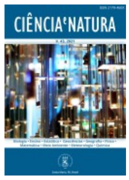
CIÊNCIA e NATURA



ISSN 2179-460X
Acesso aberto

Ci. e Nat., Santa Maria, v. 43, e42, 2021 • https://doi.org/10.5902/2179460X39332
Submissão: 16/07/2020 • Aprovação: 24/07/2020 • Publicação: 04/06/2021

Geography

Spatial autocorrelation proposal of the relationship between the socioeconomic conditions in Metropolitan Region of Sorocaba, SP, Brazil

Proposta de autocorrelação espacial da relação entre as condições socioeconômicas na Região Metropolitana de Sorocaba, SP, Brasil

Darllan Collins da Cunha e Silva[I], Vanessa Cezar Simonetti[II], Renan Angrizani de Oliveira[II], Jomil Costa Abreu Sales[II], Roberto Wagner Lourenço[I]

[I] Universidade Estadual Paulista (Unesp), Departamento de Engenharia Ambiental, Sorocaba, SP, Brasil
[II] Universidade Estadual Paulista (Unesp), Pós-Graduação em Ciências Ambientais, Sorocaba, SP, Brasil

ABSTRACT

The spatialization of social data allows to analyze some social and territorial characteristics of census tracts up to the totality of a city or metropolitan region. The objective of this study was to verify the spatial autocorrelation of data that reflect the health and income conditions of households in the Metropolitan Region of Sorocaba (MRS) and verify if there is a correlation of these indicators by a multiple linear regression test. For this, the Global and Local Moran Index was calculated, which were used to measure autocorrelation and spatial dependence among the census tracts. It was identified that there are 177 census tracts distributed by MRS that showed autocorrelations for all variables and correspond to 31.1% of the territory and 5.4% of the total population of MRS. This study can be used by public managers to develop public policies aimed at improving the quality of life of the population because allows the identification of the regions that go beyond the administrative limits of the municipalities that lack collective investment and cooperation of municipalities.
Keywords: Spatial correlation; Urbanization; Geoprocessing; Demographic census; Linear regression

RESUMO

A espacialização dos dados sociais permite analisar algumas características sociais e territoriais dos setores censitários até a totalidade de uma cidade ou região metropolitana. O objetivo deste estudo foi verificar a autocorrelação espacial dos dados que refletem as condições de saúde e renda dos domicílios da Região Metropolitana de Sorocaba (RMS) e se há correlação desses indicadores por meio de um teste de regressão linear múltipla. Para isso, foi calculado o Índice de Moran Global e Local, os quais foram utilizados para mensurar a autocorrelação e a dependência espacial entre os setores censitários. Identificou-se que existem 177 setores censitários distribuídos pela RMS que mostraram autocorrelações para todas as variáveis e correspondem a 31,1% do território e 5,4% da população total da RMS. Portanto, este estudo pode ser utilizado por gestores públicos para desenvolver políticas públicas voltadas à

melhoria da qualidade de vida da população, pois possibilita a identificação das regiões que extrapolam os limites administrativos dos municípios e que carecem de investimentos coletivos e cooperação dos municípios.
Palavras-chave: Correlação espacial; Urbanização; Geoprocessamento; Censo demográfico; Regressão linear

1 INTRODUCTION

Data spatialization allows the elaboration of thematic maps that are widely used in several areas to visualize the spatial distribution of an event, such as the percentage of people who have sanitary sewage collection in their residences, indicating possible areas of high occurrence or predominance of this event, needing an intervention or further investigation of the causes of this fact (TEIXEIRA; GUILERMINO, 2006; MONDINI; CHIARAVALLOTI-NETO, 2008; MEIRA et al., 2016).

In most cases, the data are not available with the exact position of the event, but by the number of occurrences for areas aggregated by geographically defined regions. Such aggregation data may occur for convenience or because it reflects the way which data was available. From the statistical view, these aggregates by area can present different data between them, implying that each area shows a distinct distribution probability from the others. The estimates obtained within aggregates area vary depending on the various ways that these units can be grouped and can obtain different results when changing the boundaries of these areas (ANSELIN, 1995; TEIXEIRA; GUILERMINO, 2006; ROSTAMI et al., 2017; SILVA et al., 2017).

The inclusion, application, and analysis of automated georeferenced database using geotechnologies enable to improve the spacial understanding, allowing the integration of socioeconomic, environmental and demographic information, in order to capture existing inequalities without dissociating the territorial space, especially when it comes to analyses of spatial autocorrelation of these data (RICHARDS et al., 1999; KOCH; DENIKE, 2004; SILVA et al., 2017).

The study of spatial correlations is complex and can be used in many different scenarios. Rostami et al. (2017) used mathematical modeling and spatial analysis in the investigation of space-time variations related to substance abuse in Iran, finding evidence of spatial heterogeneity as well as gender inequality associated with the deaths caused by the disorders related to the use of these substances. Hallisey et al. (2017) analyzed the distribution of cancer mortality in adolescents in Georgia (United States), testing different statistical methods, and concluded that the interpolation of census tract data and georeferenced information has significant accuracy and can be an important tool for health professionals and other researchers who often need to work with those data at different scales.

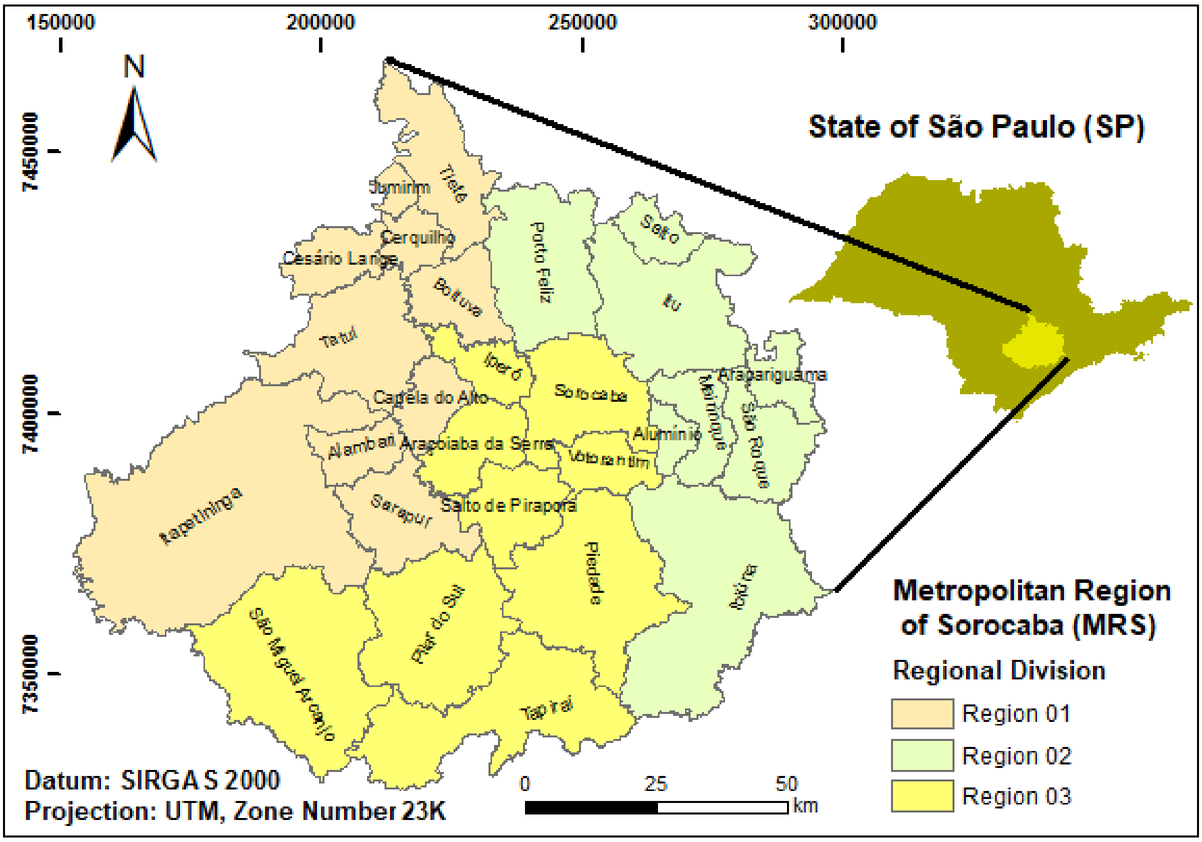
Thus, the objective of this study is to verify the spatial autocorrelation data which reflect the health and income conditions from residences located in the census tracts of the municipalities of the Metropolitan Region of Sorocaba (MRS) using mathematical modelling techniques with the aid of the Geographical Information Systems (GIS), and verify the correlation between these data by a multiple linear regression analysis.

2 MATERIAL AND METHODS

2.1 Study area

The Metropolitan Region of Sorocaba (MRS) located in São Paulo State in Brazil, has 27 municipalities (Figure 1), with a population of 2.1 million for the year 2018 and borders the São Paulo and Campinas metropolitan regions. It has an area of 11,611.48 km², which is equivalent to approximately 4.7% of the territory of the State of São Paulo and a demographic density of approximately 182.6 inhabitants / km². Twelve municipalities are located on the axis of the Castello Branco and/or Raposo Tavares highways, with economies based on industrial activities, of which six are of particular relevance in the São Paulo state economy: Sorocaba, Itu, Itapetininga, Salto, Cerquilho and Votorantim. Also, Sorocaba has the largest Gross Domestic Product (GDP) in the metropolitan region, being the 19th largest GDP in Brazil (EMPLASA, 2020; NERY et al., 2020).

Figure 1 – Location of the Metropolitan Region of Sorocaba (MRS) in the State of São Paulo



The MRS was institutionalized by State Complementary Law No. 1.241, dated May 8, 2014, strategically located between the metropolitan areas of São Paulo and Curitiba in Brazil, having a process of the conurbation and territorial limit with the Metropolitan Region of Campinas, occupying the 15th place in the national economy with a Gross Domestic Product (GDP) equal to 67.24 billion Brazilian Reais, equivalent to 4% of the GDP of São Paulo State, the richest Brazilian State (SÃO PAULO, 2014; LOURENÇO et al., 2015; EMPLASA, 2020).

In the Brazilian scenario, MRS stands out for intense and diversified economic activity, characterized by highly developed industrial production, predominantly in the metallurgical, electro-electronic, textile and agribusiness sectors, being the largest agricultural producer among the areas of the metropolitan region of the São Paulo State, with high diversity. It has an important contribution in the state production of ores, such as cement, limestone, ornamental rock, gravel stone, clay, carbonate, and tungsten (EMPLASA, 2013; OLIVEIRA et al., 2016).

2.2 Spatial Data Analysis

The data used was available by the Demographic Census of the year 2010, made by the Brazilian Institute of Geography and Statistics (IBGE), and the data collection was performed by census tracts of the municipalities present in the Metropolitan Region of Sorocaba (MRS).

The average per capita household income of permanent private households (INCOME) was one of the information collected to verify how the income of the inhabitants of these regions is linked to the infrastructure of these areas.

The data collected was transformed into a percentual, such as permanent private households with garbage collection (WASTE), permanent private households with water supply from the general network, well or spring on the property (WATER SUPPLY), and permanent private households with bathrooms exclusively for residents or sanitary sewage via the general sewage network or septic tank (SEWAGE), to verify the locations that need a greater investment of the municipalities.

The census tracts with several households equal to 0 or omitted data, were disregarded and replaced by data from the adjacent census tract.

In order to verify the spatial autocorrelation of the studied variables for the MRS, was necessary to estimate the spatial variability of these data for each municipality. For this, was necessary to construct a Neighborhood Matrix, and the adopted criterion was the contiguity. Also, the matrix was normalized per line with the sum of the weights of each line being equal to 1 (SHEKHAR et al., 2011; SILVA et al., 2014; RIGHETTO; TACHIBANA, 2015; PENSO; PÉRICO, 2016).

From the Proximity Matrix, the Global Moran Index (GMI) was calculated as a global measure of the spatial autocorrelation that was used in this study to measure the autocorrelation and spatial dependence between the census tracts of the municipalities and their respective variables. The GMI value is equal to 0, indicating that there is spatial independence. However, the positive values (> 0 to 1) indicate a direct correlation, that is, the value of a given variable of a census tract tends to be like the values of the neighboring census tracts, whereas, negative values (<0 to -1) indicate an inverse correlation (MONDINI; CHIARAVALLOTI-NETO, 2008; SILVA et al., 2014; RIGHETTO; TACHIBANA, 2015). The GMI was calculated by Equation 1.

$$GMI = \frac{n}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}} \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \text{ to } i \neq j \quad (1)$$

Where:

GMI is the Global Moran Index;

n corresponds to the number of census tracts in the Metropolitan Region of Sorocaba;

yi is the value of the attribute considered in the census tract i;

yj is the value of the attribute considered in the census tract j;

$\bar{y}$ represents the average value of the attribute for the Metropolitan Region of Sorocaba;

wij are the weights assigned in determining the proximity matrix according to the connection between areas i and j.

Once the GMI is calculated, it is important to establish the statistical validity, and, for this, a pseudo-significance test was used for several permutations equal to 99 (MONDINI; CHIARAVALLOTI-NETO, 2008; SILVA et al., 2014; RIGHETTO; TACHIBANA, 2015).

After the analysis of the global autocorrelation, the local spatial association of the studied variables that showed significant global autocorrelation was verified through the Local Moran Index (LMI). This procedure was performed to group the census tracts with the greatest spatial similarity, making possible to highlight trends in the studied variables and their spatial distribution patterns.

Differently than the GMI that provides a single value as a measure of spatial association for the study area, the LMI produces a specific value for each census tract, enabling the identification of clusters in census tracts with values of similar indexes (clusters), census tracts with anomalous rates (outliers) and more than one spatial regime (ANSELIN, 1995; SILVA et al., 2014; RIGHETTO; TACHIBANA, 2015). Therefore, this ability to detect significant clusters allows identifying the locations where such characteristics predominate spatially. The LMI was calculated by Equation 2.

$$LMI = \frac{(y_i - \bar{y}) \sum_{j=1}^{n} w_{ij} (y_j - \bar{y})}{\frac{\sum_{i=1}^{n} (y_i - \bar{y})^2}{n}} \qquad 2)$$

Where:

LMI is the Local Moran Index;

n corresponds to the number of census tracts in the Metropolitan Region of Sorocaba;

y_i is the value of the attribute considered in the census tract i;

y_j is the value of the attribute considered in the census tract j;

$\bar{y}$ represents the average value of the attribute for the Metropolitan Region of Sorocaba;

w_{ij} are the weights assigned in determining the proximity matrix according to the connection between areas i and j.

The mathematical operations described, and the database was manipulated in the TerraView 4.2.2 software, provided by the Brazilian National Institute of Space Research (INPE), while the map layout was generated in ArcGIS 10.4 software.

For the census tracts where was a spatial correlation between the rates studied, a statistical analysis was performed to verify if there is a correlation between these rates. Therefore, the census tracts that showed spatial correlation for each variable were overlapped and the coincident census tracts were identified for all variables. With the data from these census tracts, a multiple linear regression analysis was performed. This analysis presupposes the construction of multiple linear equations with independent variables that can explain the variation of a dependent variable, as described in Equation 3.

$$Y=a+\sum_{i=1}^{n} b_i \times x_i \qquad 3)$$

Where, in this study, Y is the dependent variable represented by INCOME, while x_i is one of the independent variables inserted in the model and corresponds to the variables related to basic sanitation, a is the intercept related to the load of the independent variable when the dependent variable is zero, and bi are the individual explanatory coefficients of the independent variables, which define the increase (or decrease) of the influence on the variable Y per unit change in xi.

Data processing was performed in Microsoft Excel 2010 software, whereas Pearson's correlation coefficient and multiple linear regression calculations were performed in BioEstat 5.0 software.

3 RESULTS AND DISCUSSION

Table 1 shows the values of the Global Moran Index (GMI) for each variable studied in the 2735 census tracts present in the MRS.

Table 1 – Global Moran Index values for the studied variables

Index	Income	Waste	Water supply	Sewage
Global Moran Index Value	0.38	0.32	0.18	0.53
p-value	0.01	0.01	0.01	0.01

It was verified that there is spatial autocorrelation presence of all the variables studied by the GMI for a number of permutations equal to 99, since the GMI presented values on average above 0.30, with the significance of 0.01, an exception was the WATER SUPPLY variable that presented few census tracts with spatial autocorrelation, thus, this variable was not used for the Local Moran Index analysis.

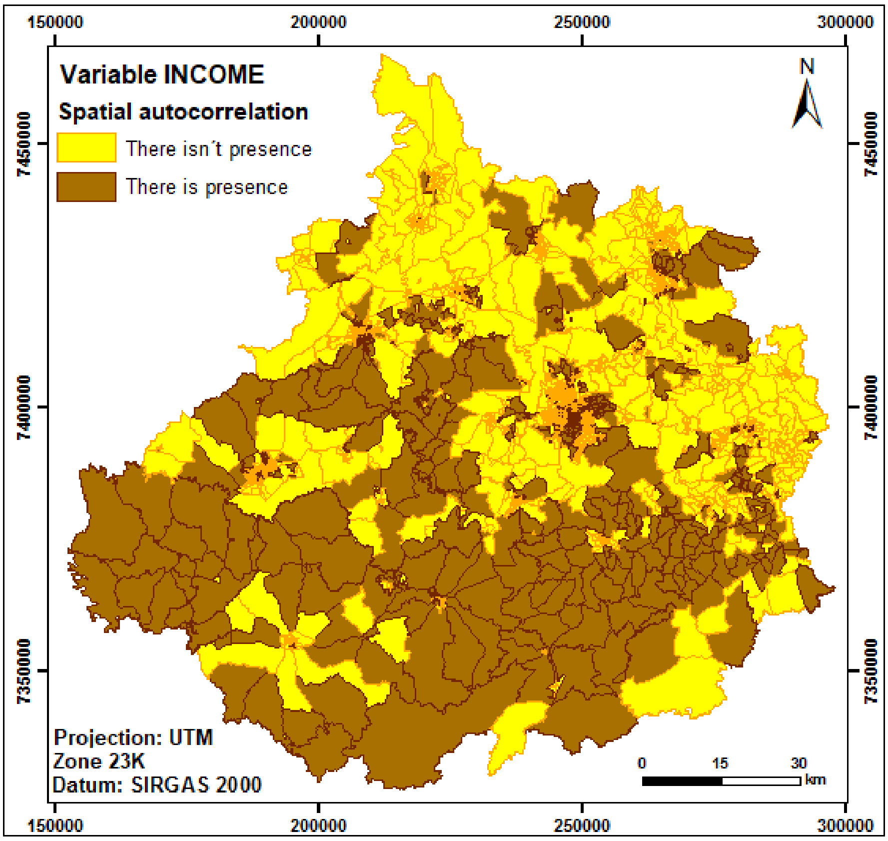
Nunes (2013) investigated the spatial autocorrelation of census tracts with permanent private households with garbage collection for the Regions North and Northeast in Goias State in Brazil, with GMI values of 0.06 and p-value equal to 0.22. Differently, from this study, the results found by Nunes (2013) did not present spatial autocorrelation, evidencing that these regions present a probable autonomy of the municipal administrative policies that interfere in this variable. The author also analyzed the variable income, being found the GMI value of 0.31 and p-value equal to 0.01, like the value found for the MRS, indicating spatial autocorrelation for this variable.

Ribeiro et al. (2013) characterized the spatial and temporal distribution of Brazilian spotted fever (BSF) in the São Paulo State from 1998 to 2010. These authors used the data of precipitation, temperature, land use and soil cover in their correlations, finding a coefficient of 0.22 with a p-value of 0.01, indicating that the areas with the highest population density and lower vegetation cover, rural areas, and planted areas, presented a high potential for the onset of spotted fever.

To identify the census tracts that present similar and clustered values for the studied variables, the Local Moran Index was calculated, where areas with significance values less than 0.05 are represented in the maps of Figures 2 to 4 in brown and indicate spatial autocorrelation, whereas areas with significance values greater than 0.05 are represented by yellow color and indicate census tracts that do not show spatial autocorrelation.

Figure 2 shows the map with the spatial autocorrelation distribution for the INCOME variable for MRS.

Figure 2 – Spatial autocorrelation of the variable INCOME for the Metropolitan Region of Sorocaba



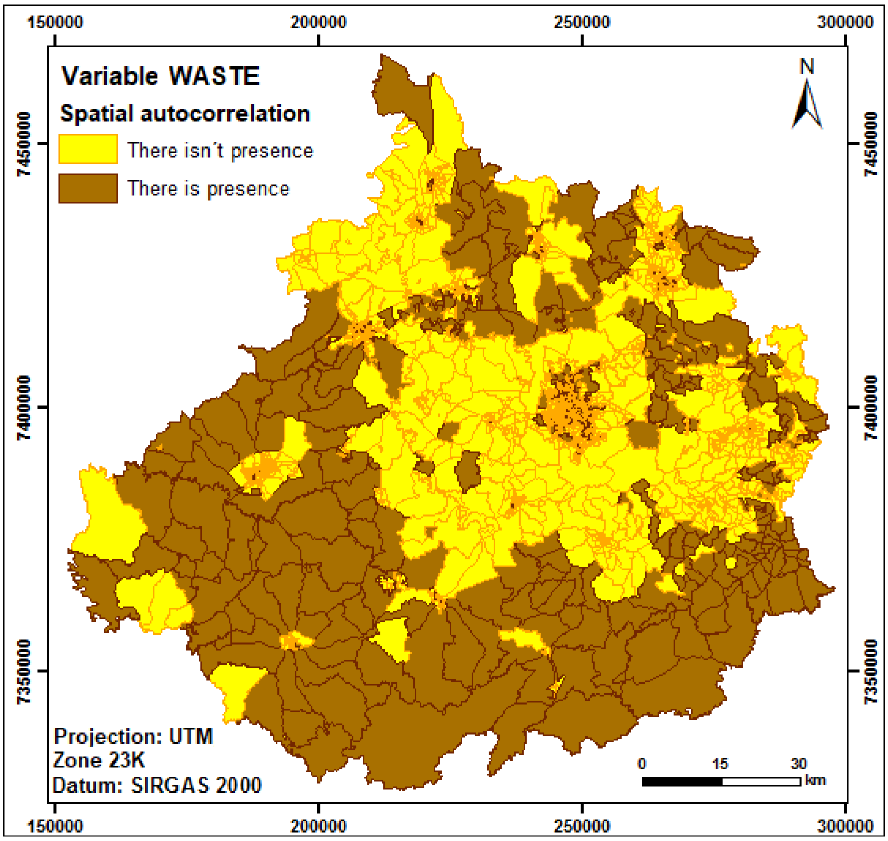
Was observed that the average per capita income of the resident population in the MRS is equal to 1,413.20 Brazilian Reais with a standard deviation equal to 1,057.40 Brazilian Reais, showing the non-homogeneous distribution of income, since the coefficient of variation is equal to 74.8%. Approximately 74.5% of the census tracts evaluated did not present spatial autocorrelation, but they are not representative since they represent 46.8% of the area. Regarding the population, the percentage of people in census tracts that do not show autocorrelation rises to 76.2%, indicating that most of the population is in areas that do not have a spatial distribution pattern.

In general, was possible to verify that the north, east and northeast regions showed a high homogeneity of census tracts with the absence of spatial autocorrelation with the variable INCOME. Whereas, the central region of the MRS presented greater heterogeneity in relation to the spatial distribution of the variable; while in the south, southeast and southwest there is a prevalence of spatial autocorrelation.

Therefore, census tracts with a similar average income are generally closer to each other on the map presented and are in municipalities where agricultural activity is an important source of income for their residents. Similar study was carried by Righetto and Tachibana (2015) using the Global Moran Index and Local Moran Index for some independent variables related to the quality of life and income of the residents of the municipality of Presidente Epitácio in Sao Paulo State (Brazil), where a positive significance value was found between variables quality of life and income, representing spatial autocorrelation between the income variable and the resident population in the south-central regions, with GMI being 0.1769, below the values found in this study for variables sanitary sewage, income, and garbage collection.

Figure 3 shows the map with spatial autocorrelation for the WASTE combination for MRS.

Figure 3 – Spatial autocorrelation of the WASTE variable for the Metropolitan Region of Sorocaba



The mean percentage of census tracts with permanent private households served with garbage collection at MRS is 92.7%, with a standard deviation equal to 20.6%, showing the non-homogeneous distribution of the WASTE variable, since the coefficient variation is equal to 22.2%.

Approximately 84.4% of the census tracts evaluated did not present spatial autocorrelation, but they are not representative since they represent 42.5% as can be observed in Figure 3.

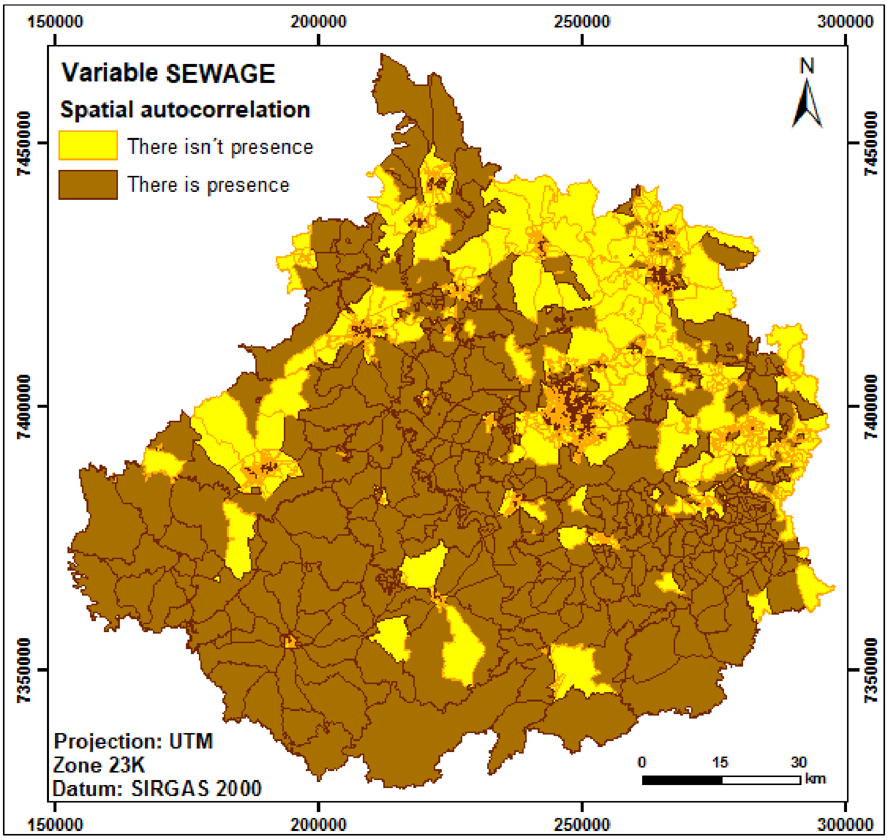
When it comes to the population, the percentage of people that are in census tracts that do not show autocorrelation rises to 87.1%, representing that most of the population is in areas that do not have a spatial distribution pattern. Of the census tracts that present spatial autocorrelation for the WASTE variable, 10.9% are in census tracts classified as rural, but when is considered the area occupied by the census tracts in the rural area, these values increase to 76.3% due to the size of census tracts in rural areas.

The central and northern zone of MRS presented the largest cluster of census tracts with the absence of spatial autocorrelation, whereas the sectors located to the southeast, south, and west of the MRS had more clusters of census tracts with the presence of autocorrelation among the variable WASTE.

The areas that present a constant and permanent low collection rate are generally closer to each other in the presented map (Figure 3), and the areas with high rates of garbage collection are found in the census tracts that do not have such characteristics and do not present relation in space. This lack of spatial relation in the central region of MRS is due to the difference between the infrastructure presented by the urban area of the city of Sorocaba and Votorantim to the others adjacent municipalities, is clearly demonstrated by the absence of a spatial relation between the census tracts present in this region.

Figure 4 shows the map with spatial autocorrelation distribution for SEWAGE variable for MRS.

Figure 4 – Spatial autocorrelation of the variable SEWAGE for the Metropolitan Region of Sorocaba



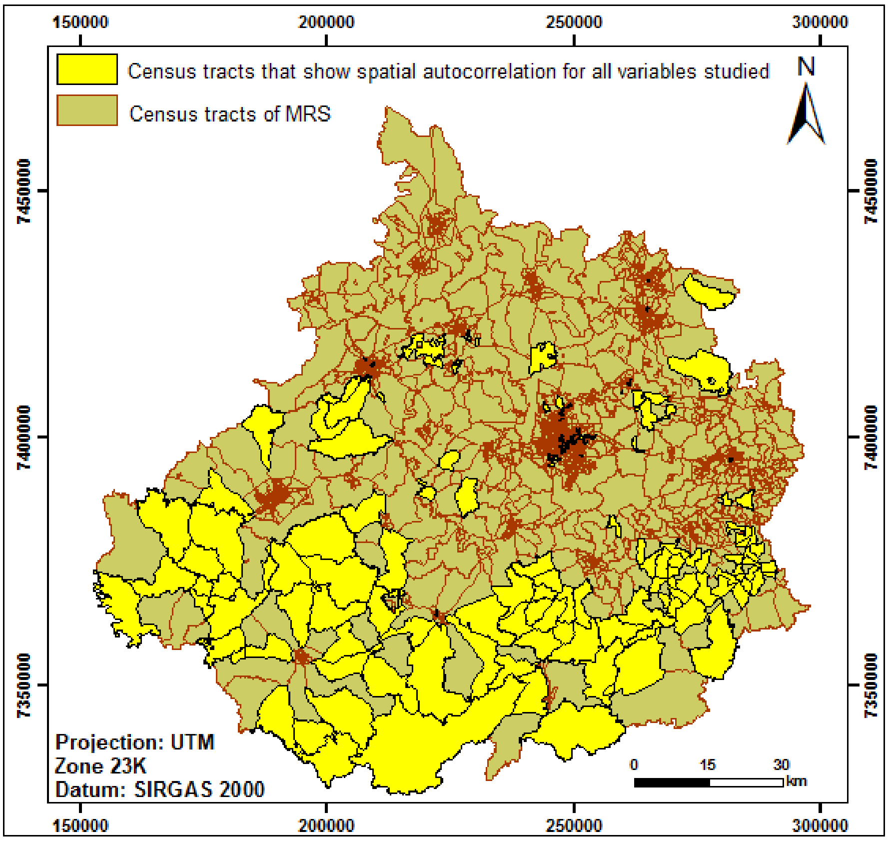
When verifying in Figure 4, was observed that the mean percentage of census tracts with permanent private households with sanitary sewage served with the general sewage network or septic tank in MRS is equal to 82.5%, with a standard deviation equal to 33%, evidencing the non-homogeneous distribution of the variable SEWAGE, since the coefficient of variation is equal to 40%. Approximately 63.7% of the census tracts evaluated do not have spatial autocorrelation, but are not representative, since they represent 29.2% of the occupied area.

When evaluating the population, the percentage of people who are in census tracts that do not show autocorrelation is 67.6%. However, in the census tracts with spatial autocorrelation for SEWAGE variable, 8% are in census tracts classified as rural, but when is considered the area occupied by the census tracts in the rural area, these values increase to 68.9% due to the size of census tracts in rural areas.

The census tracts with a similar sewage collection rate are closer to each other in the presented map and are homogeneously located in almost the entire extent of MRS, except for the northeastern region of the study area contemplated by the municipalities of Porto Feliz, Itu, and Salto.

Figure 5 shows the spatial distribution of the studied variables with coincidence or overlapping of the areas with spatial autocorrelation.

Figure 5 – Census tracts that show spatial autocorrelation and overlap between the variables studied in MRS



There are 177 census tracts distributed by MRS that present autocorrelation for all variables studied and was used to verify if there is a relation between the data of these respective areas. These census tracts correspond to 31.1% of the MRS territory and have a population of 100,587 inhabitants, corresponding to 5.4% of the total MRS population. The occupied area is extensive, but the population is smaller because these census tracts (62.7%) are present in rural areas.

For these census tracts where was a spatial correlation and overlap between the studied variables, a multiple linear regression analysis was performed to verify if there is a correlation between them. It was verified that the per capita income in these census tracts is dependent on other variables.

The correlation coefficient (r) was 0.428, showing that the variable INCOME is dependent to the other variables since the value of F (Regression) was 19.5 with very significant value ($p = 0.0001$), rejecting the null hypothesis. It is also verified by the coefficient of determination (r^2) that shows only 18.3% of the per capita income of the residents of these census tracts is explained by the sanitation conditions in which they live, suggesting that other external factors, not evaluated in the present study can act as predictors of income variation.

However, was correlation only between the INCOME and SEWAGE variables, with $p < 0.0001$. This explains that the per capita income increases as the percent of households connected to the sewage system or septic tank grow or vice versa. Thus, this correlation found among the studied variables can be considered statistically significant for this study.

Approximately 50% of census tracts with spatial autocorrelation (Figure 5) have a sanitary sewage collection rate of over 80% and are areas with a good infrastructure of sanitary sewage. Concluding with this result, approximately 85% of census tracts with spatial autocorrelation have a rate of over 80% for access to the drinking water but adequate sanitary sewage does not follow this efficiency.

The mean per capita income of the census tracts with spatial autocorrelation (Figure 5) is close to 1,415.00 Brazilian Reais, and approximately 26% of the census tracts present a higher income average, which presents a spatial correlation between them and development in basic infrastructure, resulting in the opportunity to attract financial resources to the residents higher than MRS.

The other census tracts that presented this spatial relation (74%) have deficits in infrastructure and low income per capita, requiring financial resources that would allow greater access to quality water and its dump more appropriately, however, regarding collection of waste, there is a need for greater intervention of the public managers to guarantee access to this basic right for the population.

Also, in Figure 5, it is possible to infer that the census tracts with spatial autocorrelation extrapolate the political-administrative limits of the municipalities present in the MRS and, for this reason, there is the possibility of external investments to the MRS by the public managers, that can be destined for mitigating the impacts of low infrastructure in basic sanitation.

The studies that deal with spatial correlations of census data and socio-environmental conditions are extremely important since they infer about the social and environmental conditions of the regions and municipalities studied, providing a general overview of the quality of life of the population and the precarious infrastructure of these localities. Teixeira and Guilhermino (2006) affirm that investments in sanitation are essential, as they verified in their studies that the absence of sanitary sewage systems has a direct relation with the mortality due to infectious and parasitic diseases in the Brazilian States.

Therefore, spatialization of sanitary conditions makes possible to associate the purchasing power of residents in a given area, as well as the number of the occurrences of waterborne diseases that can affect the residents, regardless of social condition, when exposed to this water with low quality due to precarious health conditions.

In this scenario of pluralities, development is closely linked to issues concerning human health (CARRARA; VENTURA, 2012; UNDP, 2014). In this sense, this study can be taken as guide, providing subsidies to the governments in decision-making in the face of the resources to be used to improve the infrastructure conditions of states and municipalities, since it aim metropolitan management performed by the government, integrating actions that aim at social and economic improvements among municipalities.

According to Castro and Santos Junior (2017), due to the socio-spatial complexity present in the metropolitan regions, it is necessary to implement public policies to assist urban problems that are very noticeable in these spaces, as identified in this study, such as the strong spatial correlation between the low-income census tracts and the census tracts with low infrastructure of the sanitary sewage system. This correlation between the census tracts with low sanitary sewage infrastructure and with low per capita income corroborates with the Castro and Santos Junior (2017) studies, which verified a continuity of socio-spatial inequalities in the expansion of the São Paulo metropolis in Brazil, showing that metropolization is accentuating these social inequalities.

Also, this spatial analysis technique allows from a set of any variables presuppose some part of the reality, but it is not able to translate completely the complexity of the processes in the production of space. In this sense, when verifying where is a low income, garbage collection and sanitary sewage that are spatially coincident, these do not reflect the absolute reality of these census tracts, as there are households that do not fit into this profile and, therefore, this homogenization of the spatialization of the phenomena by census tracts do not faithfully portray the real characteristics of the living conditions of all residents in these sectors.

However, what is sought in this study is to homogeneously evaluate these areas within the scope of public investments for improving the quality of life of the residents in these areas, since for public managers it is necessary to establish zones with differentiated degrees of priorities investments.

There are significant questions that help to understand the geographic reality under analysis, such as the physical characteristics of the area of study, type of predominant economic activity, among others. However, the fact that there is a deficit of sanitary sewage collection already reflects the need for changes and greater investment for the region, as well as a greater understanding of the governments of these areas regarding the need for joint actions to minimize the impacts of this situation that extrapolates the municipal boundary.

Regarding the study area, was observed that the presence of spatial autocorrelation for all variables are in a region where the municipalities have the lowest per capita income in the MRS, except for some small census tracts which are distributed in the north and east of the city of Sorocaba. These locations, except for the census tracts located in the city of Sorocaba, are mostly located in rural areas where has less investment in infrastructure, and the number of residents in these areas is much lower than in other census tracts present in the study area. These census tracts present the largest dimensions, which makes more difficult and expensive the public investments compared to the smaller census tracts, being all these factors presented relevant to justify the numerical heterogeneity presented by the studied variables.

Therefore, this study presents a proposal of extreme relevance that brings aid to the reality of the socio-environmental conditions of the municipalities inserted in the MRS, making possible the identification of the regions that extrapolate the administrative limits of the municipalities that lack collective investments and cooperation of the municipalities, since it is impossible to discontinue these socio-environmental inequalities by working only the administrative political limits. Thus, methodologies such as the presented in this study are important instruments of public management, since they help in the elaboration of public policies and orientation of public expenditures.

4 CONCLUSION

Was verified that there is a spatial correlation for all variables studied, however, the variable related to water supply presented low spatial autocorrelation, while the sanitary sewage variable presented the most significant spatial autocorrelation. Regarding the correlation between the variables for the census tracts, it was observed that there is a correlation only between the SEWAGE and INCOME variables.

There are 177 census tracts distributed by MRS that present spatial autocorrelation for all variables studied. These census tracts correspond to 31.1% of the MRS territory and have a population of 100,587 inhabitants, corresponding to 5.4% of the MRS total population. For these census tracts, a correlation coefficient (r) of 0.428 was found, demonstrating that the variable INCOME is dependent on the other variables studied. It is also verified by the coefficient of determination (r^2) that only 18.3% of the per capita income of the residents of these census tracts is explained by the basic sanitation conditions which they live, suggesting that other external factors, not evaluated in the present study, can act as predictors of income variation.

The census tracts that do not have the sanitary infrastructure and low-income are in rural areas and are much larger than the census tracts present in urban areas. For public managers, this study shows where there is a deficit in the services offered and enables public managers to make strategic decisions since it is possible to know the peculiarities of the different census tracts.

Therefore, the objective of using the variable space in this study was to identify if there are census tracts that present similar characteristics in relation to the studied rates, being these considered high or low and being indicative of precarious conditions or not. From the moment that most of the population resides in census tracts that do not show spatial autocorrelation between them, it can be concluded that the rates studied are independent of their geographical position and are determined by other factors than space.

This study can be used by public managers of metropolitan regions to focus their efforts and resources for the development of these areas and to elaborate public policies that aim at the improvement of the quality of life of the population, since they can prioritize the investment and carry out directed actions from the intersection of census data with information about per capita income, garbage collection, sanitary sewage, as well as the destination of the sewage, serving as a methodological proposal for the analysis of metropolises aiming to promote public policies aimed at improving the quality of life and the environment in these areas.

ACKNOWLEDGEMENT

Thanks to reviewers, collaborators, and São Paulo State University (Unesp).

REFERENCES

Anselin, L. (1994). Local indicators of spatial association-LISA. **Geographical analysis**, 27(2), 93-115.

Carrara, B. S., Ventura, C. A. A. (2012). A saúde e o desenvolvimento humano. **Saúde & Transformação Social**, 3(4), 89-96.

Emplasa, Empresa Metropolitana de Planejamento da Grande São Paulo. (2013). **Plano de Ação da Macrometrópole Paulista 2013/2040**. Emplasa.

Castro, H. R., Santos Júnior, W. R. (2017). A expansão da macrometrópole e a criação de novas RMs: um novo rumo para a metropolização institucional no estado de São Paulo. **Cadernos Metrópole**, 19(40), 703-720.

Emplasa, Empresa Metropolitana de Planejamento da Grande São Paulo. (2020). **Sobre a RMS**, URL https://emplasa.sp.gov.br/RMS.

Hallisey, E., Tai, E., Berens, A., Wilt, G., Peipins, L., Lewis, B., Graham, S., Flanagan, B., Lunsford, N. B. (2017). Transforming geographic scale: a comparison of combined population and areal weighting to other interpolation methods. **International Journal of Health Geographics**, 16(29), 1-16.

Koch, T., Denike, K. (2004). Medical Mapping: The Revolution in Teaching – and Using – Maps for the Analysis of Medical Issues. **Journal of Geography**, 103(1), 76-85.

São Paulo. (2014). Lei complementar n. 1.241 de 8 de maio de 2014 - Cria a Região Metropolitana de Sorocaba e dá providências correlatas. **Diário Oficial da União** (São Paulo).

Lourenço, R. W., Silva, D. C. C., Sales, J. C. A., Medeiros, G. A., Otero, R. A. P. (2015). Metodologia para seleção de áreas aptas à instalação de aterros sanitários consorciados utilizando SIG. **Ciência e Natura**, 37, 122-140.

Meira, R. T., Sabonaro, D. Z., Silva, D. C. C. (2016). Elaboração de Carta de Adequabilidade Ambiental de uma pequena propriedade rural no município de São Miguel Arcanjo, São Paulo, utilizando técnicas de geoprocessamento. **Engenharia Sanitária e Ambiental**, 21, 77-84.

Mondini, A, Chiaravalloti-Neto, F. (2008). Spatial correlation of incidence of dengue with socioeconomic, demographic and environmental variables in a Brazilian city. **Science of The Total Environment,** 393(2-3), 241-248.

Nery, L. M., Simonetti, V. C., Machado, L. P., Silva, D. C. C. (2020). Geotecnologias aplicadas na análise do risco de contaminação de poços de água no município de Sorocaba, SP. **Holos Environment,** 20, 214-230.

Nunes, F. G. (2013). Análise exploratória espacial de indicadores de desenvolvimento socioambiental das regiões de planejamento do norte e nordeste goiano. **Ateliê Geográfico**, 7(1), 237-259.

Oliveira, R. A., Silva, D. C. C., Simonetti, V. C., Stroka, E. A. B., Sabonaro, D. Z. (2016). Proposição de Corredor Ecológico entre duas Unidades de Conservação na Região Metropolitana de Sorocaba. **Revista do Departamento de Geografia**, 2, 61-71.

Ribeiro, M. D., Furtado, M. A., Ferraudo, A. S., Cesario, M., Morraye, M. A. (2013). Fatores ambientais envolvidos na epidemiologia da febre maculosa no estado de São Paulo. **Hygeia**, 9(16), 103-114.

Richards, T. B., Croner, C. M., Rushton, G., Brown, C. K., Fowler, L. (1999). Geographic information systems and public health: mapping the future. **Public Health Reports**, 114(4), 359-373.

Righetto, A. J., Tachibana, V. M. (2015). Análise de regressão linear: abordagem tradicional e espacial em um estudo de caso. **Revista Brasileira de Qualidade de Vida**, 7(3), 169-179.

Rostami, M., Mohammadi, Y., Jalilian, A., Nazparvar, B. (2017). Modeling spatio-temporal variations of substance abuse mortality in Iran using a log-Gaussian Cox point Process. **Spatial and Spatio-temporal Epidemiology,** 22, 15-25.

Silva, D. C. C., Lourenço, R. W., Donalisio, M. R., Cordeiro, R. (2014). Análise da relação entre a distribuição espacial das morbidades por obesidade e hipertensão arterial para o estado de São Paulo, Brasil, de 2000 a 2010. **Ciência & Saúde Coletiva**, 19(6), 1709-1719.

Silva, D. C. C., Albuquerque Filho, J. L., Oliveira, R. A., Lourenço, R. W. (2017). Proposta Metodológica para Análise da Inserção Social na Bacia Hidrográfica do Rio Una em Ibiúna/SP. **Revista do Departamento de Geografia**, 33, 74-84.

Shekhar, S., Evans, M. R., Kang, J. M., Mohan, P. (2011). Identifying patterns in spatial information: A survey of methods. **WIREs Data Mining Knowledge Discovery**, 1(3), 193-214.

Penso, J. M., Périco, E. (2016). Análise Espaço-Temporal da Mortalidade por Diabetes Mellitus no Rio Grande do Sul. **Revista Brasileira de Geografia Física**, 9(6), 1836-1848.

Teixeira, J. C., Guilhermino, R. L. (2006). Analysis of association between sanitation and health in brazilian states, using secondary data from data bank indicadores e dados básicos para a saúde 2003 - IDB 2003. **Revista de Engenharia Sanitária Ambiental**, 11(3), 277-282.

UNDP, United Nations Development Programme. (2014). **Human Development Reports; Sustaining Human Progress: Reducing Vulnerability and Building Resilience**, United Nations Development Programme.

AUTHORSHIP CONTRIBUTIONS

1 – Darllan Collins da Cunha e Silva
Doutor e Professor do Departamento de Engenharia Ambiental da Universidade Estadual Paulista (Unesp), Instituto de Ciência e Tecnologia de Sorocaba (ICTS).
http://orcid.org/0000-0003-3280-0478 – darllan.collins@unesp.br
Contribution: Conceptualization, Investigation, Formal Analysis, Methodology, Writing – original draft.

2 - Vanessa Cezar Simonetti
Doutoranda do Programa de Pós-Graduação em Ciências Ambientais da Universidade Estadual Paulista (Unesp), Instituto de Ciência e Tecnologia de Sorocaba (ICTS).
http://orcid.org/0000-0001-6845-4875 – va_simonetti@hotmail.com
Contribution: Data curation, Writing – original draft.

3 - Renan Angrizani de Oliveira
Doutorando do Programa de Pós-Graduação em Ciências Ambientais da Universidade Estadual Paulista (Unesp), Instituto de Ciência e Tecnologia de Sorocaba (ICTS).
http://orcid.org/0000-0001-9772-7756 – renan_angrizani@hotmail.com
Contribution: Data curation, Writing – original draft.

4 - Jomil Costa Abreu Sales
Doutor em Ciências Ambientais pela Universidade Estadual Paulista (Unesp), Instituto de Ciência e Tecnologia de Sorocaba (ICTS).
http://orcid.org/0000-0001-8722-8398 – jomilc@gmail.com
Contribution: Data curation, Writing – original draft.

5 - Roberto Wagner Lourenço
Doutor e Professor do Departamento de Engenharia Ambiental da Universidade Estadual Paulista (Unesp), Instituto de Ciência e Tecnologia de Sorocaba (ICTS).
http://orcid.org/0000-0002-5234-8944 – roberto.lourenco@unesp.br
Contribution: Supervision, Writing – review & editing.

HOW TO QUOTE THIS ARTICLEZ

SILVA, D. C. C.; SIMONETTI, V. C.; OLIVEIRA, R. A.; SALES, J. C. A.; LOURENÇO, R. W. Daily Global and diffuse radiation in the Brazilian Cerrado-Amazon transition region. Ciência e Natura, Santa Maria, v. 43, e42, p. 1-22, 2021. DOI: https://doi.org/10.5902/2179460X39332. Accessed: Month Abbreviated. day, year.