

AVALIAÇÃO DE DIFERENTES TÉCNICAS ESPACIAIS PARA ANÁLISE DA OCORRÊNCIA DE DENGUE EM RIO CLARO-SP

Jacqueline DOMINGUES¹
José Sílvio GOVONE²

- RESUMO: No presente artigo foi desenvolvido um estudo para a ocorrência de dengue para a área urbana do município de Rio Claro, São Paulo - Brasil. A Fundação Municipal de Saúde do município forneceu dados sobre os casos notificados em 2011. O objetivo principal consistiu em analisar diferentes modelos espaciais para a distribuição da doença no município, dividido por setores censitários, e também, relacionar a doença com fatores socioeconômicos. Dois tipos de modelos espaciais foram aplicados aos dados, sendo eles o modelo SAR – “Simultaneous Autoregressive Models” e SEM – “Simultaneous Error Models”, além da regressão linear clássica para comparação. Os resultados apresentam qual abordagem se mostra mais eficiente e quais fatores socioeconômicos contribuem para ocorrência da doença. Por meio de mapas temáticos identificou-se quais áreas possuem maior concentração da doença. As análises foram realizadas no *software* R.
- PALAVRAS-CHAVE: Dengue; epidemiologia espacial; simultaneous autoregressive models (SAR); simultaneous error models (SEM); análise espacial.

1 Introdução

A dengue é atualmente a arbovirose mais importante que afeta o ser humano, sendo um grave problema de saúde pública no mundo. A dengue é transmitida pelo mosquito *Aedes aegypti*, o qual prolifera-se em recipientes que acumulam água limpa, e a principal forma de transmissão é pela picada do vector. Existem quatro tipos diferentes de vírus da dengue: DENV 1, DENV 2, DENV 3 e DENV 4, de acordo com o Ministério da Saúde (2016).

¹Universidade Estadual Paulista - UNESP, Instituto de Biociências, Departamento de Bioestatística, CEP: 18618-689, Botucatu, SP, Brasil. E-mail: jacquelinedomingues.unesp@gmail.com

²Universidade Estadual Paulista - UNESP, Instituto de Geociências e Ciências Exatas, Departamento de Estatística, Matemática Aplicada e Computação, CEP: 13506-900, Rio Claro, SP, Brasil. E-mail: jsgovone@gmail.com

Através dos dados obtidos pela Fundação Municipal de Saúde de Rio Claro-SP, buscou-se identificar o comportamento espacial da doença por setores censitários, onde optou-se pelo trabalho somente com as áreas urbanas devido a baixa incidência em áreas rurais.

Recentemente a estatística e a epidemiologia espacial tornaram-se ferramentas eficientes para auxiliar a identificar possíveis epidemias. Segundo Kuhn *et al.* (2005) a identificação precoce de uma epidemia de doenças infecciosas é um importante primeiro passo para a implementação de intervenções eficazes para controlar a doença e reduzir a mortalidade e morbidade na população. Entretanto, a disseminação das doenças está em estágio avançado quando as autoridades locais são notificadas e, como consequência, as medidas de controle são implementadas tardiamente.

Em virtude do que foi mencionado, a consideração do espaço para a análise de ocorrência de dengue pode ser uma ferramenta útil para identificação de áreas de maior ocorrência, visto que os resultados podem ser apresentados por mapas temáticos.

Usualmente, quando se busca analisar a relação entre uma variável de interesse com um conjunto de variáveis explicativas, aplica-se o modelo de regressão linear clássico, sendo esta uma estrutura básica para muitos modelos espaciais. A estrutura de variância e covariância de um modelo clássico assume que os dados sejam não correlacionados; em consequência disso, se as observações estiverem espacialmente distribuídas, pelas suposições, os valores observados em determinado local deverão ser independentes das observações em locais próximos, sendo que, em contextos espaciais, essas suposições são extremamente fortes. Como alternativa a abordagem clássica, foram aplicados aos dados dois modelos espaciais, os modelos SAR – “Simultaneous Autoregressive Models” e SEM – “Simultaneous Error Models”.

2 Materiais e métodos

A análise foi realizada no município de Rio Claro, localizado a 22°24'39”S e 47°33'39”O no interior do estado de São Paulo - Brasil; em 2010 possuía 186.253 habitantes e uma área territorial de 498,422km², dividido em 197 setores censitários; neste caso, foram consideradas apenas as áreas urbanas, segundo o Censo Demográfico de 2010 (IBGE, 2016).

Obteve-se, em 2011, 1754 casos notificados de dengue no município, porém, devido a falta de informações, como endereços incompletos, foi possível georreferenciar apenas 1324, os quais foram considerados. O georreferenciamento deu-se por endereço dos pacientes, fornecidos no momento da procura por atendimento médico. Utilizou-se o aplicativo TerraView (2014). As covariáveis escolhidas para compor, inicialmente, os modelos foram obtidas pelo Censo 2010.

O estudo consistiu em analisar a doença, por setores censitários (áreas urbanas), no município de Rio Claro. Para cada setor foi atribuído o valor da

variável resposta (casos notificados de dengue) $Y = \{y_1, \dots, y_{197}\}$ e também um conjunto de covariáveis $X = \{X_1, \dots, X_6\}$, definidas conforme tabela 1.

Tabela 1 - Covariáveis (Censo demográfico, 2010)

X_i	Descrição de acordo com IBGE
X_1	Média de moradores por domicílio.
X_2	Valor do rendimento nominal médio mensal das pessoas responsáveis por domicílios particulares permanentes (com e sem rendimento).
X_3	Setores que possuem abastecimento de água diferente do convencional: abastecimento por poço, nascente, cisterna ou outra forma de abastecimento.
X_4	Setores que possuem descarte de lixo diferente do convencional: lixo queimado na propriedade, enterrado na propriedade, jogado em terreno baldio ou logradouro, jogado em rio, lago, mar, ou com outro destino.
X_5	Número total de moradores em domicílios particulares permanentes ou população residente em domicílios particulares permanentes.
X_6	Número total de analfabetos responsáveis pela família, residentes no setor censitário.

Dessa forma, pode-se classificar os dados como de área, em que $S = \{s_1, \dots, s_{197}\}$ representa os setores censitários com $s_i \cap s_j = \emptyset$ e $\bigcup s_i = S$ para quaisquer $i \neq j$, $i, j = \{1, \dots, 197\}$.

Para avaliar a possível presença de autocorrelação espacial nos dados, utilizou-se o índice de Moran. A autocorrelação espacial busca analisar quanto o valor observado de determinada variável em uma área está associado com o valor observado desta mesma variável em áreas consideradas vizinhas.

O índice global de Moran, apresentado em Lloyd (2011) e Bailey & Gratell (1995), é dado por:

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n w_{ij} (y_i - \bar{y})(y_j - \bar{y})}{\left(\sum_{i=1}^n (y_i - \bar{y})^2 \right) \left(\sum_{i \neq j} w_{ij} \right)}. \quad (1)$$

y_i : representa o valor da variável Y na área s_i ; \bar{y} : média da variável Y ; w_{ij} : refere-se ao elemento na linha i e coluna j da matriz de vizinhança $W_{n \times n}$ e n o número de setores.

Uma vez calculado, é importante estabelecer sua significância estatística. Uma possibilidade para testar o resultado encontrado é um teste de pseudo-significância. Neste caso são geradas diferentes permutações dos valores observados associados as áreas e recalculam-se as estatísticas. Esse processo é repetido por N vezes, N suficientemente grande, gerando assim uma distribuição de referência. Se o valor

do índice I calculado originalmente corresponder a um extremo da distribuição simulada através das permutações, então o valor de I possui significância estatística (ANSELIN, 2005).

A matriz de vizinhança $W_{n \times n}$, ou matriz de adjacências, busca expressar a estrutura espacial dos dados. Cada um dos elementos w_{ij} representa uma medida de proximidade/distância entre s_i e s_j . Para esta análise foi considerado uma matriz de vizinhança $W_{197 \times 197}$, padronizada, utilizando o critério de adjacência *queen*, que considera como vizinhas as unidades que possuem fronteiras ou vértices comuns, em que a unidade vizinha é definida da forma $w_{ij} = 1$, caso contrário $w_{ij} = 0$, para vizinhos de primeira ordem¹. A figura 1 apresenta o município de Rio Claro, seus 197 setores censitários e sua estrutura de vizinhança.

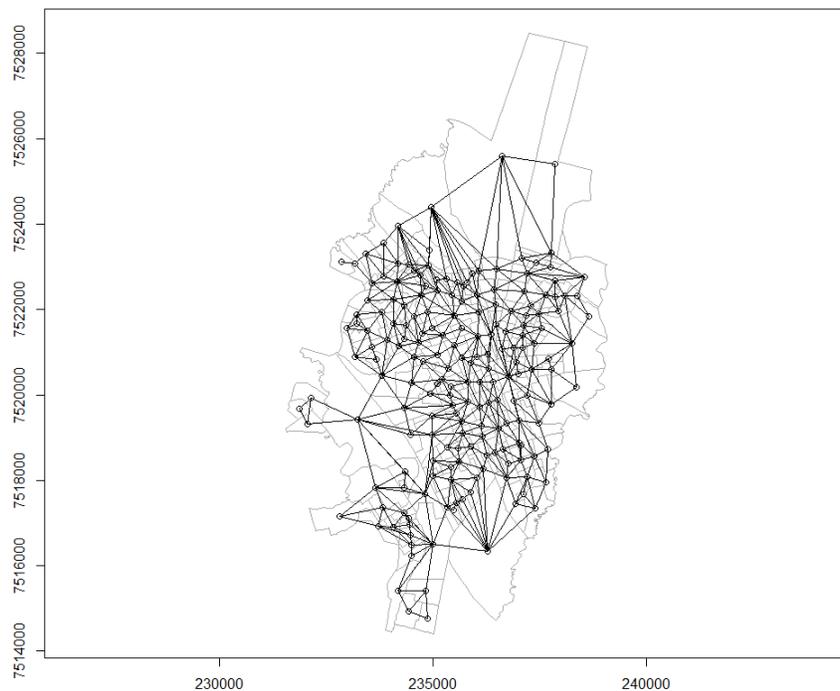


Figura 1 - Vizinhança de primeira ordem, critério de adjacência *queen*. Área urbana de Rio Claro-SP, dividida em 197 setores censitários.

Definida a estrutura espacial, os modelos utilizados buscaram incorporar a estrutura espacial presente nos dados, sendo eles os modelos SAR – “Simultaneous Autoregressive Models” e SEM – “Simultaneous Error Models”.

¹Vizinhança de primeira ordem considera como vizinhos apenas as áreas que possuem contato direto. Vizinhança de ordem dois, por exemplo, considera vizinho os vizinhos de seu vizinhos. Exemplo em Domingues (2017).

A começar pelo modelo SAR, neste a autocorrelação espacial é atribuída à variável resposta Y . Seu uso é apropriado quando os valores observados de Y em uma área s_i são diretamente influenciados pelos valores de Y observados em seus vizinhos. O modelo é definido, de acordo com Anselin (2005), por:

$$Y = \rho WY + X\beta + \varepsilon, \quad (2)$$

em que $Y_{n \times 1}$: número de ocorrências de dengue em cada setor censitário; $X_{n \times k}$: variáveis explicativas (Tabela 1); $\beta_{k \times 1}$: coeficientes de regressão; $\varepsilon_{n \times 1}$: erros aleatórios com média zero e variância $\sigma^2 I$; I matriz identidade; $W_{n \times n}$: matriz de vizinhança espacial; ρ : coeficiente espacial autoregressivo com $E(y_i y_j) \neq 0$, indicando que a observação y na área s_i está correlacionada com a observação y na área s_j ; n é o número de setores e k o número de parâmetros do modelo (número de covariáveis+1).

Se o parâmetro ρ for significativo, então parte da variação total de Y é explicada pela dependência de cada observação dos seus vizinhos. O parâmetro ρ mensura o efeito médio da variável dependente Y , em relação a sua vizinhança espacial.

Se Y não for influenciado diretamente pelo seus vizinhos, e existir algum agrupamento espacial que influencie o seu valor em s_i e seus vizinhos, porém esta característica é omitida ou não observável, convém considerar um modelo alternativo com erros espacialmente correlacionados, como o modelo SEM (WARD e GLEDITSCH, 2008).

Em relação ao modelo SEM, este considera que os efeitos espaciais são um ruído, ou perturbação, ou seja, fator que precisa ser removido (CÂMARA *et al.*, 2004). Neste caso, os efeitos da autocorrelação espacial são associados ao termo de erro ε . Bailey e Gratell (1995) o define como:

$$\begin{aligned} Y &= X\beta + \varepsilon, \\ \varepsilon &= \lambda W\varepsilon + U, \\ Y &= X\beta + (I - \lambda W)^{-1}U, \end{aligned} \quad (3)$$

em que $Y_{n \times 1}$: número de ocorrências de dengue em cada setor censitário; $X_{n \times k}$: variáveis explicativas (Tabela 1); $\beta_{k \times 1}$: coeficientes de regressão; $W\varepsilon_{n \times 1}$: erro com efeito espacial; $U_{n \times 1}$: erros aleatórios com média zero e variância $\sigma^2 I$; λ : coeficiente autoregressivo; $E(U) = 0$, $E(UU^T) = \sigma^2 I$, $E(\varepsilon_i \varepsilon_j) \neq 0$, indicando que o erro atribuído à área s_i está correlacionado com o erro na área s_j .

Os estimadores dos parâmetros, para ambos, são obtidos a partir da maximização da função logaritmo da verossimilhança.

Para a escolha do modelo mais adequado, os resíduos dos modelos ajustados foram analisados, visando identificar se há quebra de suas suposições, como presença de autocorrelação espacial. Utilizou-se também, as estatísticas do teste do multiplicador de Lagrange (LM)² que testam a hipótese de ausência

²O teste do Multiplicador de Lagrange, foi utilizado para comparar o Modelo de Regressão Linear Clássico com os modelos SAR e SEM (Anselin, 2005).

de autocorrelação espacial devido a uma estrutura SAR ou a SEM. O teste busca indicar qual modelo deve ser considerado; para isso segue-se um algoritmo de decisões, quatro estatísticas diferentes para o multiplicador de Lagrange são calculadas: LM-SAR e LM-SAR Robusto tomam o modelo SAR como alternativo. LM-SEM e LM-SEM Robusto referem-se ao modelo SEM como a alternativo. Vale ressaltar que os testes LM Robustos devem ser considerados quando as versões padrão (LM-SAR ou LM-SEM) são significativas. Quando não o forem, as propriedades das versões robustas podem não ser válidas. A rejeição da hipótese nula por ambas as estatísticas é uma situação comumente encontrada na prática, podendo então considerar as formas robustas das estatísticas (Anselin, 2005). Por fim, observou-se os valores do Critério de Informação de Akaike (AIC) e do Logaritmo da Verossimilhança (LIK) para a escolha do modelo mais adequado.

Os ajustes foram realizados no *software* R (R CORE TEAM, 2018) através das funções *lm*, *lagsarlm* e *errorsarlm*.

3 Resultados e discussão

3.1 Análise exploratória

Primeiramente analisou-se a variável resposta Y para verificar se esta segue as pressuposições necessárias, como normalidade. Observou-se alta assimetria na variável original Y . Assim, aplicou-se a transformação $Y^* = \ln(Y + 1)$, a qual diminuiu consideravelmente a assimetria presente, como mostra a figura 2. Esta transformação também adequa o ajuste para que não sejam estimados valores negativos para a ocorrência de dengue. A variável Y^* foi considerada para o estudo.

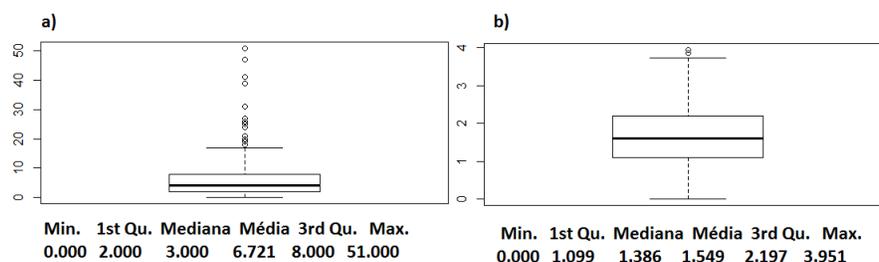


Figura 2 - Boxplot e medidas descritivas. (a) Variável Y original e (b) Variável Y^* .

Em relação as covariáveis X_i , $i = \{1, \dots, 6\}$, observou-se (figura 3): que os setores censitários possuem, em sua maioria, 2 a 4 moradores por residência; que existe discrepância na distribuição de renda entre os setores, alguns de moradores com alta renda enquanto outros, de baixa renda; sobre abastecimento de água e coleta de lixo, estas covariáveis foram categorizadas como 1 para setores que

possuem coleta de lixo diferente do convencional, 0 caso contrário, e 1 para setores que possuem abastecimento de água diferente do convencional, 0 caso contrário. Estas covariáveis não apresentaram indicativos de diferença, pois o município tem boa cobertura destes serviços e os setores que apresentam abastecimento/coleta diferentes do convencional possuem poucas residências com este comportamento. A distribuição da população indica setores com diferentes densidades populacionais e em relação ao número de analfabetos responsáveis pela família, existe discrepância desta quantidade entre os setores.

Analisando a figura 3 em relação a ocorrência de dengue, os domicílios com maior média de moradores, de três até quatro, tendem a apresentar maior ocorrência; valores de renda, coleta de lixo, abastecimento de água e população residente não apresentam grandes relações e por fim, as famílias, cujos responsáveis são analfabetos apresentam indicativos de que quanto maior o número de analfabetos que o setor possui, maior a ocorrência de dengue. As covariáveis foram escolhidas com intuito de descrever as diferenças sociais e econômicas de cada setor. Levando-se em consideração esses aspectos optou-se por trabalhar somente com duas covariáveis, sendo elas: X_1 : Média de moradores por domicílio e X_6 : número total de analfabetos responsáveis pela família, residentes no setor censitário.

Definida a variável resposta e suas covariáveis X_1 e X_6 , conforme tabela 1, prosseguiu-se a análise em busca de identificar, através do índice de Moran, a existência, ou não, de autocorrelação nos dados. Para a variável resposta Y^* observou-se um índice $I_{Y^*} = 0,3343$ com valor $p < 0,00001$ para 10000 simulações (figura 4). Como o valor verdadeiro do índice corresponde a um extremo da distribuição simulada pode-se aceitar a hipótese de autocorrelação nos dados.

Após esta primeira análise prosseguiu-se com os ajustes dos modelos, a começar pela regressão linear clássica.

3.2 Regressão linear clássica

Como já mencionado, ao considerar o modelo de regressão linear clássico, suposições sobre os erros devem ser válidas, sendo elas, $\varepsilon \sim N(0, \sigma^2)$ e $Cov(\varepsilon_i, \varepsilon_j) = 0$ para quaisquer $i \neq j$. O modelo é definido por:

$$Y^* = X\beta + \varepsilon, \quad (4)$$

em que $Y_{197 \times 1}^* = \{y_1^*, \dots, y_{197}^*\}$ casos de dengue em 2011 transformados por $\ln(Y+1)$; $\beta_{3 \times 1}$ os coeficientes de regressão e $X_{197 \times 3}$ é a matriz de incidência dos efeitos das covariáveis. A tabela 2 apresenta as estimativas dos parâmetros do modelo ajustado.

Para testar a normalidade para os termos do erro utilizou-se o teste de Jarque-Bera (ANSELIN, 2005), que tem como hipótese nula a normalidade; o valor obtido $p < 0,00001$, rejeita a hipótese, indicando que os erros não seguem uma distribuição normal. Outro fato importante observado é a existência de autocorrelação espacial nos resíduos; o índice de Moran calculado para os resíduos foi $I_{OLS} = 0,2301$ com

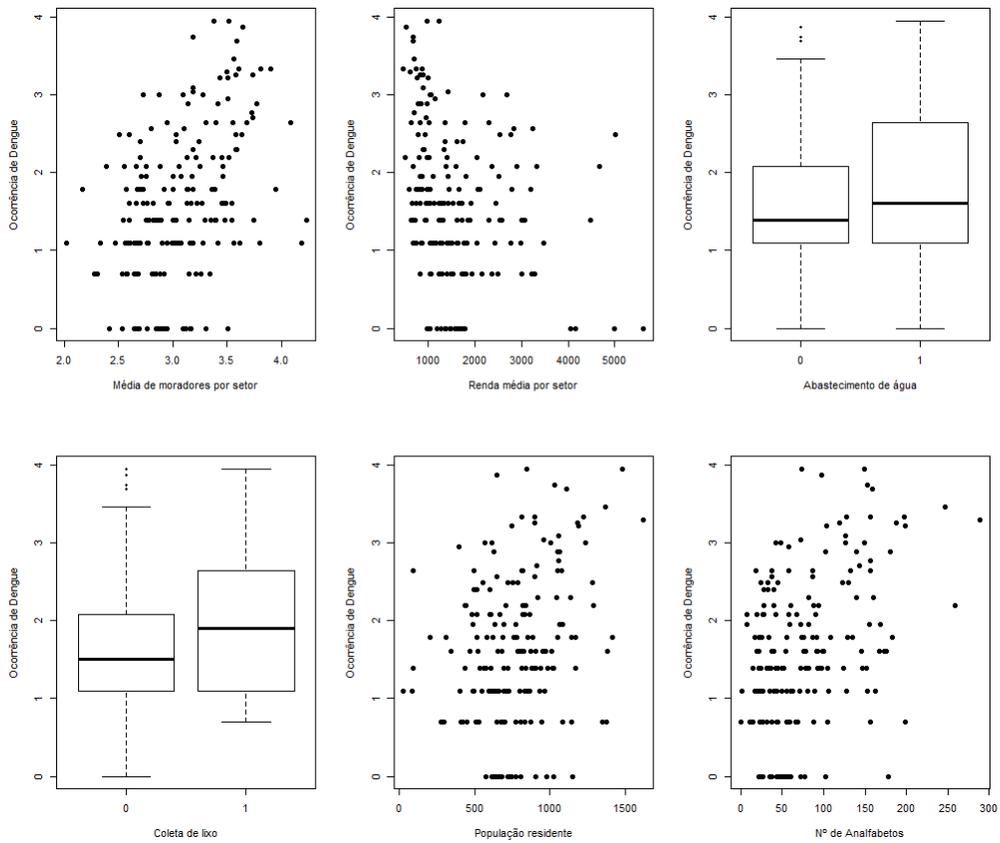


Figura 3 - Dengue em relação às covariáveis X_i , $i = \{1, \dots, 6\}$.

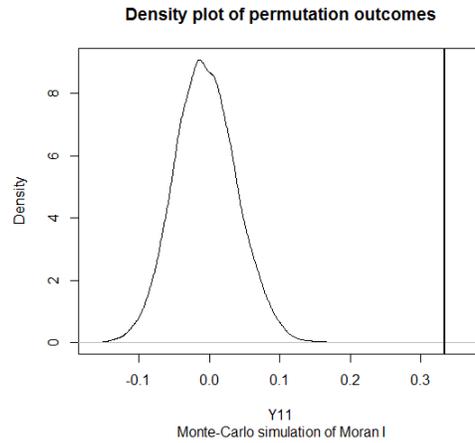


Figura 4 - Índice de Moran para ocorrência de dengue por setor censitário.

Tabela 2 - Modelo clássico ajustado

	Estimativa	Erro padrão	t	$\Pr(> t)$
Intercepto	-0,6107	0,5599	-1,09	0,2767
X_1	0,6129	0,2036	3,01	0,0030
X_6	0,0043	0,0015	2,80	0,0056
LIK	-237,8012			
AIC	491,6025			

valor $p < 0,00001$ (figura 5) indicando presença de autocorrelação espacial, ou seja, necessita-se considerar modelos espaciais para o ajuste. O teste do multiplicador de Lagrange indicou que o modelo SAR como a alternativa mais apropriada, como mostra a tabela 3. Porém, com intuito de comparar diferentes modelos, o modelo SEM também foi ajustado.

Tabela 3 - Multiplicadores de Lagrange

Diagnóstico para dependência espacial				
Modelo	LM	valor p	LM_Robusto	valor p
SAR	32,2150	<0,0001	5,9724	0,01453
SEM	26,2820	<0,0001	0,0394	0,84270

3.3 SAR – “Simultaneous Autoregressive Models”

O modelo SAR atribui a autocorrelação espacial à variável resposta Y^* , definido por:

$$Y^* = \rho WY^* + X\beta + \varepsilon, \quad (5)$$

em que $W_{197 \times 197}$ matriz de vizinhança baseada no critério de contingência *queen* para vizinhos de primeira ordem; ρ : coeficiente espacial autoregressivo; as outras variáveis como definidas para o modelo anterior. A tabela 4 apresenta as estimativas dos parâmetros do modelo ajustado.

Tabela 4 - Modelo SAR ajustado

	Estimativa	Erro padrão	z	$\Pr(> z)$
Intercepto	-0,9723	0,5129	-1,90	0,0580
X_1	0,5440	0,1866	2,92	0,0035
X_6	0,0027	0,0014	1,89	0,0483
ρ	0,4347			< 0,0000
LIK	-226,1492			
AIC	462,3000			

O índice de Moran, aplicado aos resíduos deste modelo, resultou em $I_{SAR} = -0,0152$ com valor $p = 0,5860$, ou seja, a correlação espacial é não significativa (figura 5).

3.4 SEM – “Simultaneous Error Models”

Este modelo de regressão espacial considera que os efeitos da autocorrelação espacial estejam associados ao termo de erro ε e é definido por:

$$Y^* = X\beta + \varepsilon, \quad (6)$$

$$\varepsilon = \lambda W\varepsilon + U,$$

em que λ : coeficiente espacial autoregressivo; $U_{197 \times 1}$ os erros aleatórios com média zero e variância $\sigma^2 I$ e as outras variáveis como definidas anteriormente. A tabela 5 apresenta as estimativas dos parâmetros do modelo ajustado.

Tabela 5 - Modelo SEM ajustado

	Estimativa	Erro padrão	z	$\Pr(> z)$
Intercepto	-0,4869	0,5316	-0,92	0,3597
X_1	0,6038	0,1861	3,24	0,0012
X_6	0,0029	0,0014	2,01	0,0449
λ	0,4671			< 0,0000
LIK	-227,3763			
AIC	464,7500			

Os resíduos não apresentam autocorrelação espacial, o índice de Moran resultou em $I_{SEM} = -0,0129$ com valor $p = 0,5770$ (figura 5).

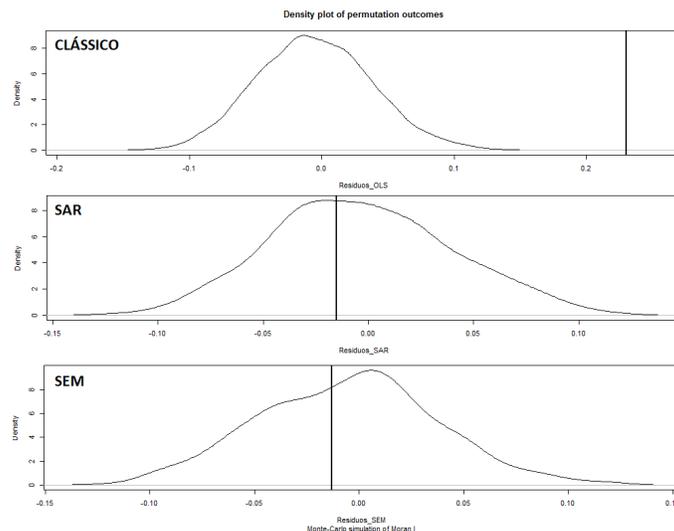


Figura 5 - Teste de permutação aleatória para os resíduos, para 10000 permutações, considerando os três modelos analisados.

Através da aplicação dos diferentes modelos, identificou-se qual melhor se adapta aos dados espacialmente distribuídos e quais variáveis socioeconômicas estão relacionadas com a incidência de dengue. As covariáveis X_1 e X_6 mostraram-se significativas, ou seja, existe relação entre a ocorrência de dengue e a média de moradores por residência, bem como o número total de analfabetos responsáveis pela família, residentes no setor censitário. Claramente essas covariáveis são relacionadas a fatores sociais e econômicos, descrevendo que setores mais carentes podem apresentar maior ocorrência de dengue.

Os modelos com efeitos globais que incorporam estrutura espacial foram satisfatórios, e se mostraram mais eficientes quando comparados à regressão linear clássica, indicando que o acréscimo do parâmetro espacial auxilia significativamente na explicação da dengue. O modelo que melhor se ajustou aos dados foi o modelo SAR, como indicado pelo multiplicador de Lagrange (Tabela 3). Além disso, analisando os valores de AIC e LIK obtidos para os modelos, o modelo SAR possui o menor AIC dentre os três e maior valor para o LIK.

O modelo SAR ajustado é dado por:

$$\widehat{Y}^* = -0,9723 + 0,5440X_1 + 0,0027X_6 + 0,43472WY^* \quad (7)$$

em que W é a matriz de adjacência *queen* para vizinhos de primeira ordem e Y^* os

casos de dengue transformados.

As covariáveis apresentaram valores positivos nos coeficientes de regressão, mas esta interpretação não é direta; a variável resposta Y foi transformada. Para interpretar as estimativas deve-se fazer a transformação inversa; assim pode-se concluir sobre o número de casos de dengue estimados para cada setor além da ponderação espacial $\hat{\rho} \approx 0.43$.

Os mapas (figura 6) apresentam a variável Y^* distribuída espacialmente (à esquerda) e Y_{SAR}^* estimado pelo modelo SAR (à direita). Nota-se que Y_{SAR}^* , suaviza a superfície; esta suavização auxilia na identificação de *clusters* que indicam as áreas de maior incidência da doença, podendo ser uma ferramenta útil para a vigilância sanitária distribuir suas equipes em áreas que necessitam de maior atenção.

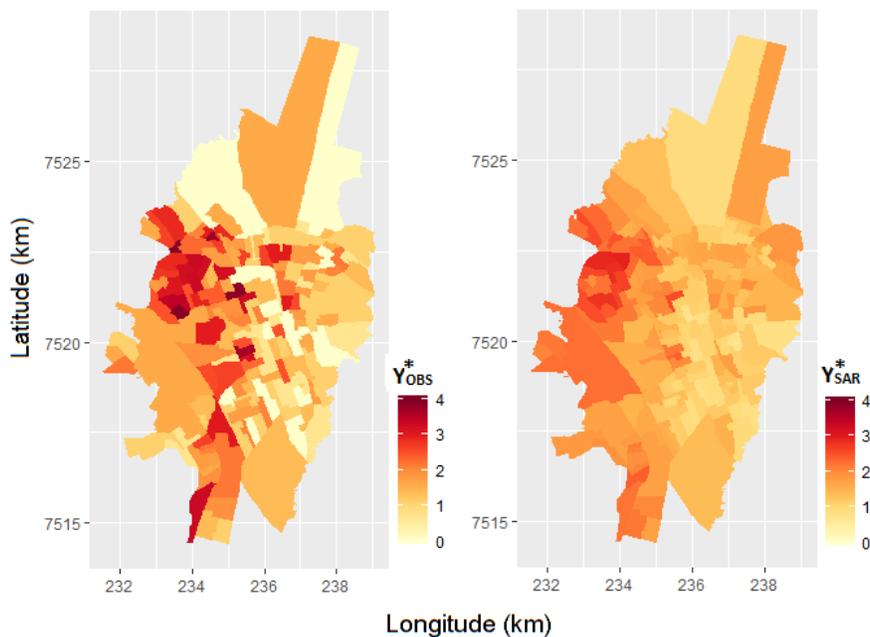


Figura 6 - À esquerda: Ocorrência de dengue (transformada, Y^*) em Rio Claro por setores censitários. À direita: Estimativas para ocorrência de dengue (Y_{SAR}^*) em Rio Claro por setores censitários através do ajuste do modelo SAR.

O modelo SAR indicou a região noroeste da cidade de Rio Claro – SP como a de maior incidência. A figura 7 apresenta os bairros que estão na região ou em áreas próximas.

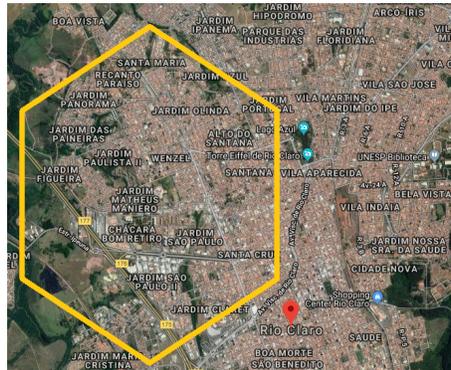


Figura 7 - Mapa da cidade de Rio Claro-SP com destaque para a área com maior estimativa para ocorrência de dengue. Fonte: google maps.

Outra preocupação são os resíduos do modelo ajustado (SAR), os quais não devem apresentar autocorrelação espacial. Como visto anteriormente o índice de Moran resultou em $I_{SAR} = -0,0152$ com valor $p = 0,5860$. Pode-se concluir que os resíduos se distribuem aleatoriamente pela região de estudo. A figura 8 apresenta a distribuição espacial dos resíduos por setores censitários, índice de Moran e teste de permutação aleatória.

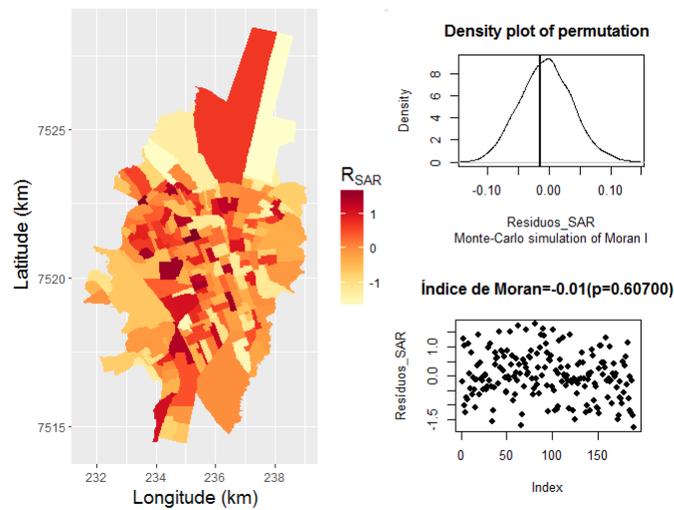


Figura 8 - Distribuição espacial dos resíduos do modelo SAR, por setores censitários.

4 Conclusões

Através das aplicações, identificou-se regiões com maior incidência de casos de dengue e com quais covariáveis esta incidência se relaciona. Com intuito de verificar a autocorrelação espacial, aplicou-se o índice de Moran, o qual indicou autocorrelação entre os setores censitários, detectando agrupamentos.

Ajustou-se três modelos de regressão: modelo linear clássico, SAR e SEM, buscando relacionar a dengue com as covariáveis socioeconômicas obtidas pelo Censo 2010. No modelo clássico os resultados obtidos foram considerados insatisfatórios devido a quebra das suposições usuais, além de não incorporar em sua estrutura a dependência espacial presente. Em virtude do que foi mencionado, os modelos com efeitos espaciais globais ajustados se mostraram mais eficientes quando comparados ao clássico.

Ao comparar os modelos alternativos com a modelagem clássica, pelos métodos dos multiplicadores de Lagrange, os melhores resultados foram obtidos pelos modelos que incorporaram a dependência espacial, indicando que o acréscimo do parâmetro autorregressivo auxilia significativamente na explicação da ocorrência de dengue, como em Domingues (2017).

O modelo que melhor se ajustou aos dados foi o SAR, indicando, através de um mapa suavizado, a região noroeste da cidade de Rio Claro – SP como a de maior ocorrência de dengue. Os mapas geralmente são de fácil interpretação, o que pode ser uma ferramenta útil para as autoridades de saúde responsáveis, que podem tratar a área em destaque com maior ênfase e auxiliar nas campanhas de prevenção para anos futuros.

As técnicas espaciais vêm se tornando uma ferramenta considerável para estudos epidemiológicos, visto que fornecem uma visão da distribuição/localização espacial das doenças. Barcellos et al. (2006), afirmam que quando espaço e saúde são relacionados, permite-se a formulação de novas hipóteses e a criação de indicadores, possibilitando a simulação de diversas situações possíveis na busca dos principais determinantes da doença em estudo.

Para o estudo de doenças, existem modelos espaciais que buscam inserir informações sobre o local de ocorrência através de distribuições *a priori* (abordagem Bayesiana). Estas distribuições (LEE, 2011) tem por objetivo aumentar o valor da densidade para áreas consideradas próximas/vizinhas. Os Modelos Bayesianos vêm se destacando e apresentam resultados satisfatórios, sendo um campo de estudo com grande potencial.

Muito se discute, também, sobre a importância de incluir, para estudos de dengue, fatores climáticos, como pode-se ver em Lowe *et al.* (2011, 2014). Recomenda-se, para futuros trabalhos, o estudo da distribuição espacial ou espaço-temporal da dengue considerando métodos Bayesianos para a análise.

5 Agradecimentos

À fundação Municipal de Saúde do município de Rio Claro, SP pelo fornecimento de dados. À CAPES pelo apoio financeiro. Ao programa de pós graduação em Biometria/IBB/UNESP campus de Botucatu - SP por possibilitar a realização do presente trabalho. Aos revisores e editores pelas sugestões.

DOMINGUES, J.; GOVONE, J. S. Evaluation of different spatial techniques to analyze the occurrence of dengue in Rio Claro - SP. *Rev. Bras. Biom.*, Lavras, v.37, n.1, p.1-16, 2019.

■ **ABSTRACT:** *In this paper we have developed a study of dengue fever in Rio Claro, São Paulo - Brazil. The Municipal Health Foundation of Rio Claro provided data about reported cases of dengue fever in 2011. The main objective was to analyze both the spatial distribution of the disease in the city, by Census tracts, and the relationship of the disease with socioeconomic factors. Two types of spatial models were applied to the data: the SAR model "Simultaneous Autoregressive Models" and SEM - "Simultaneous Error Models". We also fitted the classic linear model, just for comparison to the two spatial models. The results showed that dengue is related to socioeconomic factors and, through the models it was possible to identify which one was statistically significant. Thematic maps have identified the areas that have the highest concentration of the disease.*

■ **KEYWORDS:** *Dengue; epidemiology; simultaneous autoregressive models (SAR); simultaneous error models (SEM); spatial analysis.*

Referências

ANSELIN, L. *Exploring spatial data with GeoDaTM : A workbook for spatial analysis*. Laboratory Department of Geography - University of Illinois, Urbana-Champaign Urbana, IL 61801. Revised Version, March 6, 2005.

BAILEY, C. T.; GATRELL, C. A. *Interactive spatial data analysis*. London: Longman Scientific & Technical, 1995. 413p.

BARCELLOS, C.; SANTOS, S. M. Brasil. Ministério da Saúde. Secretaria de Vigilância em Saúde. Fundação Oswaldo Cruz. *Abordagens espaciais na saúde pública* / Ministério da Saúde, Fundação Oswaldo Cruz; Brasília : Ministério da Saúde. (Série B. Textos Básicos de Saúde) (Série Capacitação e Atualização em Geoprocessamento em Saúde; 1) ISBN 85-334-1181-2, 2006.

CÂMARA, G.; CARVALHO, M. S.; CRUZ, O. G.; CORREA, V. Análise espacial de áreas. *Embrapa Cerrados. Análise espacial de dados geográficos*, Planaltina: Embrapa; p.157-209, 2004.

DOMINGUES, J. *Técnicas de processos espaciais e espaço-temporais com aplicações*

em dados de dengue. 2017. 77f. Dissertação (Mestrado em Biometria), - Instituto de Biociências, Universidade Estadual Paulista, Botucatu, 2017.

IBGE - Instituto Brasileiro de Geografia e Estatística. *Censo demográfico de 2010* - BRASIL. <https://www.ibge.gov.br/>

KUHN, K.; CAMPBELL-LENDRUM, D.; HAINES, A.; COX, J.; CORVOLAN, C.; ANKER, M. *Using climate to predict infectious disease epidemics*. Geneva: World Health Organization. 55pp. 2005.

LEE, D. A comparison of conditional autoregressive models used in Bayesian disease mapping. *Spatial and Spatio-temporal Epidemiology*, v.2, n.2, p.79-89, 2011.

LLOYD, C. D. *Local models for spatial analysis*. London: CRC Press , 2.ed., 2011. 336p.

LOWE, R.; BAILEY, T. C.; STEPHENSON, D. B.; GRAHAM R. J.; COELHO, C. A. S., SÁ CARVALHO, M. S.; BARCELLOS, C. Spatio-temporal modelling of climate-sensitive disease risk: Towards an early warning system for dengue in Brazil. *Computers & Geosciences*, v.37, n.3, p.371-381, 2011.

LOWE, R.; BARCELLOS C.; COELHO, C.; BAILEY, T. C.; COELHO, G. E.; GRAHAM, R.; JUPP, T.; RAMALHO, W. M.; SÁ CARVALHO, M. S.; STEPHENSON, D. B.; RODÓ, X. Dengue outlook for the World Cup in Brazil: an early warning model framework driven by real-time seasonal climate forecasts. *The Lancet, Infections Diseases*, v.14, n. 7, p.619-626. 2014.

MINISTÉRIO DA SAÚDE. www.saude.gov.br, 2016.

R CORE TEAM. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>, 2018.

TERRAVIEW 4.1.0. São José dos Campos, SP: INPE, 2010. Disponível em: www.dpi.inpe.br/terraview. Acesso em: 11/02/2014.

WARD, M. D; GLEDITSCH, K. S. *Spatial regression models*. London: SAGE Publications, v.155, 2008. 112p.

Recebido em 22.09.2017.

Aprovado após revisão em 02.05.2018.