Research, Society and Development, v. 12, n. 2, e0212239844, 2023
(CC BY 4.0) | ISSN 2525-3409 | DOI: http://dx.doi.org/10.33448/rsd-v12i2.39844

# Proposal of a framework for improving multi-criteria decision-making related to epidemics using heterogeneous spatial data and evolutionary algorithms

**Proposta de um framework para melhorar a tomada de decisão multicritério relacionada a epidemias usando dados espaciais heterogêneos e algoritmos evolutivos**

**Propuesta de un framework para mejorar la toma de decisiones multicriterio relacionadas con epidemias utilizando datos espaciales heterogéneos y algoritmos evolutivos**

Received: 12/30/2022 | Revised: 01/09/2023 | Accepted: 01/10/2023 | Published: 01/13/2023

**Gesiel Rios Lopes**
ORCID: https://orcid.org/0000-0002-4339-2417
University of São Paulo, Brazil
E-mail: gesielrios@usp.br
**Roberto Fray da Silva**
ORCID: https://orcid.org/0000-0002-9792-0553
University of São Paulo, Brazil
E-mail: roberto.fray.silva@gmail.com
**Karina Jorge Pelarigo**
ORCID: https://orcid.org/0000-0003-2446-7687
Federal University of São Carlos, Brazil
E-mail: karoufscar@hotmail.com
**Mellina Yamamura**
ORCID: https://orcid.org/0000-0001-5228-8788
Federal University of São Carlos, Brazil
E-mail: mellina@ufscar.br
**Alexandre C. B. Delbem**
ORCID: https://orcid.org/0000-0003-1810-1742
University of São Paulo, Brazil
E-mail: acbd@icmc.usp.br
**Denise Scatolini**
ORCID: https://orcid.org/0000-0003-1729-7966
São Carlos Municipality, Brazil
E-mail: denise.scatolini@saocarlos.sp.gov.br
**Filippo Ghiglieno**
ORCID: https://orcid.org/0000-0002-7795-6858
Federal University of São Carlos, Brazil
E-mail: filippo.ghiglieno@df.ufscar.br
**Antonio Mauro Saraiva**
ORCID: https://orcid.org/0000-0003-2283-1123
University of São Paulo, Brazil
E-mail: saraiva@usp.br

**Abstract**
The decision-making of complex problems, such as epidemics monitoring and control, involves multiple heterogeneous data and spatial and temporal aspects. Most problems cannot be reduced to one objective, characterized as multi-criteria decision-making (MCDM) problems. Adding temporal and spatial aspects further increases the complexity of addressing those problems. This paper proposed a framework that uses evolutionary algorithms and map algebra for addressing spatial and temporal multidimensional complex problems. It was evaluated in a case study of dengue and tuberculosis diseases in an urban environment, considering multi-resolution data and a genetic algorithm. Several analyses were conducted, generating maps and information essential to generate insights into the problem and a better understanding of the spatial relations between the variables. The framework and the code implemented could be applied to different problems, spatial resolutions, and objectives.
**Keywords:** Decision-making; Data fusion; Spatio-Temporal data analysis; Evolutionary algorithm; Map visualization.

**Resumo**
A tomada de decisão de problemas complexos, como monitoramento e controle de epidemias, envolve múltiplos dados heterogêneos e aspectos espaciais e temporais. A maioria dos problemas não pode ser reduzida a um objetivo, caracterizados como problemas de tomada de decisão multicritério (MCDM). Adicionar aspectos temporais e espaciais aumenta ainda mais a complexidade de lidar com esses problemas. Este artigo propôs uma estrutura que usa algoritmos

evolutivos e álgebra de mapas para resolver problemas complexos multidimensionais espaciais e temporais. Foi avaliado em um estudo de caso das doenças dengue e tuberculose em ambiente urbano, considerando dados de multi-resolução e um algoritmo genético. Diversas análises foram realizadas, gerando mapas e informações essenciais para gerar insights sobre o problema e entender melhor as relações espaciais entre as variáveis. A estrutura e o código implementados podem ser aplicados a diferentes problemas, resoluções espaciais e objetivos.
**Palavras-chave:** Tomada de decisão; Integração de múltiplas fontes; Análise de dados espaço-temporais; Algoritmo evolutivo; Visualização de mapas.

**Resumen**
La toma de decisiones de problemas complejos, como el monitoreo y control de epidemias, involucra múltiples datos heterogéneos y aspectos espaciales y temporales. La mayoría de los problemas no se pueden reducir a un objetivo, caracterizado como problemas de toma de decisiones de criterios múltiples (MCDM). Agregar aspectos temporales y espaciales aumenta aún más la complejidad de abordar esos problemas. Este documento propuso un marco que utiliza algoritmos evolutivos y álgebra de mapas para abordar problemas complejos multidimensionales espaciales y temporales. Se evaluó en un estudio de caso de enfermedades de dengue y tuberculosis en un medio urbano, considerando datos multiresolución y un algoritmo genético. Se realizaron varios análisis, generando mapas e información esencial para generar conocimientos sobre el problema y comprender mejor las relaciones espaciales entre las variables. El marco y el código implementado podrían aplicarse a diferentes problemas, resoluciones espaciales y objetivos.
**Palabras clave:** Toma de decisiones; Integración de múltiples fuentes; Análisis de datos espacio-temporales; Algoritmo evolutivo; Visualización de mapas.

## 1. Introduction

Decision-making for complex problems based on heterogeneous and multiple data sources requires structuring information with adequate representation of the phenomenon under analysis. Dimensions such as temporal and spatial add complexity to data processing, information extraction, and interpretation of results (Moraes et al., 2014; Lopes et al., 2022, November).

Many works in the literature address the use of multi-criteria decision-making methods (MCDM), such as the Analytical Hierarchy Process (AHP), to overcome this challenge. In Ferreira and Silva (2020) the authors use the Analytical Hierarchy Process (AHP) method with Map Algebra to determine the environmental fragility of the Rio Brilhante watershed in Mato Grosso do Sul, Brazil, from the superposition of natural and anthropic factors. The authors use the AHP method to define a set of weights with the level of importance among the variables with the most significant influence on the definition of the environmental fragility of the basin from the paired analysis by an expert.

In Niño et al. (2020) the authors propose a multi-criteria evaluation model, integrated with a Geographic Information System, to analyze the risk of transmission of SARS-CoV-2 in the urban area of the municipality of Villavicencio, Colombia, from the use of descriptive attributes of threats and vulnerabilities of viral transmission.

This work presents a systematic approach to a cartographic representation derived from the composition of multiple thematic maps through an Evolutionary Algorithm (EA) to support decision-makers in defining mitigation strategies at the local level, facilitating the location and optimization of resources for decision-makers in complex, multidimensional scenarios.

The main research question it addresses is: “How can EA be used to improve decision-making in complex, multidimensional scenarios with spatial aspects?”. To answer this question, a framework will be proposed and implemented in a case study of dengue disease prediction in an urban environment, considering data on different spatial resolutions and types.

The rest of the paper is organized as follows. Section 2 presents works related to the use of the MCDM in Spatial Analysis. Section 3 introduces essential concepts of the method. A case study to apply the methods is presented in Section 4. The results and discussions about the method are described in Section 5. Finally, Section 6 points out conclusions and future research lines.

## 2. Related Works

The application of MCDM methods for complex geospatial scientific investigations, as in the context of epidemiological surveillance, has been widely used in the literature (Malczewski & Rinner, 2015; Ferreira & Silva, 2020; Niño et al., 2020; Lopes et al., 2022, November).

Chabuk et al., (2017) use the AHP method to define the hierarchical importance of physical and environmental factors that contribute to the potential environmental fragility assessment of the river basin of the Jequitinhonha River Basin, Minas Gerais, Brazil. Another work with a similar proposal is (Hongoh et al., 2011), where the authors propose an MCDA-based approach to developing geospatial models and spatially explicit decision support tools for managing vector-borne diseases.

Vanolya et al., 2019 use citizen-generated information as baseline data to validate the results of the MCDA and Geographic Information Systems (GIS) processes. The validation proposed by the authors is performed using specific spatial indices, including total coverage, geometric intersection, central distances and statistical indices. Ur Rahman (2022) proposes investigating data visualization techniques for disease mapping, to create awareness about the disease for the guidance of patients, health professionals, and government agencies.

Covre et al. (2022) carry out an epidemiological study, of the ecological type, with data from the Epidemiological Report provided by the State Department of Health of Paraná on confirmed cases of covid-19, in the period from March 12, 2020, to January 18, 2021, where they apply spatial analysis in the distribution of intensive care beds to identify priority critical and transition areas related to the spread and control of the disease. Yalew et al., 2016 use the AHP method to define the degree of importance of multiple global data from different sources on agriculture through the Google platform in a tool called AgriSuit framework. Another work that uses the AHP method as an MCDA technique to aid decision-making in spatial data is (Duan et al., 2022), where the authors use the AHP method to calculate the weight and generate a map of results of hazards, exposure, vulnerability, and emergency responses and recovery capacity in areas at risk of urban flooding.

Thus, it is observed that works use subjective MCDM methods (Odu, 2019), such as AHP, to determine criteria weights according to the priority of criteria, which can make the solution depends on the expert's vision. In this paper, we propose a systematic approach, independent of an expert's view, to define data influence through an EA and map algebra to combine these data and build thematic maps from multiple sources to support decision-makers in defining mitigation strategies.

## 3. Materials and Methods

The use of hierarchical theme approaches to define the degree of importance of each thematic layer in the weighted composition of these layers is recurrent; however, these approaches are subjective and highly centered on the judgment of who defines the paired comparison of the essential criteria, making the solution skewed (Malczewski & Rinner, 2015; Lopes et al., 2022, November).

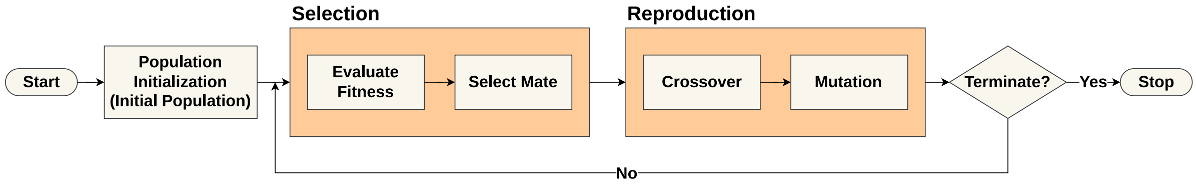
Finding the user's balance and the optimal relationship between the thematic layers from a spatial point of view is an excellent and computationally complex challenge (Pfeiffer et al., 2008; Malczewski & Rinner, 2015). In order to mitigate it, we propose a systematic approach to building a cartographic representation derived from the composition of multiple thematic maps through a genetic algorithm to support decision-makers in defining different strategies at the local level in the area under study, facilitating the location and optimization of resources.

Genetic Algorithms (GAs), introduced by John Holland (Holland, 1973), are among the most successful metaheuristics in solving optimization problems. Inspired by the evolution of species, GAs perform search procedures in the space of viable solutions, using probabilistic rules to combine solutions to obtain quality improvements. GAs work by maintaining a population, called individuals or chromosomes, operating on them similarly to the evolution of species. The so-called genetic operators are

Research, Society and Development, v. 12, n. 2, e0212239844, 2023
(CC BY 4.0) | ISSN 2525-3409 | DOI: http://dx.doi.org/10.33448/rsd-v12i2.39844

applied to these populations, such as crossover and mutation, to simulate the survival of the fittest, making the evolution of the species (Eiben & Smith, 2003; Goldbarg et al., 2017). Figure 1 shows a flowchart that describes how the GAs work.

**Figure 1 -** Flowchart with the operation of the GAs.




Source: adapted from Goldbarg et al., (2017).

GAs start from the following principle: given a population of individuals (i.e., a set of solutions), environmental pressures trigger a process of natural selection, that is, a process that generally favors the best solutions found so far, which causes an increase in the adequacy rate of the solutions. Given a function to be optimized, a set of solutions is randomly generated that are elements belonging to the domain of an objective function that, in turn, must be used to evaluate the quality of the generated solutions. The value resulting from the adequacy assessment is called fitness.

The proposed thematic map $S$ to be modeled from GA consists of the $i$-th thematic maps weighted by their respective weighting factors, according to Equation 1.

$$S = \sum_{i=1}^{n} v_i \, w_i \tag{1}$$

where $v_i$ is the $i$-th thematic map and $w_i$ is the $i$-th weighting factor of the importance degree of the $i$-th thematic map, with $\sum w_i = 1$.

The proposed GA uses Equation 1 as fitness and aims to find the best combination of weights in the combination of thematic layers that optimizes the relations from a spatial point of view between them, since one of the first steps when analyzing a phenomenon from a spatial point of view is to test the hypothesis that the spatial data are randomly distributed, that is, if the values of the attribute under study in a given region do not depend on the values of the attributes in the neighboring regions (Almeida, 2012).

Moran's global index (Moran's I) is used as an evaluation criterion for the weighted composition generated by the GA since it is the expression of the autocorrelation, considering only the first neighborhood level, whose null hypothesis is the independence of the data (Moran, 1950; Li et al., 2007). Global indicators such as Moran's I provide a single value as a measure of spatial association for the entire dataset, which helps characterize the region as a whole (Câmara et al., 2004).

The Moran's I is given by Equation 2.

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

where $n$ is the number of areas, $z_i$ is the value of the attribute considered in the area $i$, $\bar{z}$ is the average of the assigned values in the study region, and $w_{ij}$ the elements from the normalized spatial proximity matrix (Moran, 1950; Câmara et al., 2004; Li et al., 2007).

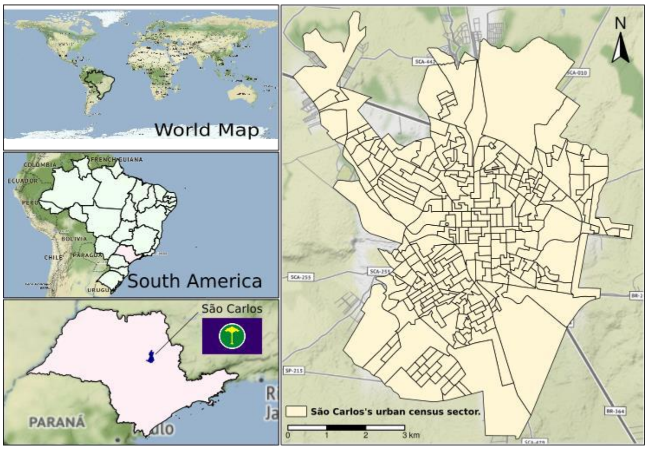
## 4. Case Study: Dengue and Tuberculosis in an Urban Environment

Transmission of infectious diseases is closely linked to spatial and spatio-temporal proximity, as the transmission is more likely if individuals at risk are close in a spatial and temporal sense. Understanding, mapping, and early identification of the risks and difficulties of coping with diseases considering socio-environmental factors is highly complex and essential for the effective implementation of interventions aimed at reducing the incidence of human mortality in the population of any region (Pfeiffer et al., 2008).

The proposed GA is applied in the construction of a cartographic representation derived from population information, social factors, and multiple epidemics to support epidemiological analyzes and, consequently, the behavior of decision-makers, defining mitigation strategies at the local level, facilitating the optimization of resources.

In this case study, the census sectors of the urban area of the municipality of São Carlos (Figure 2) were used, a medium-sized city in the interior of the state of São Paulo, southeastern Brazil. Located in the east-central region of the state, 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. Concerning socioeconomic aspects, the municipality had a Gini index of 0.63, a Human Development Index of 0.805, and a gross domestic product of R$ 6,712,498.00 in 2010.

**Figure 2 -** Location of the municipality and urban perimeter of São Carlos – SP.




Source: Authors.

A set of thematic maps necessary to model the mapping of critical regions about multiple epidemics were raised, depending on the availability of data and spatial bases from the opinion of a group of multidisciplinary experts, with the presence of three physicians, two epidemiologists, two nurses, three statisticians and two specialists in geoprocessing. Table 1 shows the set of variables used.

As described in Table 1, for this case study, the population database of the IBGE demographic census, carried out in 2010, was used, with aggregation at the level of census tracts. With the total population aggregated for the sectors, the demographic density in inhabitants per square kilometer was calculated according to the value of the area of each polygon corresponding to the sectors.

**Table 1** – Selected variables.

| Variable | Description of Variables |
|---|---|
| Demographic density | The population density per square kilometer was extracted from the IBGE population census database, carried out in 2010, with aggregation at the level of census sectors. |
| Average Residents per Household | Also extracted from the IBGE 2010 census database. This variable considers the total population of the sector by the respective number of households. |
| Percentage of population aged over 60 years | From the census data, the cumulative population over 60 years was determined. Then the percentage value was calculated in relation to the total population for each sector. |
| Paulista Social Vulnerability Index | The São Paulo Social Vulnerability Index (IPVS) is an index that was conceived by the SEADE foundation of the state of SP and implemented from the set of existing information in the IBGE 2010 Demographic Census database, which consists of socioeconomic information and demographics investigated by the Census, and aggregated at the level of census sectors |
| Dengue Case Count | Reported and confirmed dengue cases from the SINAN-Dengue Notifiable Diseases Information System of residents of the urban area of the municipality from January 1 to December 31, 2019, aggregated by census sectors. |
| Tuberculosis case count | All tuberculosis cases registered in the tuberculosis patient control system of the state of São Paulo (TB-WEB) from January 1 to December 31, 2019, residents of the municipality, aggregated by census sectors. |

Source: Authors.

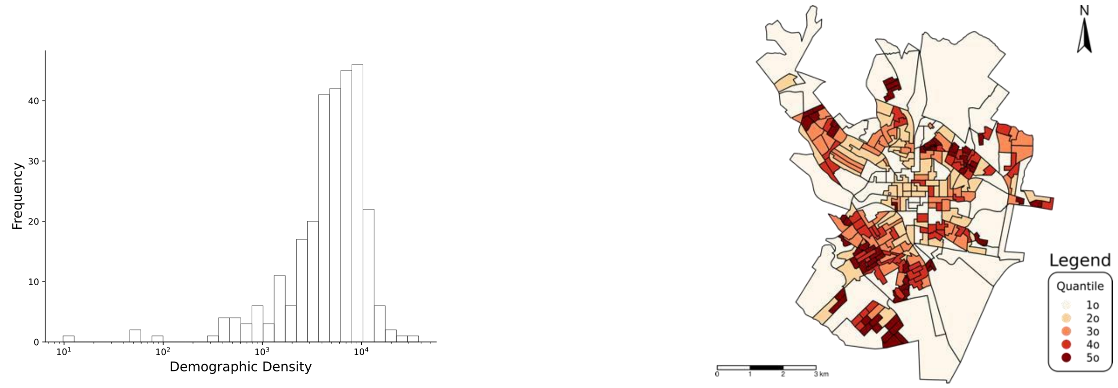
Since we are dealing with almost three hundred sectors, and due to the amplitude of these results, we need, in order to be able to rank them and assign scores to them to represent them in a classification of the variable by intervals. There are different data classification methods described in the literature. Matsumoto et al., (2017) describe several of these methods, highlighting the importance of choosing an adequate method for the desired faithful representation. In this study, we used the quantile method to classify data to assign scores to them, whose main characteristic is to form classes with an approximate number of assigned features. Demographic density values were then sliced into five groups according to their quintiles.

Figure 3(a) shows the histogram of demographic density data, and Figure 3(b) shows the distribution of demographic density in sectors according to these groups. Analogously to demographic density, it is to be expected that within the household, we also have a higher probability of contagion with a more significant number of residents. This is also in line with the thought brought to the discussion by health professionals that the number of people sleeping in the same bedroom should be considered. In this case, the IBGE 2010 census database was also used. This data considers the sector's total population and the respective household count. The same slicing procedure used for the previous variable was adopted for this variable.

Figure 4(a) shows the histogram of data on the average number of residents per household, and Figure 4(b) shows the distribution of the average number of residents per household in the sectors according to these groups. Each group was assigned a score ranging from 1 to 5 points according to the increase in the average number of residents per household.

From the point of view of the criticality of the epidemic, it is known that age is one of the most relevant risk factors. In order to contemplate this characteristic in the modeling, from the census data, the accumulated population aged over 60 years was accounted for, and the percentage value was calculated in relation to the total for each sector. The same slicing procedure used for both previous variables was adopted in this case.

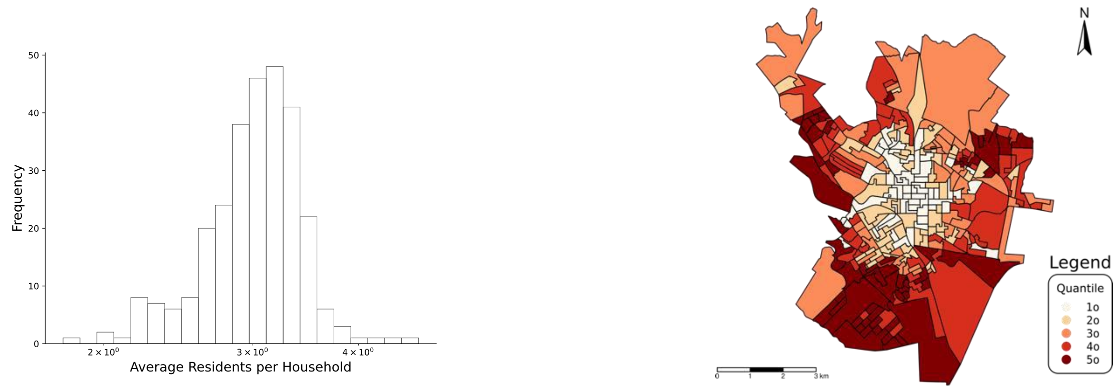
**Figure 3 -** (a) Histogram of demographic density data and (b) Choropleth representation of population density distribution.




(a) (b)

Source: Authors.

**Figure 4 -** (a) Histogram of data on the average number of residents per household and (b) Choroplectic representation of the distribution of the average number of residents per household.



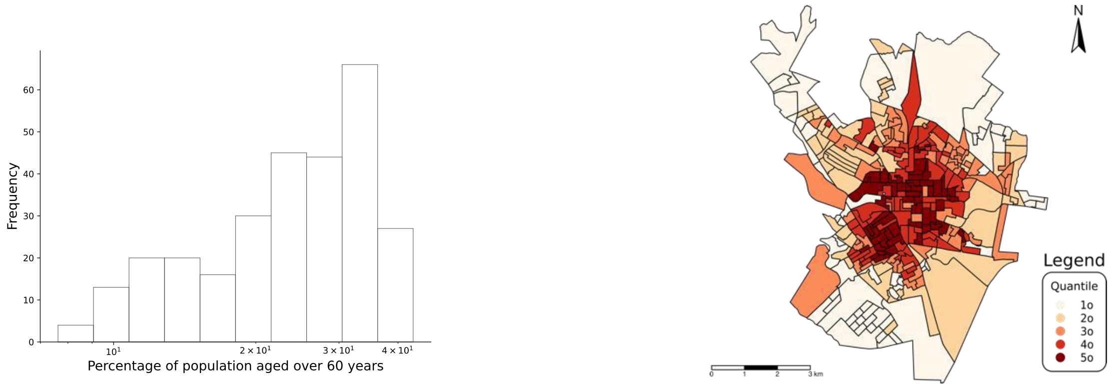

(a) (b)

Source: Authors.

Figure 5(a) shows the histogram of data on the percentage of the population over 60, and Figure 5(b) shows it is possible to visualize the distribution of the population over 60 years in the sectors according to these groups. Each group was assigned a score ranging from 1 to 5 points according to the increase in the percentage of the population aged over 60 years. The previous existence of cases in a particular region puts it under an even more severe risk condition. The greater the number of these cases, the worse the condition of the region.

For the dengue cases was considered all cases reported and confirmed from the SINAN-Dengue Notifiable Diseases Information System of residents of the urban area of the municipality from January 1 to December 31, 2019 and all tuberculosis cases registered in the tuberculosis patient control system of the state of São Paulo (TB-WEB) from January 1 to December 31, 2019, residents of the municipality. With the data on the epidemics in hand, a step was first taken to clean and anonymize these data, removing duplicate data, inconsistent dates, and unwanted and invalid information, in addition to any data related to the patient, such as name, telephone number, date of birth, for example. After this stage, the cases of dengue and tuberculosis were georeferenced using the Google Maps Geocoding API to obtain the respective geographic coordinates referring to the notified residential addresses. Once the georeferenced notifications are made using the geolocalized data with the mesh of the urban census sector in the municipality of São Carlos through a spatial join function, removing geolocalized data outside the area under study, this step is called Geolocation Data Cleaning. Figure 6 shows the workflow used to perform this geocoding process.

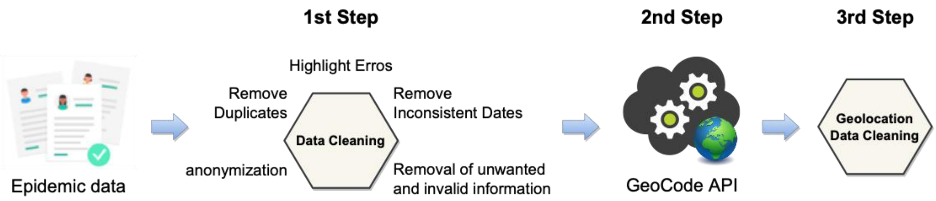
**Figure 5 -** (a) Histogram of data on the percentage of the population aged over 60 years and (b) Choropleth representation of the distribution of the percentage of the population aged over 60 years.




(a) (b)

Source: Authors.

**Figure 6 -** Workflow used to perform this geocoding process.



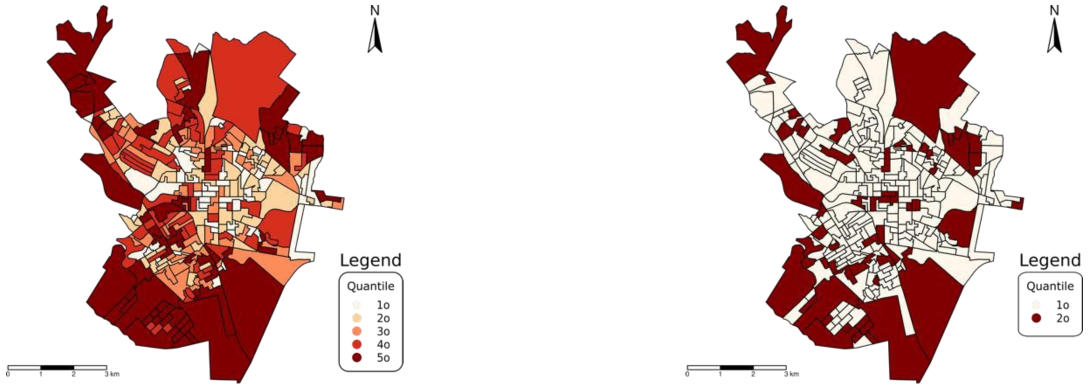

Source: Authors.

Figure 7 show the choropleth representation of the georeferenced cases of Dengue and Tuberculosis for São Carlos-SP in 2019.

To describe the social vulnerability was used data from the "Fundação Sistema Estadual de Análise de Dados (SEADE)", referring to the São Paulo Social Vulnerability Index (IPVS) for 2010. This index classifies the census sectors based on a combination of demographic and socioeconomic dimensions.It identifies 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 (Alves, 2020).

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 (Alves, 2020), 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. Figure 8 shows the choropleth representation of the IPVS distribution for São Carlos-SP.

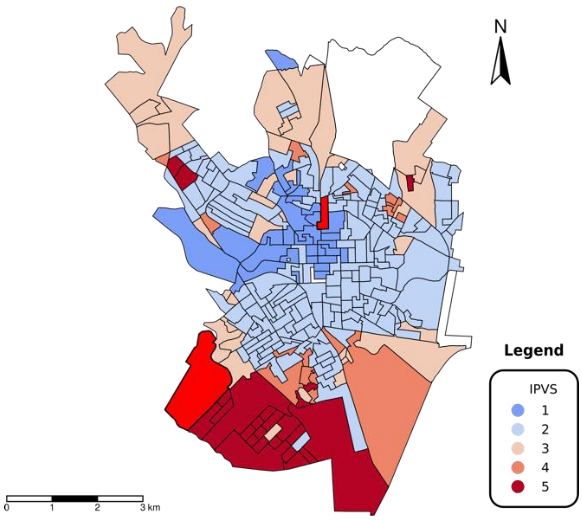
**Figure 7 -** (a) Choropleth representation of the spatial distribution of confirmed Dengue cases in 2019 and (b) Choropleth representation of the spatial distribution of Tuberculosis cases in 2019.




(a) (b)

Source: Authors.

**Figure 8 -** Choropleth representation of the IPVS distribution for São Carlos-SP.




Source: Authors.

## 5. Results and Discussion

In order to determine the relative contribution of each thematic map and thus validate the proposed approach, it was decided to classify the thematic maps into three groups according to demographic, social, and epidemiological aspects. With the definition of the weight of each group of variables, a global score is calculated by crossing the different thematic layers weighted by their respective weights using a weighted linear combination, which can be expressed by Equation 3.

$$Global\ Score = w_1 \times demographic\ score + w_2 \times social\ score + w_3 \times epidemiological\ score \tag{3}$$

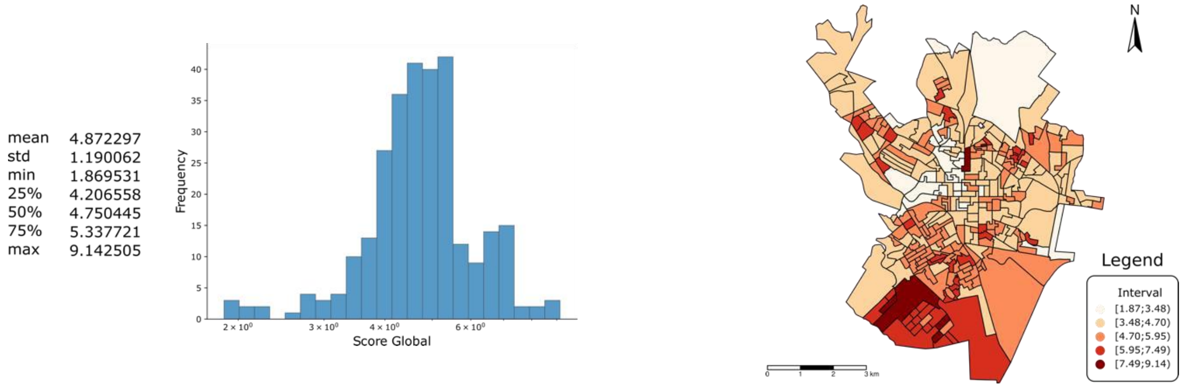
From the calculated global scores data, the first analysis carried out was to verify its descriptive statistics. These are shown in Figure 9(a).

The global scores values were then sliced into five classes according to their natural break values. In this case, the classification method adopted has the characteristic of similar grouping values and maximizing differences between classes, with established limits where there are considerable differences between data values. Thus, this method represents the natural scaling of the data series, grouping them according to (Matsumoto et al., 2017) similarity. Figure 9(b) presents the classification of census sector in relation to the model's global scores.

To ensure that the modeling represents the phenomenon from a spatial point of view and thus validates the mapping in question, one of the necessary aspects that we have to answer is whether the event under study and the factors related to it have a spatially conditioned distribution.

In this case, we have to use spatial statistics, in particular, the study of spatial dependence, to demonstrate how values are correlated with space, that is, if and how they depend on values of the same variable in neighboring regions.

**Figure 9 -** (a) Descriptive Statistics of the Global Score and (b) Choropleth representation of the global scores.




(a) (b)

Source: Authors.

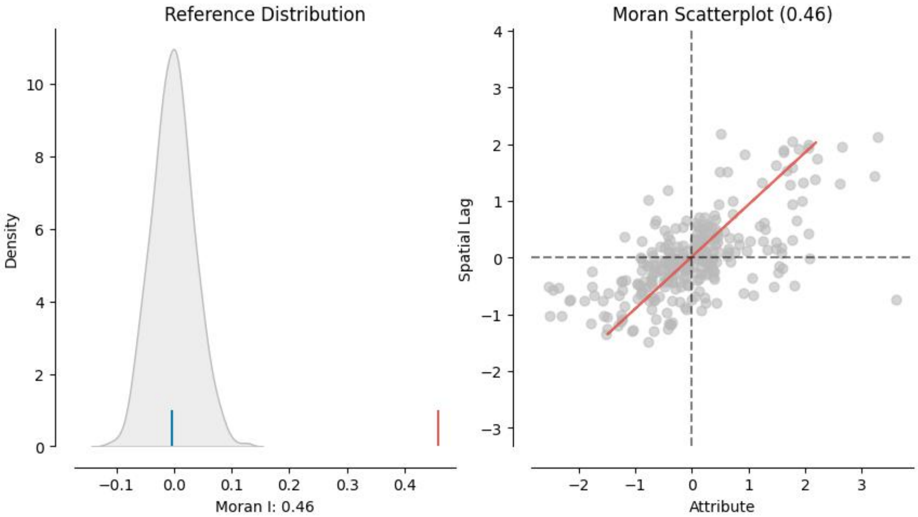
Moran's global index is the autocorrelation expression considering only the first neighborhood level and works as a test whose null hypothesis is the independence of the data (Moran, 1950; Câmara et al., 2004; Li et al., 2007; Lopes).

The global score values of each sector were used to calculate Moran's I. Figure 10 illustrates the test result. Given the Moran's I of 0.46, there is less than a 1% probability that this clustering pattern could be a random result. Thus, the spatial distribution of high and/or low values in the data set is more spatially clustered than expected if the underlying spatial processes were random.

Global indicators such as Moran I provide a single value as a measure of spatial association for the entire dataset, which is helpful in characterizing the region. Thus, we can reject the null hypothesis, but we still have no way of significantly analyzing the pattern of the clusters (Moran, 1950; Câmara et al., 2004; Li et al., 2007).

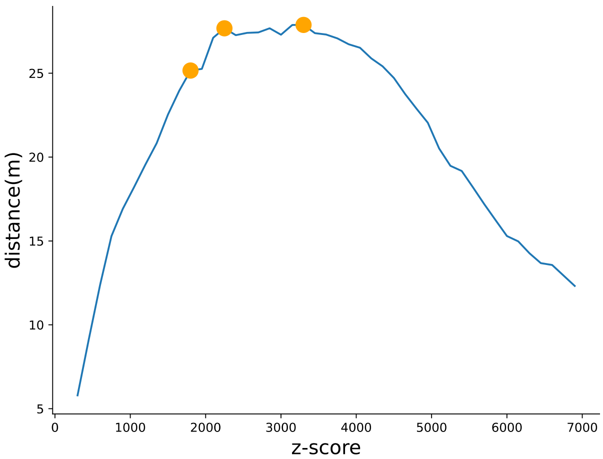
Using the centroids of the census sector as a starting point, we can calculate the neighborhood distance to other centroids to calculate Moran I and thus analyze how spatial dependence changes as a function of this distance. For this, we analyzed the distances between the centroids pair by pair. We found that the minimum distance found was around 130 meters, and the maximum was around 1600 meters, with an average distance of around 300 meters. In this way, Moran's I was calculated for distances every 150 meters starting from 300 meters in 30 intervals. In Figure 11, we present the graph of z-score variation as a function of distance. The peaks reflect the distances at which the spatial processes that promote clustering are most pronounced. The color of each point on the graph corresponds to the statistical significance of the z-score values.

**Figure 10 -** Moran's I test result.




Source: Authors.

**Figure 11 -** Special autocorrelation as a function of distance.



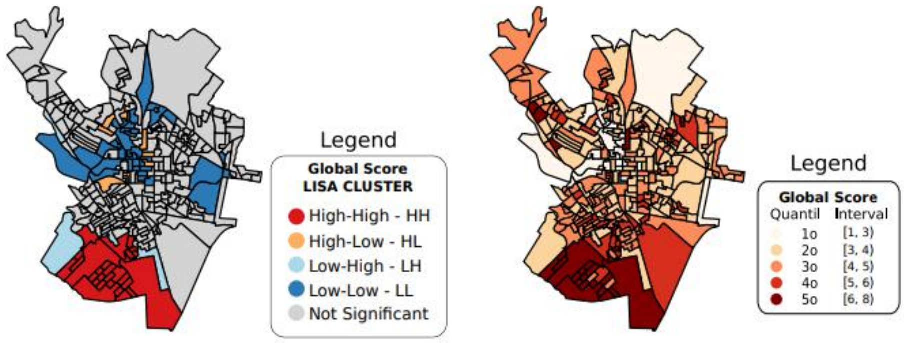

Source: Authors.

We aim to identify sharp limits by identifying high-value and low-value concentrations, and spatial discrepancies for the global score values. For this, the analysis by local identifiers produces a specific value for each area, thus allowing the identification of these groupings. We applied Moran's local identifier presented by (Anselin, 1995). For the calculation of the $p$ -value, 999 permutations were adopted, with a precision of 0.001 for the significance level. Figure 9 compares the local and global Moran's I results.

## 6. Conclusion

This paper explored the problem of improving MCDM of complex spatial problems. A framework that uses evolutionary algorithms and considers multiple heterogeneous variables with spatial correlation was proposed. It was explored in a case study of dengue and tuberculosis diseases in an urban environment, considering multiple-resolution data and a genetic algorithm. The solution generated by the proposed GA from multiple data sources helps decision-making, generating insights into the problem and the spatial relationships between the variables. It is vital to note that the same framework and code can be applied to different problems, spatial resolutions, and criteria.

**Figure 12 -** Comparison between results using Global and Local Moran's I.




Source: Authors.

## Acknowledgments

The authors of this work would like to thank the Center for Artificial Intelligence (C4A-IUSP) and the support from the São Paulo Research Foundation (FAPESP grant #2019/07665-4, #2020/16578-5) and from the IBM Corporation. We would like also to thank the CEPID-CeMEAI/ICMC-USP (CEPID, FAPESP grant #2013/07375-0)); the support of José Galizia Tundisi from the Science, Technology and Economic Development Department from São Carlos Municipality.

## References

Almeida, E. (2012). Econometria espacial. Alínea.

Alves, J. D. G. (2020). O Índice Paulista de Vulnerabilidade Social (IPVS) como ferramenta para promoção políticas públicas: aplicação do índice no município de Piracicaba–São Paulo. Anais, 1-6.

Anselin, L. (1995). Local indicators of spatial association—LISA. Geographical analysis, 27(2), 93-115.

Câmara, G., Monteiro, A. M., Fucks, S. D., & Carvalho, M. S. (2004). Análise espacial e geoprocessamento. Análise espacial de dados geográficos. Brasília: EMBRAPA, 21-54.

Chabuk, A., Al-Ansari, N., Hussain, H. M., Knutsson, S., Pusch, R., & Laue, J. (2017). Combining GIS applications and method of multi-criteria decision-making (AHP) for landfill siting in Al-Hashimiyah Qadhaa, Babylon, Iraq. Sustainability, 9(11), 1932.

Covre, E. R., Pereira, N. D., Oliveira, N. N. D., Charlo, P. B., Oliveira, M. L. F. D., Oliveira, R. R. D., ... & Salci, M. A. (2022). Correlação espacial da covid-19 com leitos de unidades de terapia intensiva no Paraná. Revista de Saúde Pública, 56.

Duan, C., Zhang, J., Chen, Y., Lang, Q., Zhang, Y., Wu, C., & Zhang, Z. (2022). Comprehensive Risk Assessment of Urban Waterlogging Disaster Based on MCDA-GIS Integration: The Case Study of Changchun, China. Remote Sensing, 14(13), 3101.

Eiben, A. E., & Smith, J. E. (2003). Introduction to evolutionary computing (Vol. 53, p. 18). Berlin: springer.

Ferreira, P. S., & Silva, C. A. D. (2020). O método AHP e a Álgebra de Mapas para determinar a fragilidade ambiental da bacia hidrográfica do Rio Brilhante (Mato Grosso do Sul/Brasil), proposições para a gestão do território. Confins. Revue franco-brésilienne de géographie/Revista franco-brasilera de geografia, (46).

Goldbarg, E., Goldbarg, M., & Luna, H. (2017). Otimização combinatória e metaheurísticas: algoritmos e apliacações. Elsevier Brasil.

Holland, J. H. (1973). Genetic algorithms and the optimal allocation of trials. SIAM journal on computing, 2(2), 88-105.

Hongoh, V., Hoen, A. G., Aenishaenslin, C., Waaub, J. P., Bélanger, D., & Michel, P. (2011). Spatially explicit multi-criteria decision analysis for managing vector-borne diseases. International Journal of Health Geographics, 10(1), 1-9.

Li, H., Calder, C. A., & Cressie, N. (2007). Beyond Moran's I: testing for spatial dependence based on the spatial autoregressive model. Geographical analysis, 39(4), 357-375.

Lopes, G. R., Delbem, A. C., da Silva, R. F., Júnior, C. B., de Mattos, S. H. V. L., Scatolini, D., ... & Saraiva, A. M. (2022, November). MultiMaps: a tool for decision-making support in the analyzes of multiple epidemics. In Proceedings of the 3rd ACM SIGSPATIAL International Workshop on Spatial Computing for Epidemiology (pp. 22-25).

Lopes, G. R., Pelarigo, K. J., Delbem, A. C., & de Sousa, J. B. (2022). Análise Exploratória de Dados Espaciais com Python. Sociedade Brasileira de Computação.

Malczewski, J., & Rinner, C. (2015). Multicriteria decision analysis in geographic information science (Vol. 1, pp. 55-77). New York: Springer.

Matsumoto, P. S. S., da Castro Catão, R., & Guimarães, R. B. (2017). Mentiras com mapas na Geografia da Saúde: métodos de classificação e o caso da base de dados de LVA do SINAN e do CVE. Hygeia: Revista Brasileira de Geografia Médica e da Saúde, 13(26), 211.

MORAES, R. M., NOGUEIRA, J. A., & SOUSA, A. C. (2014). A new architecture for a spatio-temporal decision support system for epidemiological purposes. In Decision Making and Soft Computing: Proceedings of the 11th International FLINS Conference (pp. 17-23).

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

Niño, L., Gutiérrez-Lesmes, O., Diaz-Celis, C. A., & Manrique-Abril, F. (2020). Riesgo de transmisión de SARS-CoV-2: evaluación espacial multicriterio en un municipio de Colombia, 2020. Revista de Salud Pública, 22(2).

Odu, G. O. (2019). Weighting methods for multi-criteria decision making technique. Journal of Applied Sciences and Environmental Management, 23(8), 1449-1457.

Pfeiffer, D. U., Robinson, T. P., Stevenson, M., Stevens, K. B., Rogers, D. J., & Clements, A. C. (2008). Spatial analysis in epidemiology. OUP Oxford.

ur Rahman, A. (2022). Geo-Spatial Disease Clustering for Public Health Decision Making. Informatica, 46(6).

Vanolya, N. M., Jelokhani-Niaraki, M., & Toomanian, A. (2019). Validation of spatial multicriteria decision analysis results using public participation GIS. Applied Geography, 112, 102061.

Yalew, S. G., Van Griensven, A., & van der Zaag, P. (2016). AgriSuit: A web-based GIS-MCDA framework for agricultural land suitability assessment. Computers and Electronics in Agriculture, 128, 1-8.