confirmed worldwide, with almost 400,000 deaths. In Brazil, the first documented case was in the city of São Paulo on February 26th, 2020. Since then, there are about 615,000 confirmed cases and 34,000 deaths in the national territory <sup>2</sup> (Worldometers COVID-19: coronavirus pandemic. <https://www.worldometers.info/coronavirus/>, accessed on 15/May/2020; Ministério da Saúde. Painel coronavírus. <https://covid.saude.gov.br/>, accessed on 14/May/2020).

The inter-cities mobility network serves as a proxy for the transmission network, vital for understanding outbreaks, especially in Brazil, a continental-dimension country <sup>3,4,5,6,7</sup>. The complex network approach <sup>8</sup> emerges as a natural mechanism to handle mobility data computationally, taking areas as nodes (fixed) and movements between origins and destinations as connections (flows). Some networks' measures can be used to find the structurally more vulnerable areas in the context of the current study. The degree of a node is the number of cities that it is connected to, showing the number of possible destinations. The strength captures the total number of people that travel to (or come from) such places in each time frame. From a probability perspective, the cities that receive more people are more vulnerable to SARS-CoV-2. The betweenness centrality, on the other hand, considers the entire network to depict the topological importance of a city in the routes that are more likely to be used.

In this context, this work aims to investigate the correspondences of the measures of networks with the emergence of cities with confirmed cases of COVID-19 in two scales: Brazil and the State of São Paulo. Specifically, we analyze (i) the Brazilian inter-cities mobility networks under different flow thresholds to neglect the lowest-frequency travels, especially in the beginning of the outbreak, when the interiorization of the disease was not yet in progress; and (ii) the correspondence between the statistics of networks and the spreading of COVID-19 in Brazil.

## Methods

The most common mobility data used in studies of this nature in Brazil are the pendular travels, from the 2010 national census <sup>9</sup>. In this research, we use the Brazilian Institute of Geography and Statistics (IBGE) roads data from 2016 <sup>10</sup>, which contains the flows between cities considering terrestrial vehicles in which it is possible to buy a ticket (mainly buses and vans). This information seeks to quantify the interconnection between cities, the movement of attraction that urban centers carry out for the consumption of goods and services, and the long-distance connectivity of Brazilian cities. The North region is not included in this paper, because neither the fluvial nor the air models are covered, and their roles are crucial to understand the spreading process there, especially in the Amazon region. According to an investigation of seroprevalence of antibodies to SARS-CoV-2 <sup>11</sup>, Northern cities are among the ones with the highest values, and six of them are located along a 2,000km stretch of the Amazon river.

The above-cited IBGE data <sup>10</sup> contains the travel frequency (flow) between pairs of Brazilian cities/districts in a general/typical week, considering only origins and destinations, without any information about possible connections between them. The frequencies are aggregated within the round trip, which means that the number of travels from city A to city B is the same as from B to A. We produce two types of undirected networks with a different number  $N$  of nodes to capture actions in two scales (country and state):

- (1)  $N = 4,987$  – Brazil without the North region: nodes are cities and edges are the flow of direct travels between them.
- (2)  $N = 620$  – São Paulo State: a subset of the previous network, containing only cities within the São Paulo State. For simplicity, no further analysis is performed to evaluate the dependency of the network in relation to the state neighboring cities.

Some cities are not present in our networks, due to a simplification that IBGE does: it groups small neighboring municipalities with almost no flow into single nodes. For simplicity, and considering that such places do not contain cases in the first days of the outbreak, they are not individually accounted for in our analysis.

We focus on two versions of each network for certain flow thresholds  $\eta$ , the  $\eta_0$  ( $\eta = 0$ ) that is the original network from the IBGE data and  $\eta_d$  ( $\eta = d$ ). The  $d$  corresponds to the higher flow threshold that produces the network with the largest diameter. The motivation behind  $\eta_d$  is to get a threshold high enough to not consider the least frequent connections and to not disregard the most frequent ones <sup>6,12</sup>.

We also use COVID-19 data from the state daily bulletins and Brazilian Ministry of Health <sup>2</sup>, which are reported by place of residence and notification date. It shows that, until June 4th, 2020, the number of cities with at least one confirmed patient with COVID-19 is 3,851 in the Brazil without the North region network, which corresponds to 77% of the nodes. The analogous for São Paulo is 535 (86% of the nodes). With this data, we track the response of each measure in detecting vulnerable cities according to the evolution of the virus spreading process, as each city notified the first case.

### Complex network measures

The topological degree ( $k$ ) <sup>8</sup> of a node is the number of links it has to other nodes. As here the networks are undirected, there is no distinction between incoming and outgoing edges.

In a connected graph, there is at least one shortest path  $\sigma_{vw}$  between any pair of nodes  $v$  and  $w$ . The betweenness centrality <sup>8</sup> ( $b$ ) of a node  $i$  is the rate of those shortest paths that pass through  $i$ :

$$b_i = \sum_{v \neq w \neq i} \frac{\sigma_{vw}(i)}{\sigma_{vw}} \quad (1)$$

Although it is a pointwise measure, it considers non-local information related to all shortest paths on the network. It is worth highlighting that in this context this centrality index is not a transportation (physical) measure but a mobility (process) one. Besides, both degree and betweenness do not account for the network flows here, but the binary (weightless) networks. The diameter of a network is the distance between the farthest nodes, given by the maximum shortest path.

The strength ( $s$ ) <sup>8</sup> of a node on the other hand is the accumulated flow from incident edges:

$$S_i = \sum_{j=1}^N F_{ij} \quad (2)$$

In which  $F_{ij}$  is the flow between nodes  $i$  and  $j$ .

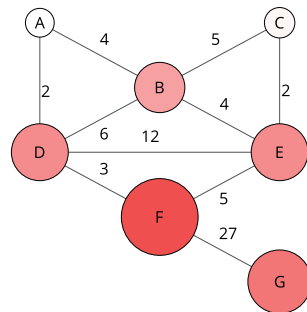
Figure 1 presents illustrations of quite simple networks with the aforementioned measures calculated for each node. The bigger and redder the nodes, the higher the values associated with them. In Figure 1a, node  $F$  has the highest strength, meaning that it receives more flow than any other. Nodes  $B$ ,  $D$  and  $E$  are the most connected, each with exactly four incident edges (Figure 1b). Lastly, node  $F$  of Figure 1c has the higher betweenness value, since all shortest paths with end on  $G$  pass through it. Nodes  $D$  and  $E$  both divide the very same load of shortest paths and have intermediate values.

We assess which of the computed measures ( $s$ ,  $k$ , and  $b$ ) of the mobility networks better approximates the spreading of COVID-19 in Brazil. We compare the top-ranked  $n$  cities of each measure with the  $n$  cities that contain confirmed cases. We vary  $n$  from 1 to the number of cities with confirmed cases to chase the transmission dynamics. In order to verify whether the rate of correspondence between the top-ranked cities from the networks' measures and the cities with COVID-19 cases has statistical significance, we verify what are the results of picking cities at random instead of under the measures' guidance via a hypothesis test with simulated distributions <sup>13</sup>. We perform 100,000 simulations for each  $n$ , choosing  $n$  nodes by chance and monitoring what is the rate of positive cases.

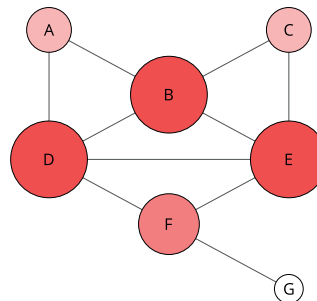
**Figure 1**

Illustration of network measures: strength, degree, and betweenness.

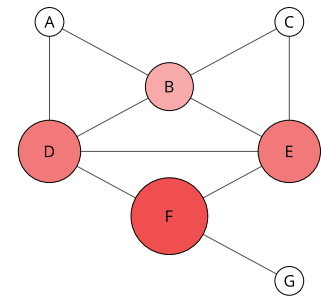
1a) Strength



1b) Degree



1c) Betweenness



Notes: the bigger and redder the nodes, the higher the values associated with them. The network of subfigure a is weighted, while the others are not.

### Geographical visualization

A geographical approach for complex systems analysis is especially important for mobility phenomena. Santos et al.<sup>14</sup> proposed a graph where the nodes have a known geographical location, and the edges have spatial dependence, the (geo)graph. It provides a simple tool to manage, represent, and analyze geographical complex networks in different domains<sup>6,12</sup> and it is used in this work. The geographical manipulation is performed with the PostgreSQL Database Management System (<https://www.postgresql.org/>) and its spatial extension PostGIS. Lastly, the maps are produced using the Geographical Information System ArcGIS (<http://www.esri.com/software/arcgis/index.html>).

### Results

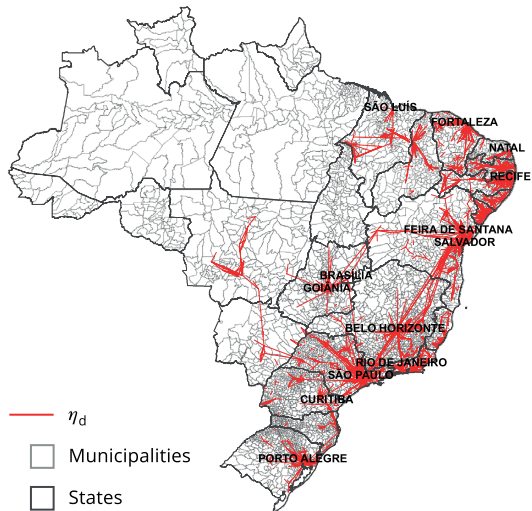
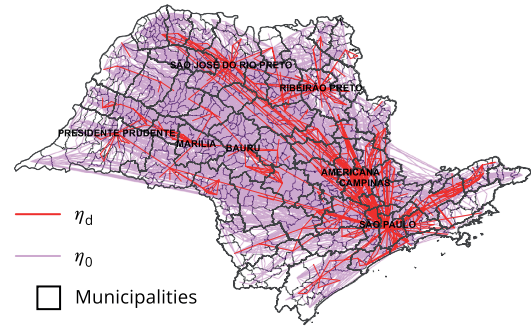
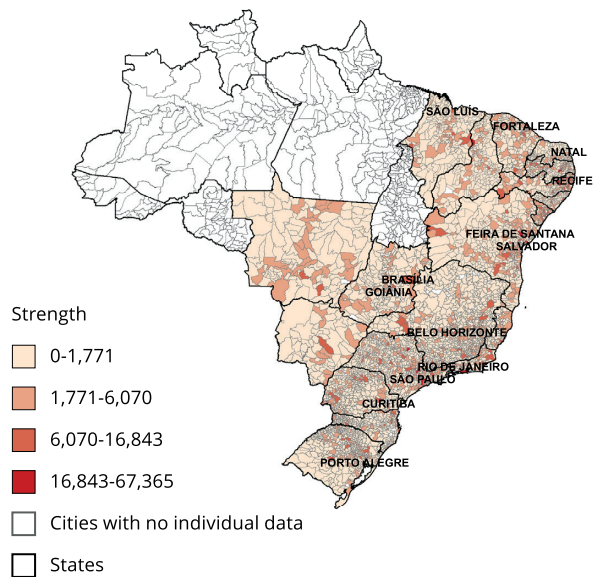
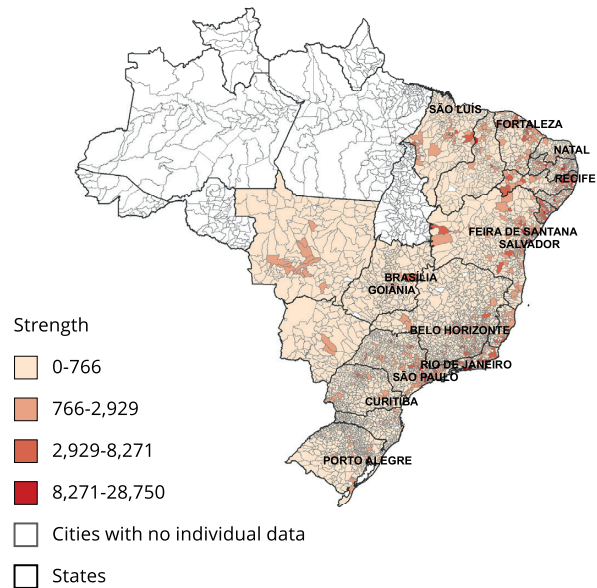
This section presents the results of the topological analysis for the previously mentioned networks. Under  $\eta_0$  (original flows), the Brazil without the North region network has  $N = 4,987$  nodes,  $|E| = 59,453$  edges,  $\langle s \rangle = 1,169.4$  of average strength,  $\langle k \rangle = 23.8$  of average degree and  $\langle b \rangle = 5,219.4$  of average betweenness. When the threshold  $\eta_d$  (the higher threshold with maximum diameter) is considered, it has  $|E| = 2,482$ ,  $\langle s \rangle = 414$ ,  $\langle k \rangle = 1$  and  $\langle b \rangle = 1,385.6$ . The São Paulo State network, on the other hand has  $N = 620$  nodes,  $|E| = 4,796$  edges,  $\langle s \rangle = 1,132.4$  of average strength,  $\langle k \rangle = 15.5$  of average degree and  $\langle b \rangle = 504.2$  of average betweenness, under  $\eta_0$ . The more restricted version, with  $\eta_d$ , possesses  $|E| = 486$ ,  $\langle s \rangle = 535$ ,  $\langle k \rangle = 1.6$  and  $\langle b \rangle = 169.9$ .

Two nodes are connected when between them there is a nonzero flow, which means that the number of connections  $|E|$  decreases for increasing threshold ( $\eta$ ). The resulting networks are undirected and, throughout the paper, both the degree and the betweenness measures do not account for the flows, but weightless edges instead. The diameter of the networks for varying  $\eta$  is computed and the higher threshold with maximum diameter is found for both networks:  $\eta_d = 507.55$  for Brazil without North region and  $\eta_d = 169.9$  for São Paulo State.

Following the (geo)graphs approach, it is possible to visualize nodes and edges of the Brazilian mobility network in the geographical space for  $\eta_d$  in Figure 2a. The edges for  $\eta_0$  are not plotted, because there are more than 59,000 and the visualization is not clear. It is important to highlight

**Figure 2**

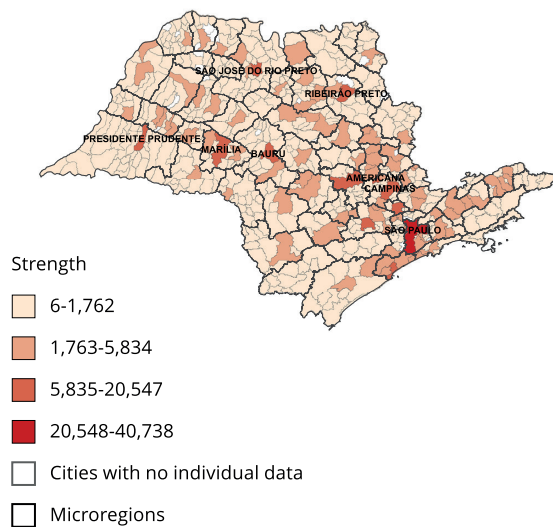
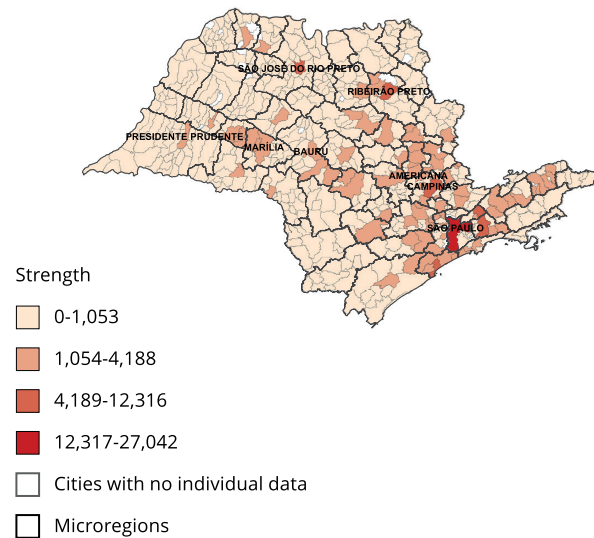
Maps of the Brazil without the North region and São Paulo State networks and the topological strength associated to each node/city for thresholds  $\eta_0$  and  $\eta_d$ .

2a) Brazilian municipalities with the sets of edges for  $\eta_d$ 2b) Municipalities of São Paulo State with the sets of edges for  $\eta_0$  and  $\eta_d$ 2c) Strength for the Brazil without the North region network under  $\eta_0$ 2d) Strength for the Brazil without the North region network under  $\eta_d$ 

(continues)



Figure 2 (continued)

2e) Strength for the São Paulo State network under  $\eta_0$ 2f) Strength for the São Paulo State network under  $\eta_d$ 

some key cities like Belo Horizonte (Minas Gerais State), Rio de Janeiro, São Paulo and Salvador (Bahia State), and the high number of connections between them. Figure 2b depicts the geographical graph regarding the State of São Paulo.

Figure 2c shows the map of the topological degree related to each node/city, considering all original flows ( $\eta_0$ ), and in Figure 2d there is the equivalent for  $\eta_d = 207.55$ . Key cities are labeled in the maps.

Figures 2e and 2f present the strength for the São Paulo State network, with  $\eta_0$  and  $\eta_d$ , respectively. Some cities with high strength also appear in a report<sup>15</sup> of most vulnerable cities to COVID-19 due to their intense traffic of people, namely São Paulo, Campinas, São José do Rio Preto, São José dos Campos, Ribeirão Preto, Santos, Sorocaba, Jaboticabal, Bragança Paulista, Presidente Prudente, Bauru, and many others. Currently, they all have a significant number of confirmed cases.

Figure 3 presents the correspondence of the first  $n$  cities with COVID-19 documented cases with both the simulated data and the top-ranked nodes under  $s$ ,  $k$ , and  $b$ . The  $n$  varies from 1 to 3,851 in Figure 3a and from 1 to 535 in Figure 3b. The gray region represents 95% of the occurrences of rates in the simulations for each  $n$ .

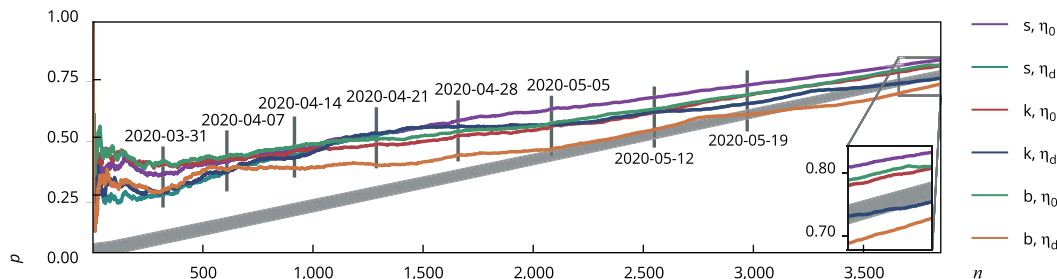
According to Figure 3, on June 4th, about 95% of the simulations have matching rates within  $0.77 \pm 0.01$  for the Brazil without North region network, and the same volume is within  $0.86 \pm 0.03$  for the São Paulo State. The results for node selection during the first days via the network indexes all lie above the gray region, which means that all indexes are a better heuristic than picking nodes at random in the beginning. However, immediately after May 5th,  $b$  with  $\eta_d$  starts to touch the region in São Paulo, having, therefore, results compared to the simulations. It has the worst results for Brazil without North region as well, after a transient.

High-frequency oscillations are perceived in Figure 3 for small  $n$ , but they stabilize afterwards and follow a tendency. In Figure 3a they are pronounced until March 24th ( $n = 150$ ). The matching  $p$  is at maximum in the beginning, because the first documented case was in the city of São Paulo, which is the first ranked city in all measures. The curve then decreases until reaching a region where the oscillations take place. Furthermore, the network quantifiers pose good correspondences already in the beginning of the spreading process, as the gray regions are not touched until  $n$  approaches the number of cities with at least one confirmed patient with COVID-19.

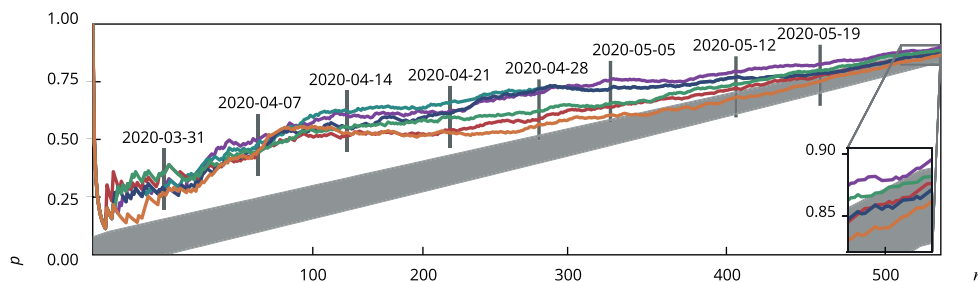
**Figure 3**

Correspondence (rate) between the  $n$  top ranked cities for different network criteria:  $s$ ,  $k$ , and  $b$ , and cities that have at least one patient with COVID-19 in Brazil without the North region and São Paulo State.

3a) Brazil without the North region



3b) São Paulo State



Notes: the gray region represents the correspondences with randomly selected cities. The inset is a zoomed area of the last days until June 4<sup>th</sup>. In both panels, the  $s$  and  $k$  with  $\eta_d$  are overlapping in the end.

Interestingly, on March 31<sup>st</sup>, the high-frequency oscillations start to diminish in São Paulo State. A few days further, after April 7<sup>th</sup>, the betweenness centrality with  $\eta_d$  starts to be a bad predictor for Brazil without North region and then for São Paulo State.

Table 1 enumerates the first twenty ordered cities according to the best-evaluated measures and compares them side-by-side with the first twenty cities with COVID-19 cases in the Brazil without North region network. The best measures for São Paulo State are compared with each other in Table 2 as well. Although in this case  $s$  with  $\eta_0$  presents good correspondences, we present the ones with  $\eta_d$  due to its importance until the end of April. In both networks, although  $p$  produces high-frequency oscillations in the beginning, as shown in Figure 3, the metrics still pose some correspondences with the first confirmed cases.

## Discussion

We present a complex network-based analysis in the Brazilian inter-cities mobility networks towards the identification of cities that are vulnerable to the SARS-CoV-2 spreading. The networks are built with the IBGE terrestrial mobility data from 2016 that have the flow of people between cities in a general/typical week. The cities are modeled as nodes and the flows as weighted edges. The geographical graphs, (geo)graphs, are visualized within Geographical Information Systems.

**Table 1**

Cities with at least one case of COVID-19 in Brazil (Brazil without the North region) in the order they were documented <sup>8</sup>, side-by-side with the top-ranked cities regarding  $s$ ,  $k$  and  $b$  for  $\eta_0$  and  $\eta_d$ .

	Cities ordered according to the appearance of the first case of COVID-19	$s, \eta_0$	$k, \eta_0$	$b, \eta_0$
1	São Paulo (SP)	São Paulo (SP)	São Paulo (SP)	São Paulo (SP)
2	Barra Mansa (RJ)	Belo Horizonte (MG)	Belo Horizonte (MG)	Belo Horizonte (MG)
3	Feira de Santana (BA)	Salvador (BA)	Campinas (SP)	Goiânia (GO)
4	Rio de Janeiro (RJ)	Aracaju (SE)	Brasília (DF)	Brasília (DF)
5	Brasília (DF)	Campinas (SP)	Goiânia (GO)	Rio de Janeiro (RJ)
6	Santana de Parnaíba (SP)	Fortaleza (CE)	Rio de Janeiro (RJ)	Porto Alegre (RS)
7	Divinópolis (MG)	Porto Alegre (RS)	Porto Alegre (RS)	Fortaleza (CE)
8	Maceió (AL)	Campina Grande (PB)	Salvador (BA)	Curitiba (PR)
9	Campo Bom (RS)	Feira de Santana (BA)	Feira de Santana (BA)	Campinas (SP)
10	Porto Alegre (RS)	Recife (PE)	Curitiba (PR)	Natal (RN)
11	Caxias do Sul (RS)	Caruaru (PE)	Cascavel (PR)	Feira de Santana (BA)
12	Cianorte (PR)	Goiânia (GO)	Ribeirão Preto (SP)	Salvador (BA)
13	Curitiba (PR)	Rio de Janeiro (RJ)	Fortaleza (CE)	Teresina (PI)
14	Ferraz de Vasconcelos (SP)	Carpina (PE)	Maringá (PR)	Ribeirão Preto (SP)
15	Florianópolis (SC)	Teresina (PI)	Carazinho (RS)	São Luís (MA)
16	Goiânia (GO)	Ribeirão Preto (SP)	Anápolis (MG)	Campina Grande (PB)
17	Ipatinga (MG)	Curitiba (PR)	São José do Rio Preto (SP)	Recife (PE)
18	Natal (RN)	Limoeiro (PE)	Vitória da Conquista (BA)	João Pessoa (PB)
19	Niterói (RJ)	Paudalho (PE)	Barreiras (BA)	Aracaju (SE)
20	Recife (PE)	Patos (PB)	Campo Grande (MS)	São José do Rio Preto (SP)

States: AL: Alagoas; BA: Bahia; CE: Ceará; DF: Distrito Federal; GO: Goiás; MA: Maranhão; MG: Minas Gerais; MS: Mato Grosso do Sul; PB: Paraíba; PE: Pernambuco; PI: Piauí; PR: Paraná; RJ: Rio de Janeiro; RN: Rio Grande do Norte; RS: Rio Grande do Sul; SC: Santa Catarina; SP: São Paulo.

Notes: the best combination is  $s$  with  $\eta_0$  (second column). Matching cities are colored alike.

Two scales are investigated, the Brazilian cities without the North region, and the State of São Paulo. The former does not account for the North due to the high number of fluvial routes and some intrinsic local characteristics that are not represented with the terrestrial data. The State of São Paulo is crucial in the ongoing pandemic since the first documented case was in the state capital, and it is currently one of the main focuses of the virus spreading.

Three network measures are studied, namely the strength, degree, and betweenness centrality, under two flow thresholds to account for different mobility intensities, the original flow data and networks with only the edges with higher weights. Other network measures were preliminarily tested, including the weighted version of the betweenness centrality. However, the integrals of the correspondence curves of Figure 3 related to the strength, degree, and betweenness are higher than them. Besides, the three chosen measures capture the network properties in two different viewpoints: a local (strength and degree) and a global (betweenness) perspective. We verified that the strength has the best matching to the cities with COVID-19 confirmed cases. Moreover, the strength measure with the original flows showed to be the best option for Brazil. Differently, a more restricted threshold culminates in better correspondences at the beginning of the pandemic in São Paulo. Probably due to the interiorization of the spreading process, a transition is observed after a certain point, around the first days of May 2020, when the original flows have better results as the connections to smaller cities are only present when they are accounted for.

Regarding Table 1, the three measures capture some cities that do not appear in the first column, namely Fortaleza (Ceará State), Salvador (Bahia State), Campinas (São Paulo State), Ribeirão Preto (São Paulo State) and Belo Horizonte (Minas Gerais State). They soon had patients with COVID-19,



**Table 2**

Cities with at least one case of COVID-19 in the State of São Paulo in the order they were documented <sup>8</sup>, side-by-side with the top-ranked cities regarding  $s$ ,  $k$  and  $b$  for  $\eta_0$  and  $\eta_d$ .

	Cities ordered according to the appearance of the first case of COVID-19	$s, \eta_d$	$k, \eta_d$	$b, \eta_0$
1	São Paulo	São Paulo	São Paulo	São Paulo
2	Santana de Parnaíba	Campinas	Campinas	Campinas
3	Ferraz de Vasconcelos	Ribeirão Preto	Ribeirão Preto	São José do Rio Preto
4	Mauá	Peruíbe	São José do Rio Preto	Ribeirão Preto
5	Santo André	São José do Rio Preto	Sorocaba	Presidente Prudente
6	São Bernardo do Campo	Santa Isabel	Bauru	Bauru
7	São Caetano do Sul	Mogi das Cruzes	Presidente Prudente	Sorocaba
8	Guarulhos	Jundiaí	Marília	Piracicaba
9	Barueri	Sorocaba	Tupã	Marília
10	Campinas	Piracicaba	São José dos Campos	Araçatuba
11	Cotia	São José dos Campos	Piracicaba	Americana
12	Jaguariúna	Presidente Prudente	Jundiaí	Assis
13	Osasco	Praia Grande	Taubaté	Catanduva
14	São José do Rio Preto	Lorena	São Roque	Jales
15	São José dos Campos	São Roque	Santa Isabel	Taubaté
16	Suzano	São Carlos	Porto Ferreira	São José dos Campos
17	Taubaté	Taubaté	Peruíbe	Jundiaí
18	Vargem Grande Paulista	São Vicente	Lorena	Franca
19	Hortolândia	Registro	Jaboticabal	Jaboticabal
20	Mogi das Cruzes	Sertãozinho	Guaratinguetá	Santo André

Notes: the best combination is  $s$  with  $\eta_d$  (second column) until the end of April. Matching cities are colored alike.

though. Interestingly, the city of Feira de Santana (Bahia State) appears in all columns – it is the second-largest city of the state and connects the capital to the countryside of Bahia <sup>16</sup>. Oppositely, the city of João Pessoa, capital of Paraíba State does not appear in the top 20 of the second column (best measure – see Figure 3), but two other cities from the state do, namely Campina Grande and Patos. Campina Grande and Patos are among the five richest cities of Paraíba <sup>17</sup>. It should be noted that within the context of an epidemic, such cities are potential super spreaders along with the states' capitals. Five cities of Pernambuco State appear in the second column (best measure – Figure 3), namely Caruaru, Carpina, Limoeiro, Paudalho, and Recife. Pernambuco is currently ranked as the second state in the number of confirmed cases of the Northeast region <sup>2</sup>.

Table 2, as in Table 1, also displays cities that are captured by the three rightmost columns that do not appear in the first, showing their high level of vulnerability: Ribeirão Preto, Jundiaí, Sorocaba, Piracicaba, and Presidente Prudente. They all have documented cases before June 4th. Our study also captured the most influential cities that had cases already in the beginning, like São Paulo, Campinas, São José do Rio Preto, São José dos Campos and Taubaté. Other cities appear in the second column (best metric) but not in the first: Praia Grande, São Vicente, São Carlos, Registro, Sertãozinho.

Due to their importance in mobility, many cities of Table 2, especially in the second column, appear in the report <sup>15</sup> on the vulnerability of microregions of São Paulo State to the SARS-CoV-2 pandemic of April 5th either as potential spreaders or places with a high probability of receiving new cases. They all have notified cases by June 4th and some have the highest numbers of the São Paulo State <sup>18</sup>.

Both  $s$  and  $b$  with  $\eta_0$  pose good results at the beginning of the pandemics for the Brazil without North region network, but  $s$  alone started to be the best predictor from the end of April. The most important cities, due to their high flow of travelers and their role in the most used routes, are reached

first, followed by those with smaller flows, probably because of the interiorization of the virus – the outbreak reaching the countryside cities. This behavior is even more pronounced in São Paulo State, in which  $s$  under  $\eta_d$  is the best option at first, neglecting lower flow venues, especially in April, but the  $\eta_0$  started to be the best option from May onwards.

In the ongoing pandemics, from May 1st, the  $s$  index with  $\eta_0$  is currently the best predictor and may help to figure out which countryside cities are about to receive new cases. Moreover, it may help in the following waves of the disease. In the case of another pandemic, one could first compute the strength of the networks according to the last updated data from IBGE and identify the top-ranked cities. In Brazil, it is enough checking on strength at the original data, as we presented, since it produces similar results as the betweenness centrality and is computationally cheaper to obtain. Regarding the State of São Paulo, one better checks on the strength index with threshold  $\eta_d$  in the first weeks and only then switch to  $\eta_0$ . As our results show, the correspondence has statistical significance and, along with other information about the regions such as where are the first notified cases, the pandemic could be closely traced.

It is worth mentioning that the COVID-19 data comes from the Brazilian Health Surveillance System, which is fed with data provided by each city in the country. The information update is a complex and dynamic process and there may be delays or errors in the data transfer. Moreover, considering the size and heterogeneity of Brazil, it is important to highlight that there are differences in the capacity to detect cases opportunely and in the quality of the information<sup>19</sup>. On the other hand, in late January 2020, almost a month before the first Brazilian COVID-19, the epidemiological surveillance guidelines and the National Contingency Plan for COVID-19 were published. One of the main objectives of these documents was to provide early guidance to the Brazilian Unified National Health System (SUS) service network, to act in the identification of COVID-19 cases<sup>20</sup>. Besides the limitations in the health surveillance system, there is a lack of information about possible intermediate stops between origins and destinations in the IBGE data, as it gives only the travelers' initial and final positions.

As future work, we intend to analyze fluvial and aerial mobility data as well, as they include valuable information about the transport of people and goods. The former is fundamental to the discussion of the dynamics for the Brazilian North region, especially the Amazon, and the latter captures long-range connections which are relevant in a possible future moment of re-emergence of the disease in the country, especially by foreign travelers. Lastly, one could check for correspondences between the measures of networks and data from other epidemics, and analyze control measures based on topological properties associated with the mobility network<sup>21</sup>.

## Contributors

V. L. S. Freitas and L. B. L. Santos conceived the original idea, developed the methodology, collected the data, interpreted results and wrote the paper. T. C. R. O. Konstantyner developed the methodology, interpreted results and wrote the paper. J. F. Mendes and C. S. N. Sepetauskas developed the methodology and reviewed the paper.

## Acknowledgments

We would like to thank Jussara R. Angelo (Oswaldo Cruz Foundation, Rio de Janeiro, Brazil) for the valuable discussions.

## Additional informations

ORCID: Vander Luis de Souza Freitas (0000-0001-7989-0816); Thais Cláudia Roma de Oliveira Konstantyner (0000-0002-2289-1866); Jeferson Feitosa Mendes (0000-0002-0943-4820); Cátia Souza do Nascimento Sepetauskas (0000-0003-1194-9564); Leonardo Bacelar Lima Santos (0000-0002-3129-772X).

## References

- Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N Engl J Med* 2020; 382:1199-207.
- Cota W. Monitoring the number of COVID-19 cases and deaths in Brazil at municipal and federative units level. *SciELO Preprints* 2020; 7 may. <https://doi.org/10.1590/SciELOPreprints.362>.
- Lana RM, Gomes MFC, Lima TFM, Honorio NA, Codeço CT. The introduction of dengue follows transportation infrastructure changes in the state of Acre, Brazil: a network-based analysis. *PLoS Negl Trop Dis* 2017; 11:e0006070.
- Angelo JR, Katsuragawa TH, Sabroza PC, Carvalho LAS, Silva LHP, Nobre CA. The role of spatial mobility in malaria transmission in the Brazilian Amazon: the case of Porto Velho municipality, Rondônia, Brazil (2010-2012). *PLoS One* 2017; 12:e0172330.
- Saba H, Moret MA, Barreto FR, Araújo MLV, Jorge EMF, Nascimento Filho AS, et al. Relevance of transportation to correlations among criticality, physical means of propagation, and distribution of dengue fever cases in the state of Bahia. *Sci Total Environ* 2018; 618:971-6.
- Santos LBL, Carvalho LM, Seron W, Coelho FC, Macau EE, Quiles MG, et al. How do urban mobility (geo) graph's topological properties fill a map? *Appl Netw Sci* 2019; 4:91.
- Coelho FC, Lana RM, Cruz OG, Codeço CT, Villela D, Bastos LS, et al. Assessing the potential impact of COVID-19 in Brazil: mobility, morbidity and the burden on the health care system. *medRxiv* 2020; 26 mar. <https://www.medrxiv.org/content/10.1101/2020.03.19.20039131v2>.
- Estrada E. The structure of complex networks: theory and applications. Oxford: Oxford University Press; 2012.
- Instituto Brasileiro de Geografia e Estatística. Censo Demográfico 2010: resultados gerais da amostra. <https://censo2010.ibge.gov.br/resultados.html> (accessed on 10/Jun/2020).
- Cavararo R. Ligações rodoviárias e hidroviárias: 2016. Rio de Janeiro: Instituto Brasileiro de Geografia e Estatística; 2017.
- Hallal P, Hartwig F, Horta B, Victora GD, Silveira M, Struchiner C, et al. Remarkable variability in SARS-CoV-2 antibodies across Brazilian regions: nationwide serological household survey in 27 states. *medRxiv* 2020; 30 may. <https://www.medrxiv.org/content/10.1101/2020.05.30.20117531v1>.
- Ceron W, Santos LB, Dolif Neto G, Quiles MG, Candido OA. Community detection in very high-resolution meteorological networks. *IEEE Geosci Remote Sens Lett* 2019; <https://ieeexplore.ieee.org/document/8930617>.
- Diez DM, Barr CD, Çetinkaya-Rundel M. Introductory statistics with randomization and simulation. <https://www.openintro.org/book/isrs/> (accessed on 09/Jul/2020).
- Santos LBL, Jorge AA, Rossato M, Santos JD, Candido OA, Seron W, et al. (Geo)graphs-complex networks as a shapefile of nodes and a shapefile of edges for different applications. *arXiv* 2017; 16 nov. <https://arxiv.org/abs/1711.05879>.
- Guimarães Jr. PR, Muniz D, Giacobelli L, Maia K, Gaiarsa M, Assis AP, et al. Vulnerabilidade das microrregiões do Estado de São Paulo à pandemia do novo coronavírus (SARS-CoV-2). *SciELO Preprints* 2020; 15 apr. <https://doi.org/10.1590/SciELOPreprints.49>.
- IBGE divulga as estimativas da população dos municípios para 2019. Agência IBGE Notícias 2019; 28 aug. <https://agenciadenoticias.ibge.gov.br/agencia-sala-de-imprensa/2013-agencia-de-noticias/releases/25278-ibge-divulga-as-estimativas-da-populacao-dos-municipios-para-2019>.
- Instituto Brasileiro de Geografia e Estatística. Produto Interno Bruto – PIB. <https://www.ibge.gov.br/explica/pib.php> (accessed on 10/Jun/2020).
- Secretaria da Saúde, Governo do Estado de São Paulo. Novo coronavírus (COVID-19): situação epidemiológica. [http://www.saude.sp.gov.br/resources/cve-centro-de-vigilancia-epidemiologica/areas-de-vigilancia/doencas-de-transmissao-respiratoria/coronavirus/coronavirus010520\\_65situacao\\_epidemiologica.pdf](http://www.saude.sp.gov.br/resources/cve-centro-de-vigilancia-epidemiologica/areas-de-vigilancia/doencas-de-transmissao-respiratoria/coronavirus/coronavirus010520_65situacao_epidemiologica.pdf) (accessed on 10/Jun/2020).
- Teixeira MG, Costa MCN, Carmo EH, Oliveira WK, Penna GO. Vigilância em saúde no SUS – construção, efeitos e perspectivas. *Ciênc Saúde Colet* 2018; 23:1811-8.
- Croda J, Oliveira WK, Frutuoso RL, Mandetta LH, Baia-da-Silva DC, Brito-Sousa JD, et al. COVID-19 in Brazil: advantages of a socialized unified health system and preparation to contain cases. *Rev Soc Bras Med Trop* 2020; 53:e20200167.
- Freitas VLS, Feitosa J, Sepetauskas CSN, Santos LBL. Robustness analysis in intercities mobility network: modeling municipal, state and federal initiatives as failures and attacks. *arXiv* 2020; 8 apr. <https://arxiv.org/abs/2004.02648>.

## Resumo

*A rede de mobilidade intermunicipal é de suma importância para a compreensão de surtos, sobretudo no Brasil, um país com dimensões continentais. Os autores adotaram os dados do Ministério da Saúde e informações sobre o fluxo de pessoas entre cidades, da base de dados do Instituto Brasileiro de Geografia e Estatística, em duas escalas: cidades brasileiras, sem a região Norte, e do Estado de São Paulo. Com base na abordagem de redes complexas, e considerando que a rede de mobilidade serve como proxy para a propagação do SARS-CoV-2, os nós e arestas representam cidades e fluxos, respectivamente. As medidas de centralidade de rede, como força e grau, são ranqueadas e comparadas à lista das cidades, de acordo com o dia da confirmação do primeiro caso de COVID-19. A medida de força capta as cidades com maior vulnerabilidade à pandemia, além de acompanhar o processo de interiorização do SARS-CoV-2 no Estado de São Paulo quando os fluxos de rede estão acima de limiares específicos. Algumas cidades do interior, como Feira de Santana (Bahia), Ribeirão Preto (São Paulo) e Caruaru (Pernambuco) mostram forças comparáveis às capitais estaduais. Nossa análise oferece ferramentas adicionais para a compreensão e o apoio para a tomada de decisões sobre intervenções na mobilidade intermunicipal em relação ao SARS-CoV-2 e outras epidemias.*

COVID-19; Vigilância em Saúde Pública; Epidemias

## Resumen

*La red de movilidad entre ciudades es de vital importancia para la comprensión de los brotes, especialmente en Brasil, un país con dimensiones continentales. Consecuimos los datos del Ministerio de Salud Brasileño y el flujo terrestre de gente entre ciudades a partir de la base de datos del Instituto Brasileño de Geografía y Estadística en dos escalas: ciudades de Brasil, sin la región Norte, y Estado de São Paulo. Basado en un planteamiento de redes complejas, y considerando que la movilidad de la red sirve como un proxy para la propagación del SARS-CoV-2, los nodos y extremos representan ciudades y flujos, respectivamente. Las medidas de centralidad de la red como la fuerza y el grado se clasificaron y compararon con la lista de ciudades, ordenadas según el día en que confirmaron el primer caso de COVID-19. La medida de fuerza captura las ciudades con la mayor vulnerabilidad en recibir nuevos casos. Asimismo, le sigue la interiorización del proceso de SARS-CoV-2 en el Estado de São Paulo, cuando los flujos de la red están por encima de determinados umbrales. Algunas ciudades en áreas rurales como Feira de Santana (Estado de Bahía), Ribeirão Preto (Estado de São Paulo), y Caruaru (Estado de Pernambuco) poseen una fuerza comparable a las capitales de los estados. Nuestro análisis ofrece herramientas adicionales para la comprensión y apoyo en la toma de decisiones, respecto a las intervenciones de movilidad entre ciudades, en relación con el SARS-CoV-2 y otras epidemias.*

COVID-19; Vigilancia en Salud Pública; Epidemias

Submitted on 27/Jun/2020

Final version resubmitted on 15/Jul/2020

Approved on 20/Jul/2020