

## **Análise de regressão linear: abordagem tradicional e espacial em um estudo de caso**

### **Linear regression analysis: traditional and space approach in a case study**

Ana Julia Righetto  
Universidade de São Paulo/Escola Superior de Agricultura Luiz de Queiroz – USP/ESALQ – Piracicaba – São Paulo -  
Brasil  
[ajrighetto@gmail.com](mailto:ajrighetto@gmail.com)

Vilma Mayumi Tachibana  
Universidade Estadual de São Paulo – UNESP – Presidente Prudente – São Paulo - Brasil  
[vilma@fct.unesp.br](mailto:vilma@fct.unesp.br)

#### **RESUMO:**

**OBJETIVO:** Analisar a qualidade de vida e renda dos municípios da cidade de Presidente Epitácio, São Paulo, por meio de métodos estatísticos.

**MÉTODOS:** O estudo foi realizado com base em um questionário aplicado aos moradores de Presidente Epitácio, São Paulo, no ano de 2008, como parte da pesquisa desenvolvida por Santos (2010), na qual foram considerados os 47 setores censitários do município definidos pelo Instituto Brasileiro de Geografia e Estatística (IBGE) para o censo demográfico de 2000. Para a análise estatística, foram utilizados os modelos de regressão linear simples e múltipla e, posteriormente, foi considerada a dependência espacial das observações por meio do modelo espacial autorregressivo misto (*Spatial AutoRegressive - SAR*) e do modelo de erro espacial (*Conditional AutoRegressive - CAR*). Todos os procedimentos estatísticos e computacionais foram realizados com auxílio dos programas Minitab 17 e GeoDa.

**RESULTADO:** Após uma análise exploratória prévia e ajuste do modelo de regressão linear clássico, verificou-se a necessidade da inclusão da dependência espacial nas variáveis em estudo. Dentre os modelos espaciais ajustados, verificou-se, por meio de critérios estatísticos, que o modelo CAR é o mais adequado para descrever o problema considerado. O modelo foi capaz de apontar que a região mais ao sul da cidade de Presidente Epitácio, São Paulo, possui qualidade de vida mais baixa, ao passo que as regiões norte e nordeste do município possuem qualidade de vida mais alta.

**CONCLUSÕES:** A metodologia aplicada se mostrou uma poderosa ferramenta no que tange à distinção de áreas que apresentam maior ou menor qualidade de vida no município de Presidente Epitácio, São Paulo. Com isto, esta pesquisa pode oferecer subsídios para a criação de políticas públicas específicas para cada uma das regiões (setores censitários) da cidade em estudo.

**PALAVRAS-CHAVE:** Regressão linear. Qualidade de vida. Políticas públicas.

## ABSTRACT:

**OBJECTIVE:** Analyze the quality of life and income of the residents of the city of Presidente Epitácio, São Paulo, through statistical methods.

**METHODS:** The study was based on a questionnaire applied to the habitants of President Epitácio, São Paulo, in 2008, as part of research developed by Santos (2010), which was considered the 47 census tracts of the city defined by IBGE for the census demographic 2000. For statistical analysis, simple and multiple linear regression models were used and later was considered the spatial dependence of observations through the Spatial AutoRegressive (SAR) and Conditional AutoRegressive (CAR) models. All statistical and computational procedures were performed with the help of Minitab 17 and GeoDa programs.

**RESULTS:** After a preliminary exploratory analysis and adjustment of the classical linear regression model, there was the need to include spatial dependence on the variables under study. Among the adjusted spatial models, it was found, using statistical criteria, the CAR model is most suitable to describe the problem considered. The model was able to point out that the region's more south of Presidente Epitácio, São Paulo, has lower quality of life, while the northern and northeastern regions of the city have higher quality of life.

**CONCLUSIONS:** The methodology proved to be a powerful tool when it comes to areas of distinction that have higher or lower quality of life in the city of Presidente Epitácio, São Paulo. With this, this research can shed light on the creation of specific public policies for each region (census tracts) of the city under study.

**KEYWORDS:** Linear regression. Quality of life. Public policy.

## 1 Introdução

A qualidade de vida de uma determinada população pode ser avaliada levando-se em consideração fatores como educação, saúde, bem-estar físico, psicológico, mental, emocional, como também através da família, dos amigos, do emprego e outras circunstâncias da vida.

A qualidade de vida de uma determinada população está diretamente relacionada com o crescimento e desenvolvimento econômico da mesma. Segundo Souza (1993), o desenvolvimento econômico requer um ritmo de crescimento contínuo e superior ao crescimento da população, englobando mudanças estruturais e melhoria nos indicadores de qualidade de vida.

Atualmente existem diversos indicadores para qualificar e quantificar a qualidade de vida de uma população. Dentre os mais utilizados na literatura, destaca-se o Índice de Desenvolvimento Humano (IDH), o qual leva em consideração basicamente a expectativa de vida de uma população, a educação e o poder de compra.

Assim como o IDH, diversos índices apresentam insatisfatoriamente uma noção da qualidade de vida, pois quantificam, de modo geral, os problemas existentes em uma determinada população desconsiderando a relação espacial presente na mesma. Nesse ponto de vista, Barcellos et al. (2002) discutem o uso da estatística espacial e de indicadores na avaliação de situações de saúde. Com este tipo de análise é possível identificar e adotar medidas espaciais/regionais para a melhoria de tais indicadores, como o IDH.

Segundo Druck et al. (2004), compreender a distribuição espacial de dados provindos de fenômenos ocorridos no espaço constitui um grande desafio para esclarecer questões centrais em diversas áreas do conhecimento, como saúde, ambiente, geologia, agronomia, entre outras. Vários tipos de dados de interesse podem caracterizar problemas de análise espacial, os quais são divididos em dados de processos pontuais, de interação espacial e de área.

Os dados desse trabalho são caracterizados como dados de área pois o estudo foi feito com os setores censitários da cidade de Presidente Epitácio, São Paulo. O presente estudo tem como

objetivo avaliar a qualidade de vida dessa cidade por meio da aplicação de um questionário em que alguns atributos (atividade física, assistir ou não televisão, renda, entre outros) foram analisados.

## 2 Material e Métodos

### 2.1 Dados do estudo

Os dados utilizados neste trabalho foram retirados de uma pesquisa desenvolvida por Santos (2010), realizada em 2008, na qual foi aplicado um questionário aos moradores da cidade de Presidente Epitácio, São Paulo. Como unidades de referências para a amostragem, foram considerados os setores censitários, definidos pelo Instituto Brasileiro de Geografia e Estatística (IBGE) para o Censo Demográfico de 2000. No total são 10020 domicílios, distribuídos em 47 setores classificados em cinco classes de renda das quais foram retiradas amostras estratificadas, totalizando 596 domicílios, referentes à quantidade de questionários a serem aplicados.

O questionário, dividido em quatro partes, é composto por: dados pessoais, condição socioeconômica, meios de consumo coletivo e percepção dos moradores quanto à qualidade ambiental e a qualidade de vida. Com uma combinação linear de variáveis referentes a qualidade ambiental, a qualidade de vida e ao lazer, foi possível construir a variável resposta (Y), denominada de qualidade. As variáveis explicativas são relacionadas à renda do entrevistado, as informações sobre o imóvel que reside, às atividades de lazer que realiza e se possui convênio de saúde. Assim, as variáveis foram nomeadas por: renda ( $X_1$ ), imóvel alugado ( $X_2$ ), assistir televisão ( $X_3$ ), nenhuma ( $X_4$ ) e outra ( $X_5$ ), essas duas últimas referentes às atividades realizadas.

Do total de setores censitário da cidade de Presidente Epitácio, São Paulo, um setor não apresenta habitante. Além disso, por falta de tempo e de transporte, o questionário não foi aplicado em 5 setores com habitação, totalizando então 41 setores em que o questionário foi aplicado. Nos setores nos quais não foi aplicado, estimaram-se os valores das variáveis com base nos valores dos vizinhos de cada setor, realizando uma média dos valores dos vizinhos.

### 2.2 Metodologia

Para verificar a relação entre as diferentes características dos moradores da cidade Presidente Epitácio, São Paulo, e a qualidade de vida, foi utilizada a análise de regressão linear. Essa técnica pode ser dividida em: modelo de regressão linear simples (MRLS), que envolve apenas uma variável explicativa, e modelo de regressão linear múltipla (MRLM), com pelo menos duas variáveis explicativas. Segundo Charnet et al. (1999), tanto o MRLS como o MRLM descrevem a variável resposta como uma soma de uma parte aleatória e de uma parte determinística.

O modelo de regressão linear múltipla pode ser apresentado por:

$$y = X\beta + \varepsilon$$

Em que:

- $\varepsilon \sim N_n(0, \sigma^2 I)$ ,  $\mathbf{0}$  é o vetor nulo de dimensão  $n$ ;
- $\mathbf{I}$  é a matriz identidade  $n \times n$ ;
- $\beta$  é o vetor de parâmetros a serem estimados de dimensão  $(p+1)$ , sendo  $p$  o número de variáveis explicativas;
- $\mathbf{X}$  é a matriz das variáveis explicativas.

Um dos pressupostos que deve ser atendido pelo modelo de regressão múltipla é a independência dos erros. Para dados espaciais esse pressuposto é raramente satisfeito, pois é

comum nestes dados detectar estruturas de tendências ou autocorrelação espacial e, ao utilizar dados georreferenciados, faz-se necessário averiguar a possível existência dessas estruturas, a qual deve ser incorporada no modelo, uma vez que a dependência entre as observações altera o poder explicativo do mesmo (DRUCK et al., 2004). Caso os dados apresentem tal estrutura de autocorrelação espacial, os modelos MRLS e MRLM se tornam inadequados por não incorporarem em sua estrutura tais correlações.

Assim, faz-se necessária a criação de uma matriz de proximidade espacial, denotada por **W**, que é uma ferramenta básica para estimar variabilidade espacial de dados de área. Dado um conjunto de *n* áreas  $\{A_1, \dots, A_n\}$ , constrói-se a matriz  $\mathbf{W}^{(1)}_{n \times n}$  em que cada um dos elementos  $w_{ij}$  representa uma medida de proximidade entre  $A_i$  e  $A_j$ . A potência (1) indica que a vizinhança será de ordem 1; no caso de vizinhança de ordem 2, considera-se o vizinho do vizinho.

Essa medida de proximidade pode ser calculada a partir de um dos seguintes métodos:

- $w_{ij}=1$ , se o centroide de  $A_i$  está a uma determinada distância de  $A_j$ , caso contrário  $w_{ij}=0$ ;
- $w_{ij}=1$ , se  $A_i$  compartilha lado comum com  $A_j$ , caso contrário  $w_{ij}=0$ ;
- $w_{ij}=l_{ij}/l_i$ , em que  $l_{ij}$  é o comprimento da fronteira entre  $A_i$  e  $A_j$  e,  $l_i$  é o perímetro de  $A_i$ .

Geralmente, as linhas desta matriz são normalizadas para que a soma dos pesos de cada linha seja igual a um.

Neste trabalho, utilizou-se o modelo espacial autorregressivo misto (*Spatial AutoRegressive - SAR*) e o modelo de erro espacial (*Conditional AutoRegressive - CAR*), que segundo Wall (2004), são modelos que produzem dependência espacial nas estruturas de covariâncias como funções da matriz de vizinhança **W**.

Após construir a matriz **W**, é necessário verificar a existência de correlação espacial e, uma das formas mais utilizadas para tal verificação é utilizando o Índice de Moran (1948) e o *Local Indicators of Spatial Association* (LISA) proposto por Anselin (1995). Esses índices que medem a dependência espacial podem ser classificados como globais e locais, sendo que os globais fornecem uma medida de associação única para todo o conjunto de dados e, os locais, produzem valores específicos para cada observação, permitindo um melhor agrupamento quando há dependências espaciais acentuadas.

O Índice Global de Moran pode ser obtido por:

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

Em que:

- n* é o número de áreas (como números de setores censitários);
- $z_i$  o valor do atributo considerado na área *i*;
- $\bar{z}$  é o valor médio do atributo na região de estudo;
- $w_{ij}$  são os elementos da matriz **W**.

O Índice Global de Moran *I* é uma técnica de autocorrelação espacial e tem por objetivo indicar o quanto regiões distintas estão correlacionadas para uma mesma variável. Segundo Bailey e Gatrell (1995), esse índice assume valores no intervalo -1 e 1, sendo que 0 indica inexistência de autocorrelação (independência espacial), valores acima de 0 indicam associação positiva, ou seja, um padrão de semelhança entre as regiões e, valores abaixo de 0, associação negativa, ou seja, dessemelhança entre as regiões estudadas. A sua interpretação é igual à correlação de Pearson.

Por se tratar de estatística, segundo Lima, Caldarelli e Camara (2014), faz-se necessário verificar sua significância, que pode ser realizada através de um teste de hipótese, na qual a hipótese de nulidade  $H_0$  corresponde à ausência de padrão espacial e, a hipótese alternativa  $H_a$ , corresponde à presença de dependência espacial. As hipóteses podem ser testadas a partir da estatística de teste do Índice de Moran e dois procedimentos de testes são usuais. O primeiro é o teste baseado na distribuição normal, que se assenta na distribuição assintótica do Índice de Moran e é adequado em situações em que a variável observada possui distribuição ao menos aproximadamente normal. O

segundo e mais geral, é o baseado no procedimento Monte Carlo (teste aleatorizado), em que a distribuição amostral é obtida através do cálculo do índice após permutações dos dados entre as áreas.

O índice LISA precisa satisfazer a condição de que para cada observação o índice deve oferecer uma indicação de aglomerados espaciais significantes de valores similares. Anselin (2005) propôs apresentar esses resultados em um mapa colorindo as áreas com agrupamento espacial significativo em quatro cores diferentes. O vermelho representa agrupamento espacial de valores altos, ou seja, as áreas que o compõem têm valores da variável de interesse altos e estão cercados por áreas com valores altos. O azul marinho representa áreas que têm valores baixos, assim como seus vizinhos. Áreas com valores altos (baixos), mas com vizinhos com valores baixos (altos), são representadas em rosa (azul claro).

Neste trabalho serão adotados dois modelos espaciais, o modelo SAR e o modelo CAR. No primeiro, considera-se a dependência espacial por meio da adição, ao modelo de regressão, de um novo termo na forma de uma relação espacial para a variável resposta, isto é, quando se atribui à variável resposta a autocorrelação espacial é ignorada. O modelo é expresso por:

$$Y = \rho WY + X\beta + \varepsilon$$

Em que:

- $W$  é a matriz de proximidade espacial;
- o produto  $WY$  expressa a dependência espacial em  $Y$
- $\rho$  é o coeficiente espacial autorregressivo.

A hipótese nula para a não existência de autocorrelação é  $H_0: \rho = 0$ .

No modelo CAR os efeitos espaciais são um ruído, ou uma perturbação, que precisa ser removido(a). Neste modelo, os efeitos de autocorrelação espacial são associados ao termo do erro e o modelo é dado por:

$$Y = X\beta + \varepsilon \quad \text{e} \quad \varepsilon = \lambda W_\varepsilon + \epsilon$$

Em que:

- $W_\varepsilon$  é a componente do erro com efeitos espaciais;
- $\lambda$  é o coeficiente autoregressivo;
- $\epsilon$  é a componente do erro com variância constante e não correlacionada.

A hipótese nula, para não existência de autocorrelação é  $H_0: \lambda=0$  o que significa que o termo de erro não é espacialmente correlacionado.

Para a seleção de modelos utilizou-se o critério de informação de Akaike (AIC) e o critério Bayesiano de Schwarz (SBC), dados respectivamente por:

$$AIC = -2 \cdot \text{lik} + 2k \quad \text{e} \quad SBC = -2 \cdot \text{lik} + k \cdot \ln(N)$$

Em que:

- lik é o log da verossimilhança;
- $k$  é o número de coeficientes de regressão;
- $N$  é o número de observações.

Para ambos os critérios, em comparações de modelo, o melhor modelo será o com menos valor dos critérios.

As análises e figuras foram todas realizadas com o auxílio dos softwares Minitab 17 e GeoDa (ANSELIN, 2005).

### 3 Resultados e Discussão

A cidade de Presidente Epitácio, São Paulo, é dividida pelo IBGE em 47 setores censitários, dos quais um setor foi retirado por não possuir moradores. O questionário foi aplicado em 41 setores e nos 5 setores nos quais não foi aplicado, estimou-se os valores das variáveis com base nos valores dos vizinhos de cada setor.

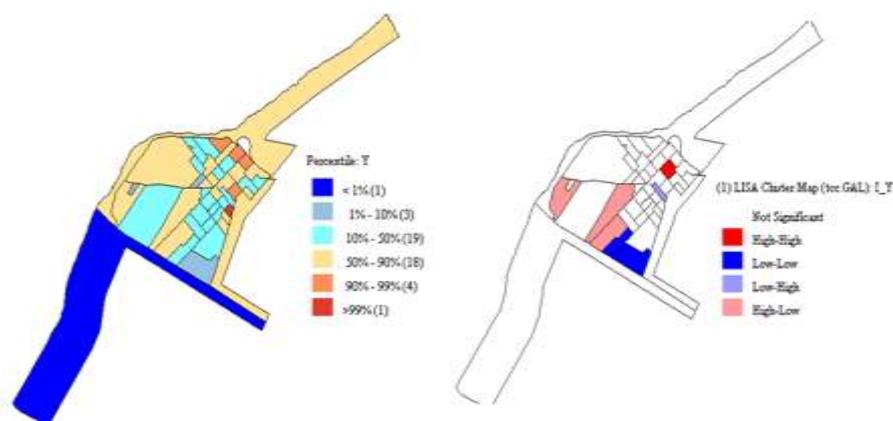
Nas aplicações, a matriz de proximidade espacial  $W$  utilizada é definida segundo o critério de contiguidade de rainha, em que cada elemento da matriz,  $w_{ij}$ , que representa uma medida de proximidade entre os setores  $i$  e  $j$ , recebe valor 1 quando esses setores compartilham campos e arestas e,  $w_{ij}=0$ , caso contrário. Esses valores foram padronizados de modo que a soma dos pesos em cada linha some 1.

As variáveis são referentes à qualidade ambiental, à qualidade de vida e ao lazer e, com pontuações atribuídas a essas variáveis, criou-se a variável resposta ( $Y$ ), denominada de 'Qualidade'. As variáveis explicativas são relacionadas a renda do entrevistado, informações sobre o imóvel que reside, atividades de lazer que realiza e se possui convênio de saúde. Assim, as variáveis foram nomeadas por: renda ( $X_1$ ), alugado ( $X_2$ ), assistir Televisão ( $X_3$ ), nenhuma ( $X_4$ ) e outra ( $X_5$ ), essas duas últimas referentes às atividades realizadas.

Realizou-se uma análise exploratória, Índice de Moran global e LISA, nas variáveis independentes e na variável resposta, para avaliar a hipótese de dependência espacial. Observando as Figuras 1 a 6, nota-se a presença de aglomerados nas variáveis, indicando a existência de dependência espacial.

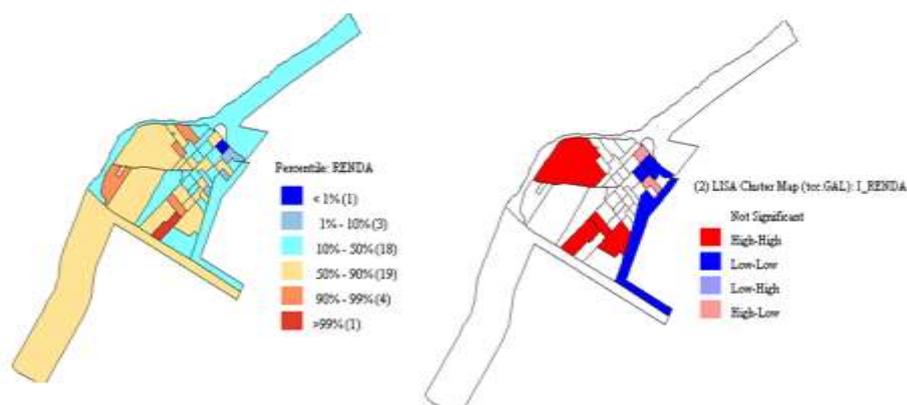
Observa-se que a qualidade  $Y$  é maior mais ao norte da cidade (Figura 1), já a variável renda é maior na região centro-sul (Figura 2). Nota-se também que a maioria dos moradores vivem em casa alugada (Figura 3), em apenas um setor poucos moradores não assistem televisão (Figura 4), em todos os setores mais da metade dos moradores não realizam atividade (Figura 5) e com exceção de um setor, todos os demais possuem moradores que realizam alguma outra atividade (Figura 6). Além disso, os valores dos Índices de Moran das variáveis  $Y$ ,  $X_1$ ,  $X_2$ ,  $X_3$ ,  $X_4$  e  $X_5$  são, respectivamente: 0,0470, 0,1769, -0,011, 0,0646, -0,1310 e -0,0429.

Figura 1 – Mapa da distribuição (direita) e da autocorrelação espacial local (LISA) (esquerda) da Qualidade  $Y$



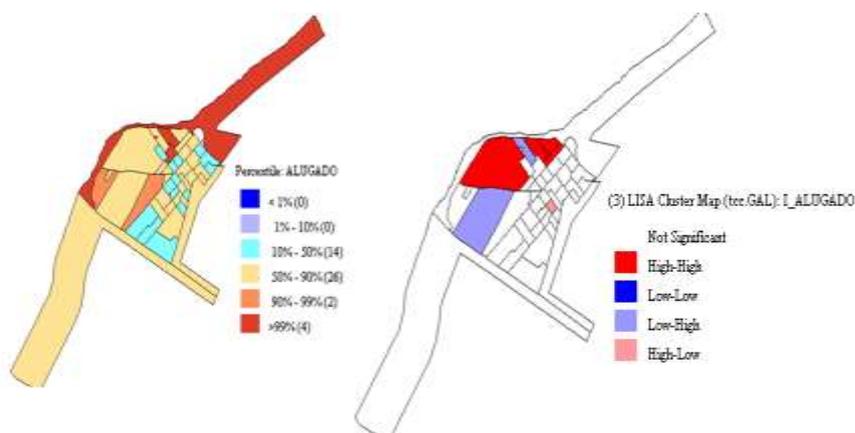
Fonte: Santos (2010).

Figura 2 – Mapa da distribuição (direita) e da autocorrelação espacial local (LISA) (esquerda) da variável renda



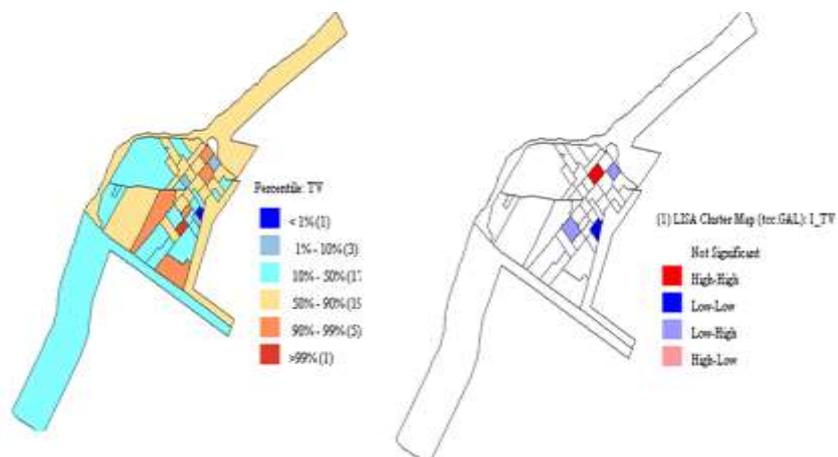
Fonte: Santos (2010).

Figura 3 – Mapa da distribuição (direita) e da autocorrelação espacial local (LISA) (esquerda) da variável imóvel alugado



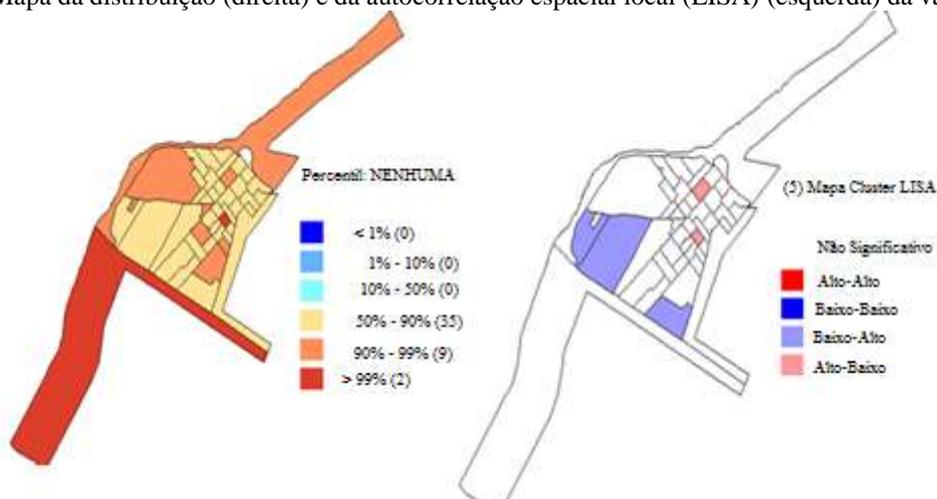
Fonte: Santos (2010).

Figura 4 – Mapa da distribuição (direita) e da autocorrelação espacial local (LISA) (esquerda) da variável Tv



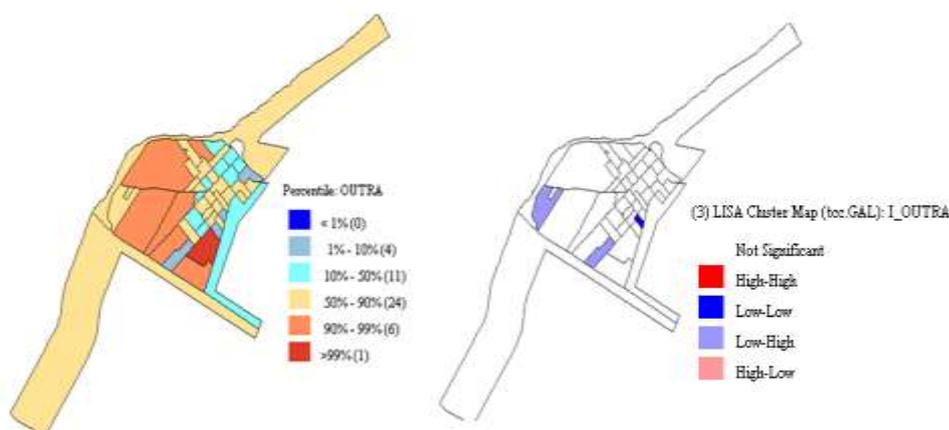
Fonte: Santos (2010).

Figura 5 – Mapa da distribuição (direita) e da autocorrelação espacial local (LISA) (esquerda) da variável nenhuma



Fonte: Santos (2010).

Figura 6 – Mapa da distribuição (direita) e da autocorrelação espacial local (LISA) (esquerda) da variável outra



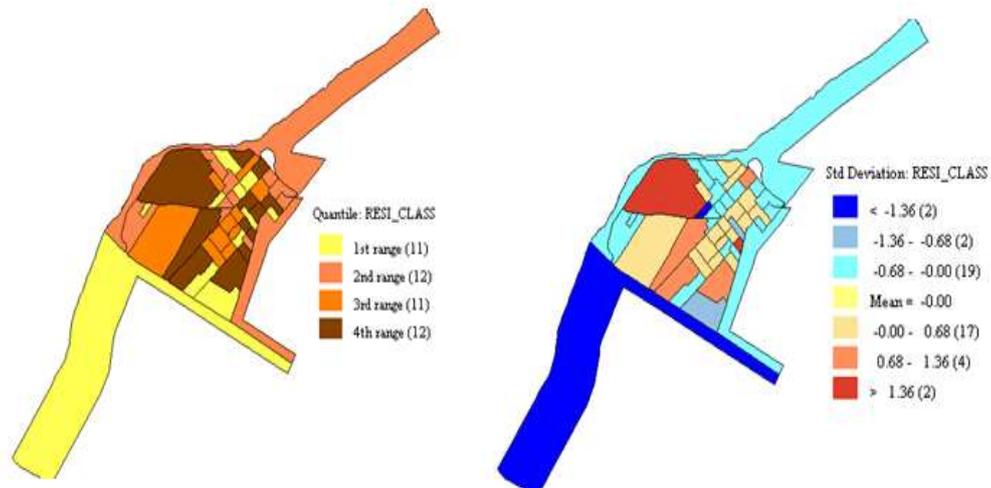
Fonte: Santos (2010).

Ajustou-se um modelo de regressão linear clássica que é dado por:

$$Y=1,402-0,973X_1 +0,153X_2-0,123X_3-0,838X_4-0,119X_5 \quad (1)$$

O coeficiente de determinação do modelo é 0,535, o AIC=105,747 e o SBC=116,719. Os resíduos do modelo (1) não estão distribuídos aleatoriamente pela região de estudo, como se pode observar na Figura 7, que representa a distribuição dos resíduos em quantis e a distribuição dos desvios padrão dos valores absolutos, o que indica a existência de dependência espacial. Desta forma, deve-se seguir com a análise levando em consideração a dependência espacial existente.

Figura 7 – Mapa da distribuição (direita) e do desvio padrão (esquerda) dos resíduos do modelo ajustado em (1)



Fonte: Santos (2010).

Por meio do diagnóstico de autocorrelação espacial sugerido por Anselin (2005), o modelo CAR foi o escolhido como melhor modelo e ficou ajustado da seguinte forma:

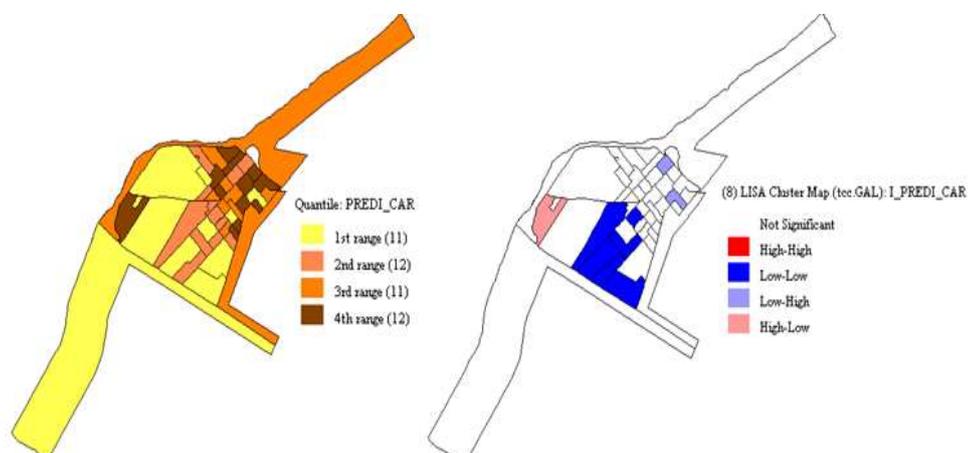
$$Y=11,361-0,953X_1+0,161X_2-0,126X_3-0,804X_4-0,132X_5-0,485W_{res} \quad (2)$$

O valor do pseudocoefficiente de determinação do modelo ajustado foi 0,578, o AIC=102,868 e SBC=113,839. Por essas três medidas, pode-se concluir que o modelo que considera a dependência espacial é o mais adequado.

A estatística I de Moran para os resíduos é -0,0081 que pode ser considerada igual a zero ao nível de significância de 10%, indicando, que a inclusão da variável  $W_{res}$  no modelo, eliminou a autocorrelação espacial que existia.

Na Figura 8 observa-se o mapa da distribuição do ajuste da variável resposta pelo modelo (2), no qual se pode notar que os setores censitários ao centro e ao norte da cidade possuem os maiores valores (cores marrom e alaranjado) da variável Y, o que indica que os moradores situados nessas regiões consideram possuir uma alta qualidade ambiental e de vida.

Figura 8 – Mapa da distribuição (direita) e da autocorrelação espacial local (LISA) (esquerda) para Y



Fonte: Santos (2010).

## 4 Considerações finais

O presente trabalho objetivou analisar a qualidade de vida dos moradores da cidade de Presidente Epitácio, São Paulo, e para tal foram utilizados modelos de regressão linear clássica e espacial. Como os dados apresentaram dependência espacial, o modelo espacial mostrou ser mais adequado, produzindo resultados satisfatórios.

Com o estudo realizado, pode-se observar que a parte mais ao sul da cidade de Presidente Epitácio, São Paulo, é onde aglomera as residências com qualidade de vida mais baixa. As partes mais ao norte e nordeste da cidade apresentam qualidade de vida mais alta. Esse tipo de análise é importante por poder contribuir com aplicações de políticas públicas na cidade para a melhoria da qualidade de vida de seus moradores.

## Agradecimentos

Agradecimentos ao Ricardo dos Santos por disponibilizar os dados utilizados neste trabalho.

## Referências

ANSELIN, L. **Exploring spatial data with GeoDa TM: a workbook**. Champaign: University of Illinois, 2005.

ANSELIN, L. Local Indicators of Spatial Association - LISA. **Geographical Analysis**, v. 27, n. 2, p. 93-115, 1995. 

BARCELLOS, C. C.; SABROZA, P. C.; PEITER, P.; ROJAS, L. I. Organização espacial, saúde e qualidade de vida: análise espacial e uso de indicadores na avaliação de situações de saúde. **Informe Epidemiológico do SUS**, Brasília, v. 11, n. 3, p. 129-138, set. 2002. 

BAILEY, T. C.; GATRELL, A. C. **Interactive spatial data analysis**. New York: J. Wiley, 1995.

CHARNET, R.; FREIRE, C. D. L.; CHARNET, E. M.; BONVINO, H. **Análise de modelos de regressão linear com aplicações**. São Paulo: Unicamp, 1999.

DRUCK, S.; CARVALHO, M. S.; CÂMARA, G.; MONTEIRO, A. M. V. **Análise espacial de dados geográficos**. Brasília: EMBRAPA, 2004.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE. **Censo Demográfico 2010**. Disponível em: <<http://www.censo2010.ibge.gov.br>>. Acesso em: 04 out. 2015.

LIMA, V. M. A.; CALDARELLI, C. E.; CAMARA, M. R. G. Análise do desenvolvimento municipal paranaense: uma abordagem espacial para a década de 2000. **Revista Economia e Desenvolvimento**, v. 26, n. 1, 2014.

MORAN, P. A. P. The interpretation of Statistical Maps. **Journal of the Royal Statistical Society**, Series B, v. 10, n. 2, p. 243-251, 1948.

SANTOS, R. **Meio ambiente e qualidade de vida na estância turística de Presidente Epitácio, São Paulo**. 2010. 374 f. Dissertação (Mestrado em Geografia) – Universidade Estadual Paulista, Presidente Prudente.

SOUZA, N. J. **Desenvolvimento econômico**. São Paulo: Atlas, 1993.

WALL, M. M. A close look at the spatial structure implied by the CAR and SAR models. **Journal of Statistical Planning and Inference**, v. 121, p. 311-324, 2004. 

Recebido em: 25 jun. 2015.  
Aprovado em: 23 jul. 2015.