Identification of risk areas as a method of surveillance of dengue cases

Gesiel Rios Lopes[1], Karina Jorge Pelarigo[1], Alexandre C. B. Delbem[1]

[1] Institute of Mathematical and Computer Sciences
University of São Paulo (ICMC-USP)
São Carlos-SP, Brazil

gesielrios@usp.br, karoufscar@hotmail.com, acbd@icmc.usp.br

Abstract. *Identifying spatial clusters of risk for dengue cases according to social vulnerability constitutes a powerful tool for effective epidemiological and urban management. In this way, this work carries out an ecological study that considered confirmed cases of dengue and actions of endemic agents in the municipality of São Carlos-SP, in the year 2019, through the application of the spatial scan technique for classification of the risk areas, computing the relative risk (RR), with a confidence interval of 95% (CI95%:) and the São Paulo Social Vulnerability Index (IPVS) to characterize these areas. Seven clusters were identified, two of which were high risk (RR=37.54 / RR=33.39), with the highest risk located in a region with high vulnerability and the second in a region with very low vulnerability. These results provide information that allows the targeting of specific control actions from the early detection of cases in places with greater dengue transmissibility.*

1. Introduction

Dengue is an acute viral infection transmitted by arthropods, mainly by the *Aedes aegypti* mosquito. It is the most frequent arboviral disease worldwide, occurring mainly in the tropics and subtropics, with an estimated burden of 390 million cases per year, of which 96 million cases manifest symptomatically, 2 million cases develop severe disease, and 21,000 deaths [do Carmo et al. 2020].

The worldwide spread of dengue is a complex problem, which can be accelerated by several factors, such as climate change, population growth, rapid and unplanned urbanization, movement of people for commerce, tourism or forced by natural disasters, weaknesses in public health and vector control programs. From 2010 to 2019, more than 16 million dengue cases were reported across the American continent, and about 10 million cases (62%) were reported in Brazil alone [Canal et al. 2018, do Carmo et al. 2020].

Dengue has a wide geographic distribution in the country, and, despite the intensification of control measures, there has been an increase in the number of severe cases, hospitalizations, and deaths in recent years, and one of the causes identified as responsible is directly related to the rampant growth of cities, accompanied by due to the lack of awareness of the population in the elimination of mosquito breeding sites, which can be any container that accumulates rainwater [de Melo et al. 2012, Canal et al. 2018].

The purpose of this work is to identify clusters of risk through spatial scanning statistics, developed by [Kulldorff and Nagarwalla 1995], based on notifications of dengue cases and actions to combat breeding sites for the *Aedes aegypti* mosquito carried out by the endemic agents, aiming to improve the visualization of the spatial distribution of dengue cases and the breeding sites of the *Aedes aegypti* mosquito, allowing health authorities to quickly identify priority areas and better direct efforts to control dengue [Estallo et al. 2014, de Melo et al. 2012, Canal et al. 2018]. Similar processes have been proposed to combat malaria in the Brazilian Amazon [Da Silva et al. 2010], and tuberculosis in the municipality of São Carlos-SP [Arroyo et al. 2017].

2. Material and methods

2.1. Study area

This is a geographic, ecological study in which the unit of analysis used was the census tracts of the urban area of the municipality of São Carlos, a medium-sized city in the interior of the state of São Paulo, Southeast Brazil. Located in the central-eastern region of the state (Figure 1), specifically at the coordinates 22°1'4" South latitude and 47°53'27" West latitude, São Carlos had an area total territorial area of 1,136,907 km^2, average altitude of 856 meters, the population density of 195.15 inhabitants/km² and the resident population of 221,950 inhabitants in 2010. Regarding socioeconomic aspects, the municipality had an index Gini index of 0.63, Human Development Index (HDI) of 0.805, and a gross domestic product of R$6,712,498.00 for the same year in 2010 [IBGE 2022].

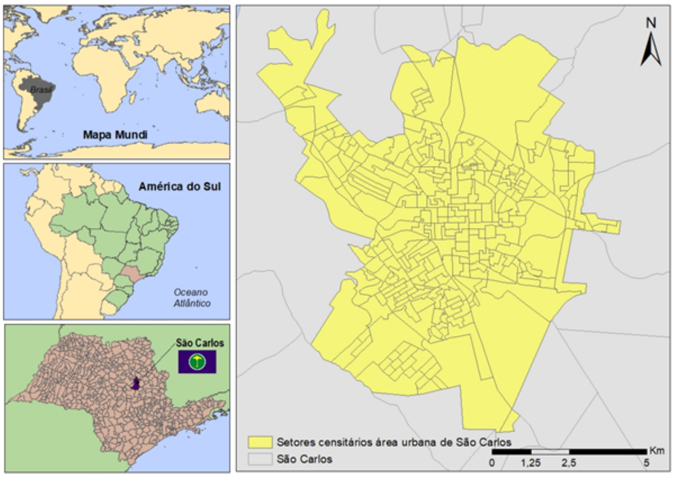


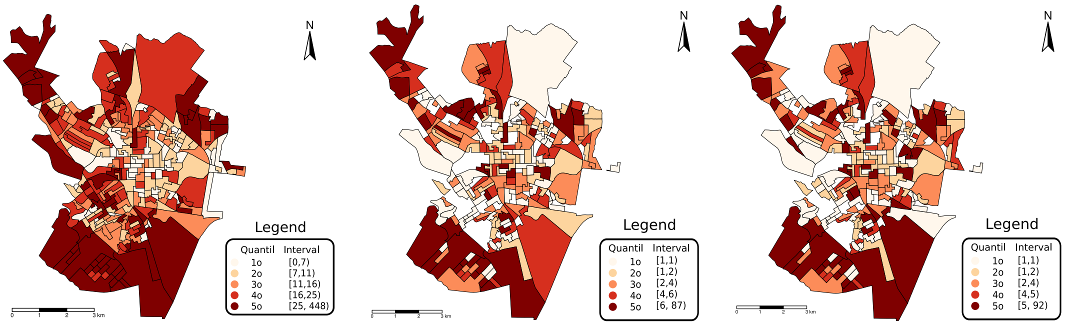
Figure 1. Municipality of São Carlos-SP and its urban perimeter.

2.2. Data source

All reported and confirmed cases of dengue from the Information System on Notifiable Diseases - Sinan Dengue/Chikungunya of residents of the urban area of the municipality of São Carlos-SP from January 1 to December 31, 2019, and information on the presence of breeding sites for the *Aedes aegypti* mosquito with larvae contained in the Bulletin of Surveillance and Control Activities of the places visited by endemic agents in two activities, called house-house and block, also in 2019.

For data analysis, the data was initially georeferenced for the notifications of dengue cases and the places visited by endemic agents. For this procedure, the authors developed a python *script* using the Google Maps geolocation API to obtain the respective geographic coordinates (latitude and longitude) of the notified addresses in the cases and the visits of endemic agents. After geolocation, the cases and actions of endemic agents are aggregated with the digital mesh of the IBGE's urban census sectors of the São Carlos through a *join spatial* function, which crosses the layers. Cases with an unspecific address, duplicates, and residents in the rural area of the municipality were excluded.

The usual way of presenting aggregated data by area is through choropleth maps with the spatial pattern of the phenomenon [Câmara et al. 2004]. Thus, Figure 2(a) presents the spatial distribution of confirmed cases of dengue, Figure 2(b) the spatial distribution with the breeding sites of the mosquito *Aedes aegypti* with larvae for blocking activity, and Figure 2(c) we have the spatial distribution with the mosquito breeding sites *Aedes aegypti* with larvae for the house-house activity. For better data visualization, aggregated data from all maps were sliced into five groups according to their quintiles to form classes with an approximate number of assigned features.



(a) (b) (c)

Figure 2. Choropletic representation of the spatial distribution: (a) of confirmed cases of confirmed dengue in 2019, (b) of breeding sites for the *Aedes aegypti* mosquito with larvae for the blocking activity and (c) of mosquito breeding sites *Aedes aegypti* with larvae for the house-house activity.

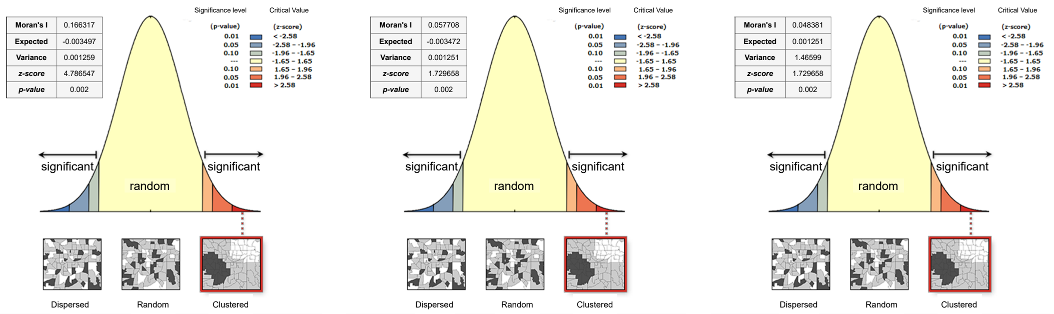
To make sure that the data aggregations in the census sectors represent a phenomenon from a spatial point of view, the spatial autocorrelation was calculated using the global Moran index [Moran 1950, Li et al. 2007] considering only the first neighborhood level, and it works as a test whose null hypothesis is data independence. Global indicators such as Moran I provide a single value to measure spatial association for the entire dataset, which helps characterize the region.

The global Moran index is given by Equation 1.

$$I = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}(z_i - \bar{z})(z_i - \bar{z})}{\sum_{i=1}^{n}(z_i - \bar{z})^2} \qquad (1)$$

where n is the number of areas, z_i is the value of the attribute considered in the area i, z is the average of the assigned values in the study region, and w_{ij} the elements from the normalized spatial proximity matrix [Moran 1950, Li et al. 2007]

The test result for reported cases of dengue can be seen in Figure 3(a), Figure 3(b) presents the test result for blocking activity, and already Figure 3(c) we have the test result for the house-house activity. Given the *z-score* of 4.786547 for dengue notifications, 1.729658 for blocking activity, and 1.46599 for house-to-house activity, respectively, for both analyzes, there is a probability less than 1% that the pattern clustering may be an unexpected result. Thus, the dataset's spatial distribution of high and low values is more spatially clustered than expected if the underlying spatial processes were random, thus rejecting the null hypothesis [Câmara et al. 2004].



(a) (b) (c)

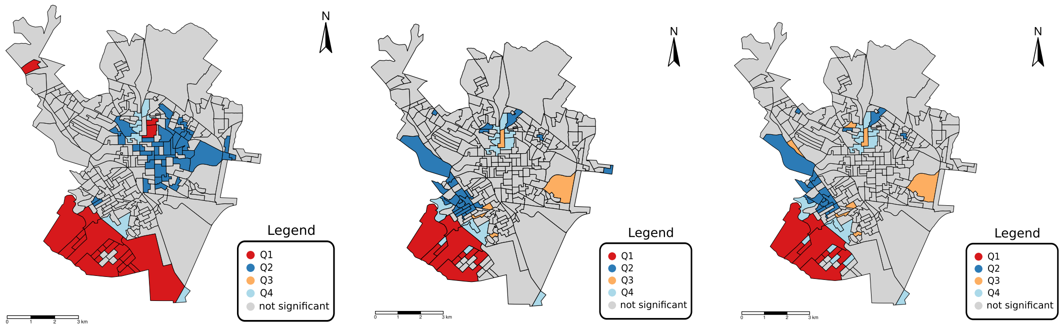
Figure 3. Analysis of the Global Moran Index: (a) of confirmed cases of dengue in 2019, (b) of mosquito breeding sites *Aedes aegypti* with larvae for the activity of blocking and (c) of mosquito breeding sites *Aedes aegypti* with larvae for the house-house activity.

Once the global dependence was verified, the Local Spatial Association Index (LISA) was calculated. LISA is a decomposition of the Moran's I, in which it is possible to analyze the local pattern of spatial data. LISA can be expressed for each area i from the normalized values of the attribute's z_i by Equation 2.

$$I_i = \frac{z_i \sum_{j=1}^{n} w_{ij} z_j}{\sum_{j=1}^{n} w_{ij} z_j^2} \tag{2}$$

Based on the LISA, the census tracts with the aggregated data are positioned in the quadrants of the Moran scatterplot as follows: Q1 (high-high), census tracts where the attribute value and the mean value of neighbors are above average as a whole, and which are therefore considered to be the highest priority for intervention; Q2 (low-low), the attribute value and the mean of the neighbors are below the mean of the set; Q3 (high-low), the attribute value is greater than that of the neighbors and the mean of the neighbors is less than the set; and Q4 (low-high), the attribute value is lower than that of the neighbors and the average of the neighbors is higher than the average of the set. Municipalities classified as high-low and low-high have intermediate priority [Monteiro et al. 2004, Andrade et al. 2007].

Figure 4(a) presents the result of the analysis of the clusters for dengue notifications, Figure 4(b) presents the result of the analysis for the blocking activity and Figure 4(c) we have the result for the house-house activity. It is possible to notice in both analyzes that the census sectors classified as Q1 (high-high), therefore the most critical, are located in the southern region of the municipality.



(a) (b) (c)

Figure 4. Cluster analysis according to the Moran I Local Index: (a) of confirmed cases of dengue in 2019, (b) of breeding sites for the *Aedes aegypti* mosquito with larvae for the blocking activity and (c) of mosquito breeding sites *Aedes aegypti* with larvae for the house-house activity.

2.3. The spatio-temporal scanning technique

The spatial scanning statistical technique was developed by [Kulldorff and Nagarwalla 1995], as already mentioned, in order to identify and locate clusters of hazards present in a given study region. For this analysis, separate spreadsheets were created to analyze dengue cases and the actions of endemic agents, considering the code of the census sector and the number of occurrences of each event.

Thus, the null hypothesis (H_0) against the alternative hypothesis (H_1) was tested differently among the diseases, highlighting that H_0 assumes that there are no clusters areas for dengue, that is, the population has the same probability of contracting dengue and H_1 assumes that one or more regions z are areas in which there would be the greater or lesser probability of contracting the diseases, compared to the that are outside that area.

In order to identify purely spatial clusters in which the distribution is heterogeneous. The events are rare in relation to the population. The discrete Poisson model will be used with requirements of non-overlapping geographical clusters, clusters with a circular shape, 999 replications, and the size of the exposed population will be stipulated by the Gini coefficient released by the software itself. In this model, the number of cases is compared to the baseline population data, and the expected number of cases in each unit is proportional to the size of the population at risk [Alves et al. 2019].

The relative risk (RR) of each cluster will be calculated, allowing the comparison of information in different areas, indicating the intensity of dengue occurrence or of having breeding sites for the *Aedes aegypti* mosquito with larvae in the municipality of São Carlos. It should be noted that the RR will be defined as the risk of having dengue or the existence of breeding sites for the *Aedes aegypti* mosquito with larvae in a risk area of the municipality about the risk of having the disease or of having breeding sites of the mosquito *Aedes aegypti* with larvae outside this area.

Areas with $p\text{-}value < 0.05$ will be considered statistically significant. The confidence interval will be calculated and estimated at 95% [Arroyo et al. 2017].

Identifying the clusters' RR provides the comparison of information in distinct areas, as the effects of different populations are disregarded, thus resulting in the intensity of occurrence of the phenomenon under analysis throughout the study area. [Arroyo et al. 2017].

The cluster detection analyzes were performed using *software* SaTScan[1] version 9.7.

3. Results

Sinan Dengue/Chikungunya reported 10,451 cases of dengue in the municipality of São Carlos in 2019, with 5,878 (56%) confirmed cases.

Applying the spatial scanning statistics for confirmed dengue cases, seven statistically significant clusters were detected (Figure 5(a)). In Table 1, the characteristics of these clusters are presented.

When looking at Figure 2(a), it is possible to identify that the sectors with the highest number of cases are found in the southern region of the municipality. When applying the spatial *scan*, it was possible to detect that the census sectors that composed the purely spatial risk clusters come from the Cidade Aracy neighborhood, also located in the southern region of the city, and from the Jardim Tangará neighborhood, this one in the northeast region of São Carlos.

On the other hand, the present analysis resulted in a cluster of low relative risk for dengue cases, with these census sectors located in the Vila Prado, Vila Boa Vista, Vila Carmem, and Jardim Beatriz neighborhoods, located in the southwest region of the municipality. These locations had an incidence rate below average, i.e., the number of space cases was lower than in any other region in the municipality, constituting areas of protection for infection by dengue cases.

Table 1. Characteristics of clusters are statistically significant in terms of risk for confirmed cases of dengue, according to the spatial scan, in the municipality of São Carlos-SP, 2019.

Cluster	Number of census sectors	Population	Number of cases	Number of expected cases	RR
1	1	553	440	14,26	33,39
2	1	138	127	3,46	37,54
3	20	21.711	962	563,55	1,85
4	1	462	63	12,31	5,16
5	19	10.860	445	276,37	1,66
6	1	527	61	13,61	4,52
7	1	80	19	2,11	9,05

In the Block action, 55,863 households were identified with containers capable of serving as breeding sites for the *Aedes aegypti* mosquito, of which 1,046 (1.87%) contained larvae (Figure 2(b)). When applying the spatial scan statistic, it was possible to detect seven statistically significant clusters (Figure 5(b)). In Table 2, the characteristics of these clusters are presented.

[1]The *software* SaTScan is a free *software* widely used by important Centers for Health Studies and can be obtained directly from textitsite from the developer group [Alves et al. 2019]

As in the clusters of risk for confirmed cases of dengue, the census sectors that made up the clusters of purely spatial risk for the blocking activity come from the region of Cidade Aracy, located in the city's southern region.

Table 2. Characteristics of statistically significant clusters for larvae in the blocking action, according to the spatial scan, in the municipality of São Carlos-SP, 2019.

Cluster	Number of census sectors	Population	Number of cases	Number of expected cases	RR
1	1	553	87	2,59	36,63
2	1	138	41	0,65	65,99
3	16	5.847	84	27,43	3,25
4	2	2.340	48	10,98	4,54
5	1	462	13	2,17	6,06
6	2	1.335	21	6,26	3,40
7	1	614	10	2,88	3,50

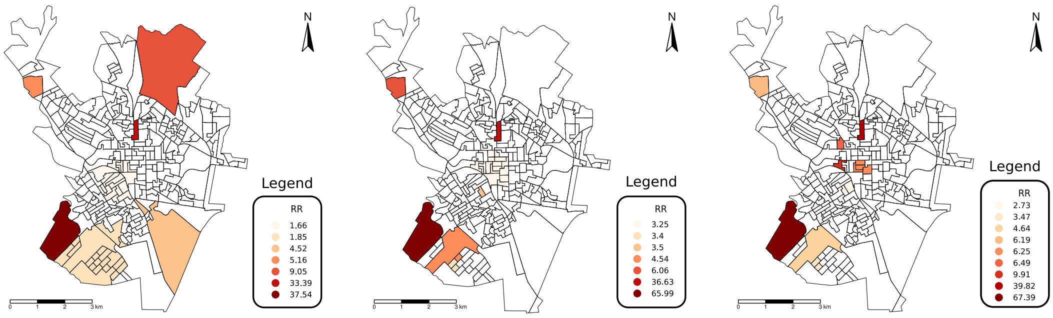
As for the house-house activity, 50,384 homes with possible breeding sites for the *Aedes aegypti* mosquito were visited, with 1,029 (2.04%) homes containing containers with larvae (Figure 2(c)). Figure 5(c) shows the nine statistically significant spatial clusters identified when applying spatial scanning statistics for house-to-house activities where breeding sites with mosquito larvae were found *Aedes aegypti* . In Table 3, the characteristics of these clusters are presented.

The census sectors that made up the purely spatial risk clusters are very similar to those that made up the purely spatial risk clusters for the blocking activity and are located in the Cidade Aracy neighborhood region.

Table 3. Characteristics of statistically significant clusters for larvae in house-to-house action, according to spatial scan, in São Carlos-SP, 2019.

Cluster	Number of census sectors	Population	Number of cases	Number of expected cases	RR
1	1	553	92	2,54	39,82
2	1	138	41	0,63	67,39
3	1	631	28	2,90	9,91
4	2	2.340	48	10,76	4,64
5	4	708	20	3,26	6,25
6	1	462	13	2,12	6,19
7	2	1.335	21	6,14	3,47
8	1	202	6	0,93	6,49
9	2	1.289	16	5,93	2,73

To describe the social vulnerability of the clusters found, data from the Fundação Sistema Estadual de Análise de Dados (SEADE), referring to the São Paulo Social Vulnerability Index (IPVS) for 2010, were used. This index classifies the census sectors based on a combination of the demographic and socioeconomic dimensions and identifies



(a) (b) (c)

Figure 5. Spatial clusters of spatial scan statistics analysis: (a) confirmed cases of dengue, (b) blocking activities and (c) house-house activities, both controlled by the population of the census sectors and by their sex and age distribution, in the city of São Carlos-SP.

the specific factors that produce the deterioration of living conditions in a community, helping to define priorities for the care of the most vulnerable population[SEADE 2010].

The IPVS incorporates the following indicators: number of inhabitants; average nominal income of households; the average age of heads of households; percentage of heads of households under 30 years of age, female heads of households under 30 years of age, and the share of children under six years of age, over the denominator of the total inhabitants of each of these segments [SEADE 2010], characterizing the census sectors in seven groups: Group 1–extremely low vulnerability; Group 2–very low vulnerability; Group 3–low vulnerability; Group 4–medium vulnerability; Group 5–high vulnerability; Group 6–very high vulnerability and Group 7–very high vulnerability.

The social vulnerability indicators of the clusters found for confirmed cases of dengue and the rest of the municipality are shown in Figure 6. When analyzing Figure 6, it is possible to notice that the census sector with the highest RR of confirmed dengue cases (Figure 5(a)) and blocking activities (Figure 5(b)) and house-house (Figure 5(c)) are found in a region with high vulnerability (Group 5) and the second cluster, confirmed cases of dengue and the activities of endemic agents, with the highest RR is located in a region with very low vulnerability (Group 1). Another point to note is that the census sector with the highest RR of confirmed dengue cases and of blocking and house-to-house activities was positioned in the Q1 (high-high) quadrant of the Moran scatterplot based on the LISA. It is also worth mentioning that according to data from SEADE, which used criteria to formulate the IPVS, the municipality under study did not present areas of very high vulnerability (IPVS 7).

As a limitation of this research, it is highlighted that in ecological studies, the identified results cannot be interpreted at the individual level. Undoubtedly, the spatial scan statistics contributed to exposing the dengue scenario in São Carlos and the presence of geographic areas of the municipality that are more susceptible to illness and need specific actions to control the disease.

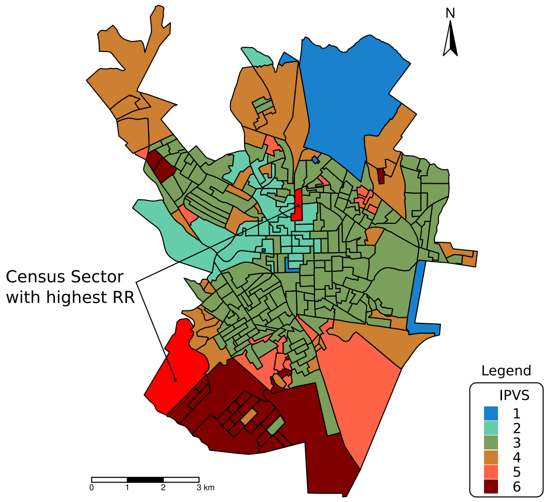


Figure 6. Spatial distribution and description of groups of the São Carlos Social Vulnerability Index (IPVS) of spatial clusters according to census sectors in the city of São Carlos-SP.

4. Conclusions

With the use of the spatial scanning technique, it was possible to more accurately identify high-risk areas for contracting the dengue virus. It was a tool to help epidemiological surveillance priority areas for intervention to combat dengue.

Furthermore, looking at the clusters in more detail can lead to developing community education and awareness programs to improve dengue prevention. However, further studies are needed to better validate spatial *clusters* as a dengue surveillance method. As future work, it is intended to carry out new studies with analytical designs capable of verifying, with more excellent reliability, possible associations between the areas of most significant risk for illness by dengue and Spatio-temporal analyses.

Acknowledgments

The authors would like to thank: CAPES (PROEX), CNPQ (grant 312605/2018-8), C4AI, FAPESP (grant 2020/16578-5), and the support of José Galizia Tundisi from the Science, Technology and Economic Development Department from São Carlos Municipality.

References

Alves, L. S., Dos Santos, D. T., Arcoverde, M. A. M., Berra, T. Z., Arroyo, L. H., Ramos, A. C. V., De Assis, I. S., De Queiroz, A. A. R., Alonso, J. B., Alves, J. D., et al. (2019). Detection of risk clusters for deaths due to tuberculosis specifically in areas of southern brazil where the disease was supposedly a non-problem. *BMC infectious diseases*, 19(1):1–13.

Andrade, A. L., Monteiro, A. M. V., Barcellos, C., Lisboa, E., Acosta, L. M. W., Almeida, M. C. d. M., Brito, M. R. V., Carvalho, M. S., Santos, M. A. d., Cruz, O., et al. (2007). Introdução à estatística espacial para a saúde pública.

Arroyo, L. H., Yamamura, M., Protti-Zanatta, S. T., Fusco, A. P. B., Palha, P. F., Ramos, A. C. V., Uchoa, S. A., and Arcêncio, R. A. (2017). Identificação de áreas de risco para a transmissão da tuberculose no município de são carlos, são paulo, 2008 a 2013. *Epidemiologia e Serviços de Saúde*, 26:525–534.

Câmara, G., Carvalho, M. S., Cruz, O. G., and Correa, V. (2004). Análise espacial de áreas. *Análise espacial de dados geográficos*, 2.

Canal, M. R., Ferreira, E. R. d. S., Estofolete, C. F., Dias, A. M., Tukasan, C., Bertoque, A. C., Muniz, V. D., Nogueira, M. L., and Silva, N. S. d. (2018). Spatiotemporal-based clusters as a method for dengue surveillance. *Revista Panamericana de Salud Pública*, 41:e162.

Da Silva, N. S., da Silva-Nunes, M., Malafronte, R. S., Menezes, M. J., D'Arcadia, R. R., Komatsu, N. T., Scopel, K. K., Braga, É. M., Cavasini, C. E., Cordeiro, J. A., et al. (2010). Epidemiology and control of frontier malaria in brazil: lessons from community-based studies in rural amazonia. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 104(5):343–350.

de Melo, D. P. O., Scherrer, L. R., and Eiras, Á. E. (2012). Dengue fever occurrence and vector detection by larval survey, ovitrap and mosquitrap: a space-time clusters analysis. *PloS one*, 7(7):e42125.

do Carmo, R. F., Júnior, J. V. J. S., Pastor, A. F., and de Souza, C. D. F. (2020). Spatiotemporal dynamics, risk areas and social determinants of dengue in northeastern brazil, 2014–2017: an ecological study. *Infectious Diseases of Poverty*, 9(1):1–16.

Estallo, E. L., Carbajo, A. E., Grech, M. G., Frías-Céspedes, M., López, L., Lanfri, M., Ludueña-Almeida, F. F., and Almiron, W. R. (2014). Spatio-temporal dynamics of dengue 2009 outbreak in córdoba city, argentina. *Acta tropica*, 136:129–136.

IBGE (2022). Instituto brasileiro de geografia e estatística cidades - são calos. https://cidades.ibge.gov.br/brasil/sp/sao-carlos/panorama.

Kulldorff, M. and Nagarwalla, N. (1995). Spatial disease clusters: detection and inference. *Statistics in medicine*, 14(8):799–810.

Li, H., Calder, C. A., and Cressie, N. (2007). Beyond moran's i: testing for spatial dependence based on the spatial autoregressive model. *Geographical analysis*, 39(4):357–375.

Monteiro, A. M. V., Câmara, G., Carvalho, M., and Druck, S. (2004). Análise espacial de dados geográficos. *Brasília: Embrapa*.

Moran, P. A. (1950). Notes on continuous stochastic phenomena. *Biometrika*, 37(1/2):17–23.

SEADE (2010). Fundação sistema estadual de análise de dados. Índice paulista de vulnerabilidade social.