



Ministério da
**Ciência, Tecnologia
e Inovação**



ANÁLISE ESPACIAL DO TIPO DE COLHEITA DA CANA-DE-AÇÚCAR NO ESTADO DE SÃO PAULO

Bruno Silva Oliveira

Relatório Final de Projeto da disciplina
SER301 – Análise Espacial de Dados
Geográficos, sob responsabilidade do Dr.
Eduardo G. Camargo e do Dr. Antônio
Miguel Vieira Monteiro.

1. Introdução

A cana-de-açúcar é uma cultura agrícola de grande importância econômica no Brasil, principalmente devido à crescente demanda de etanol, decorrente tanto da mistura do etanol na gasolina quanto da popularização dos carros *flex* após o ano de 2003 (RUDORFF et al., 2010). O aumento da demanda mundial por biocombustíveis exige a conversão de áreas de culturas convencionais e pastagens em áreas de cultivo de cana-de-açúcar, causando mudança do uso e ocupação da terra (LOARIE et al., 2011). A intervenção antrópica gerada por esta monocultura gera questionamentos socioambientais que ainda não foram bem respondidos, como a questão do desmatamento e perda de biodiversidade, o conflito da produção de biocombustíveis versus produção de alimentos, efeitos indiretos da mudança de uso da terra, poluição da área com resíduos da queimada, compactação e erosão do solo, entre outros (NASSAR et al., 2012).

A expansão desta cultura é intensa na região Centro-Sul do Brasil, em especial no Estado de São Paulo. Entre os anos de 2003 e 2013, a área plantada de cana-de-açúcar no estado aumentou 92%, em decorrência da crescente demanda do produto para produção de etanol (INPE, 2014). Segundo a União das Indústrias de Cana-de-Açúcar (UNICA, 2014), a safra de 2011/2012 indicou que a região Centro-Sul foi responsável por aproximadamente 90% da produção nacional de etanol e de 87% da produção de açúcar. Visando acabar com essa prática no Estado de São Paulo, em 2002 foi aprovada a Lei Estadual 11.241, que estipulou um cronograma para a eliminação da queima nos canaviais, com prazo final em 2021 para áreas mecanizáveis (declividade $\leq 12\%$) e em 2031 para áreas não mecanizáveis (declividade $> 12\%$). Com o intuito de acelerar o processo de diminuição dessa queima, a Secretaria de Meio Ambiente (SMA) do Estado de São Paulo e a União da Indústria de Cana-de-Açúcar (UNICA) firmaram, em agosto de 2007, um protocolo de intenções, denominado Protocolo Etanol Verde, em que a prática da queima da palha da cana-de-açúcar deve ser gradativamente reduzida até sua completa eliminação até 2017 para áreas não mecanizáveis e até 2014 em áreas mecanizáveis (RONQUIM, 2010).

A colheita da cana-de-açúcar pode ser feita com corte mecanizado (colheita verde ou crua), em que a cana-de-açúcar é colhida deixando a palha sobre o solo, ou corte manual, em que a queima da palhada é empregada de forma a facilitar o processo de colheita, porém causando danos ao meio ambiente (SUGAWARA, 2010). Entre os problemas causados pela queima da palha da cana-de-açúcar, pode-se destacar os

problemas de saúde causados aos trabalhadores do campo e a população em geral, emissão de gases de efeito estufa e aerossóis na atmosfera, que contribuem com as mudanças climáticas, entre outros. Atualmente a prática da colheita com queima é realizada em cerca de 25% do total da área cultivada no Estado de São Paulo (INPE, 2014).

Com o objetivo de avaliar o tipo de colheita da cana-de-açúcar e o padrão espacial deste, Novaes et al. (2011) propôs um índice de colheita (I_c), buscando relacionar o tipo de colheita com fatores socioeconômicos dos municípios da sua área de estudo. Sua hipótese é de que a queima da palha é mais empregada em locais onde as condições socioeconômicas da população envolvida com a cultura da cana é mais precária. Os autores afirmam que seus resultados não foram satisfatórios, sendo necessário considerar outros fatores na análise, como fatores topográficos e ambientais.

Nesse sentido, este trabalho teve como objetivo relacionar espacialmente o índice de colheita proposto por Novaes et al. (2011) com fatores topográficos e ambientais, verificando se essas variáveis podem explicar o tipo de colheita empregada.

2. Material e Métodos

2.1. Área de estudo

O Estado de São Paulo (Figura 1) situa-se entre as latitudes $19^{\circ}28'44''S$ e $25^{\circ}10'24''S$ e longitudes $44^{\circ}24'16''O$ e $53^{\circ}51'21''O$. De acordo com a Classificação climática de Köppen-Geiger, existem sete tipos climáticos em território paulista: Cwa, Cwb, Aw, Am, Cfa, Cfb e Af. A predominância é do clima tropical de altitude (Cwa), caracterizado por chuvas durante a estação do verão e períodos mais secos durante o inverno e com a temperatura média do mês mais quente ultrapassando $22^{\circ}C$ (CEPAGRI, 2013).

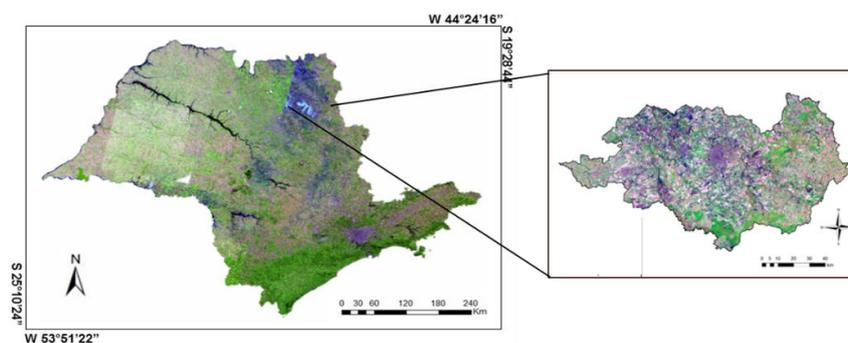


Figura 1. Estado de São Paulo e Região Administrativa de Ribeirão Preto.

A precipitação no Estado de São Paulo varia entre 1500 e 2000 mm ano⁻¹, com os valores máximos ocorrendo no litoral. A estação seca ocorre no período de junho a agosto, enquanto a estação chuvosa ocorre entre dezembro e fevereiro (JUÁREZ, 2004). Para as regiões produtoras de cana-de-açúcar, as altitudes variam entre 200 e 800 m, de acordo com os dados obtidos pela missão SRTM (*Shuttle Radar Topography Mission*) (RABUS, 2003). A cultura da cana-de-açúcar apresenta uma boa capacidade de adaptação em solos de diferentes texturas, favorecendo a expansão da cultura pelo Estado de São Paulo (KOFFLER e DONZELI, 1987). Segundo o INPE(2014), no ano de 2013 foram cultivados mais de 5,76 milhões de hectares da cultura da cana-de-açúcar no Estado de São Paulo, dos quais 5,9% corresponderam a expansão da cultura em relação ao ano anterior.

Conforme apresentado nos resultados, a análise de regressão foi restrita a Região Administrativa (RA) de Ribeirão Preto (Figura 1), por haver agrupamento da variável de interesse nessa área e por essa RA apresentar grande área de cana-de-açúcar colhida mecanicamente e com queima da palhada. A RA de Ribeirão Preto é composta por 25 municípios, que ocupam uma área total de 934 800 há (um total de 3,8% do território paulista) e população de 1.247.761 habitantes. Possui clima tropical, relevo suave e plano, com solos adequados ao desenvolvimento de agricultura tropical. Esses fatores favoreceram o seu desenvolvimento baseado na agroindústria, destacando-se a forte expansão da produção canavieira na região, o que gerou efeitos na estruturação do espaço regional, com a substituição de outras culturas pela cana-de-açúcar. O agronegócio sucroalcooleiro é o principal indutor de desenvolvimento da economia da RA, de segmentos industriais a ele relacionado, como o setor metal-mecânico ou o de turismo de negócios. (GOVERNO DO ESTADO DE SÃO PAULO, 2012).

2.2. Dados utilizados

2.2.1. Mapeamento do projeto Canasat

As áreas de cana-de-açúcar para o Estado de São Paulo foram obtidas pelos mapas temáticos do Projeto Canasat (INPE, 2014) para os anos de 2008, 2010 e 2012 (Figura 2), separadas nas classes “Cana crua” e “Cana queima”, que indicam o tipo de colheita empregada. A classe “cana bisada”, que representa as áreas de cana-de-açúcar não colhidas em cada ano foi descartada da análise, pois as áreas de interesse são as disponíveis para colheita em cada safra.

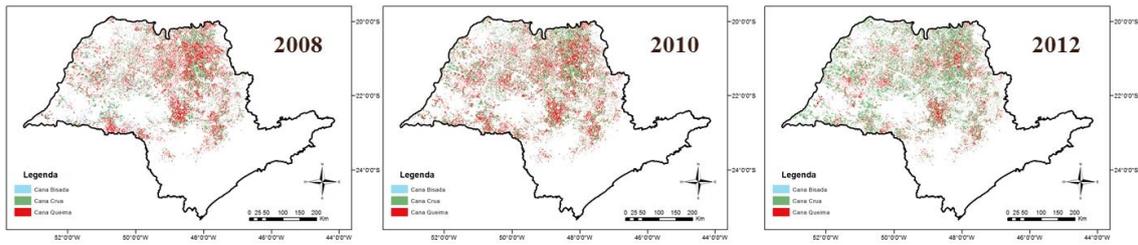


Figura 2. Mapeamento do Canasat para os anos de 2008, 2010 e 2012.

2.2.2. Dados Ambientais – Imagens MODIS:

A partir das imagens de reflectância da superfície do produto MOD09A1 (resolução espacial de 500m e resolução temporal de 8 dias), foram geradas séries temporais de albedo da superfície (segundo o modelo proposto por Liang (2000)) e do índice de vegetação EVI2 (JIANG et al., 2008). As séries temporais de albedo e EVI2 foram filtradas por meio da análise harmônica de séries temporais. Essa técnica permite representar uma função complexa como a somatória de ondas (termos) senoidais e cossenoidais, onde cada onda é definida por um valor único de amplitude e fase (LACRUZ, 2006). Para as duas variáveis foram extraídos os valores mínimos de albedo e EVI2 para cada pixel das imagens MODIS em cada ano (Figura 3). O valor mínimo de albedo e EVI2 tem como objetivo destacar as áreas de cana-de-açúcar em que se emprega a queima da palha na colheita, uma vez que áreas queimadas apresentam baixa reflectância na imagem orbital e baixo índice de vegetação. A Figura 3 ilustra o processo para gerar as imagens de albedo mínimo e EVI2 mínimo.

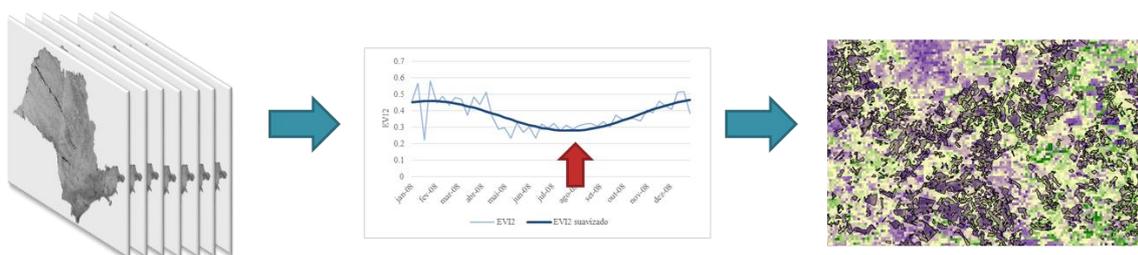


Figura 3. Processo para gerar as imagens de albedo da superfície mínimo e EVI2 mínimo para cada ano.

Além disso, foram utilizados os dados de focos de calor do produto MOD14 do MODIS. Os focos são obtidos em formato de pontos, em que foram detectadas altas temperaturas pelo sensor orbital. A Figura 4 apresenta os focos de queimada para o Estado

de São Paulo em 2008, 2010 e 2012, com os respectivos mapas de densidade de Kernel para cada ano.

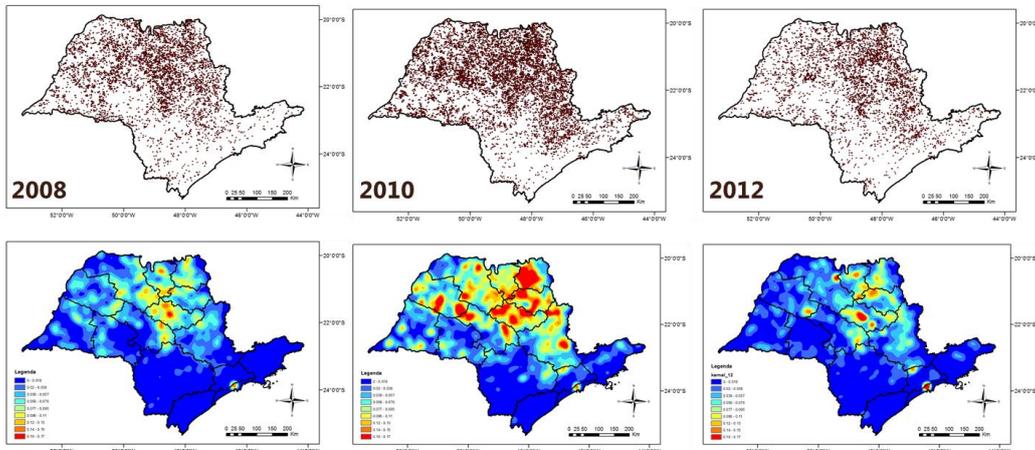


Figura 4. Focos de calor obtidos do produto MOD14; mapas de densidade de Kernel para os focos calor no Estado de São Paulo.

2.2.3. Modelo digital de elevação

Os dados topográficos foram obtidos do Topodata (VALERIANO, 2005), que é um refinamento dos dados do SRTM (*Shuttle Radar Topography Mission*) para resolução espacial de 30 metros. A Figura 5 apresenta os mapas de elevação e declividade utilizados no estudo.

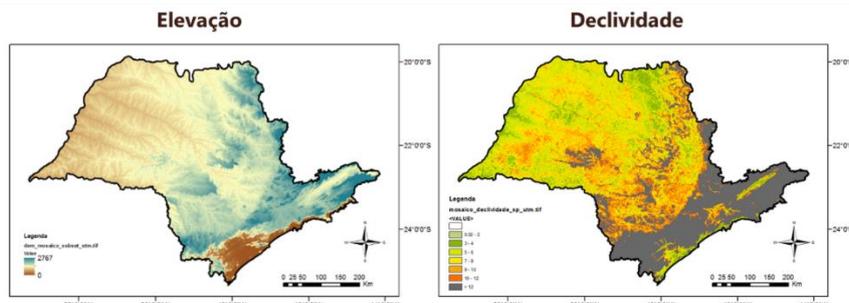


Figura 5. Modelo de elevação e declividade disponibilizados pelo Topodata.

2.3. Análise de agrupamento

Inicialmente foi realizada uma análise do agrupamento quanto a colheita com queima no Estado de São Paulo, de modo a entender como essas áreas se distribuem no espaço e se há aglomerados dentro da área de estudo. Para tanto, foi calculado o Índice Global de Moran, que é um indicador global de autocorrelação espacial e mostra como os valores da variável de interesse estão correlacionados no espaço. O índice varia entre

-1 e 1, de modo que quanto mais próximo de 1, maior a correlação espacial (objeto tem valor mais alto quando seus vizinhos também têm), valores mais próximos de -1 indicam correlação espacial inversa (objeto tem valor baixo enquanto os vizinhos têm valor alto) e valores próximos a zero indicam ausência de correlação espacial (DRUCK et al.,2004). Após o cálculo do Índice Global de Moran, estimou-se sua significância pelo teste de pseudo-significância, realizando-se 999 permutações dos valores dos atributos associados às regiões. O teste avalia a hipótese de que os dados são estacionários de primeira e segunda ordem (CÂMARA et al., 2004).

Foram gerados os indicadores locais de autocorrelação espacial – estatística LISA (*Local Indicators of Spatial Association*) (ANSELIN, 1995) para cada ano (para um intervalo de confiança de 95%), de modo a evidenciar os agrupamentos quanto ao tipo de colheita, e assim obter uma área com comportamento homogêneo para aplicar a análise de regressão para o índice de colheita. Essa etapa do trabalho foi realizada no software GeoDA (ANSELIN, 2005).

2.4. Integração dos dados em ambiente celular

Uma vez que se definiu uma área com comportamento similar quanto ao tipo de colheita, foi realizada a integração dos dados em ambiente celular no ArcGIS 10.2, com a ferramenta Hawth's Tools (BEYER, 2004). Essa técnica permite o cálculo de atributos de tabelas associados às células, homogeneizando as informações de diferentes fontes e em diferentes formatos. Após testes para o tamanho de célula, optou-se pelas células de 500 m, pois esse tamanho permite um bom desempenho computacional sem perder informações das variáveis devido a escala. A Figura 6 exemplifica a integração dos dados em ambiente celular.

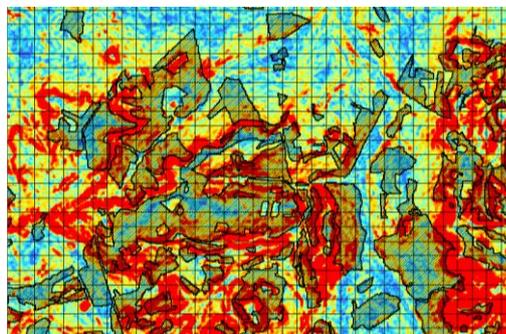


Figura 6. Integração dos dados em ambiente celular. Os polígonos hachurados pretos representam os polígonos do mapeamento do Canasat. A imagem colorida representa a declividade.

Seguindo a metodologia empregada por Novaes et al. (2011), para cada célula foi calculado o índice de colheita, que relaciona a proporção das áreas de cana-de-açúcar com cada tipo de colheita (Equação 1).

$$I_c = \frac{C_{crua} - C_{queima}}{C_{crua} + C_{queima}} \quad (1)$$

em que C_{crua} é a área de cana-de-açúcar com colheita verde e C_{queima} é a área de cana-de-açúcar colhida com queima da palha, obtidos dos mapas do Canasat. O índice varia entre -1 e 1, de modo que valores positivos mais próximos de 1 indicam predominância da colheita verde e valores mais próximos de -1 indicam que prevalece a colheita com queima.

Para os dados de elevação e declividade calculou-se a média de cada variável em cada célula, assim como para o albedo e EVI2. Já para os focos de queimada, por se tratar de dados pontuais, foi calculado o número de focos em cada célula.

2.5. Análise de regressão

Como se busca neste estudo estabelecer a relação entre o índice de colheita (variável dependente) e as variáveis ambientais e topográficas (explicativas), foram comparados três métodos de regressão: regressão clássica e dois modelos que incorporam os efeitos espaciais (Spatial Lag e Spatial Error). Os resultados da análise de regressão foram gerados no software GeoDA.

2.5.1. Regressão Linear Clássica

Inicialmente foi aplicada a regressão linear clássica, que não considera os efeitos espaciais. A regressão Linear Clássica (Equação 2) busca prever e relacionar uma variável dependente a partir de valores conhecidos de uma variável explicativa. Para aplicação dessa técnica, supõe-se os dados apresentem independência espacial. Além disso, os resíduos devem ter variância constante (homocedásticos), ter distribuição normal e ser independentes (AGUIAR et al., 2007). Dessa forma, foi realizada uma análise dos resíduos, de modo a avaliar a regressão, e verificar se esse é um modelo adequado para tratar do problema do estudo.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \varepsilon \quad (2)$$

2.5.2. Regressão Espacial

Neste estudo foram utilizados dois modelos de regressão com efeitos espaciais. Os modelos agregam os efeitos espaciais, supondo que é possível capturar a estrutura de correlação espacial através de um parâmetro adicionado ao modelo de regressão clássico (BAILEY e GATRELL, 1995; DRUCK et al., 2004). Os modelos avaliados foram o Spatial Lag e o Spatial Error, aplicados através do software GeoDA.

O modelo Spatial Lag atribui a correlação espacial à própria variável dependente Y, segundo a Equação 3.

$$Y = \rho WY + \beta_0 X + \beta_1 X_1 + \dots + \beta_n X_n + \varepsilon \quad (3)$$

em que ρ é o coeficiente espacial autoregressivo, W é a matriz de correlação espacial, WY expressa a dependência espacial em Y, β são os coeficientes associados às respectivas variáveis explicativas e ε é o erro aleatório.

O modelo Spatial Error atribui a autocorrelação espacial ao erro (Equação 4), sendo os efeitos espaciais considerados como ruídos ou perturbações que devem ser removidos (ANSELIN, 1988).

$$Y = X\beta + \varepsilon \quad \varepsilon = \rho W\varepsilon + \xi \quad (4)$$

em que β é o coeficiente associado a cada variável explicativa, ρ é a medida de correlação espacial, $W\varepsilon$ é o erro com efeitos espaciais e ξ é a componente do erro com variância constante e não correlacionada.

Para os modelos de regressão espacial também foi realizada análise dos resíduos, verificando a homocedasticidade, distribuição normal e independência espacial dos resíduos, comparando-se ao final qual modelo apresentou melhor desempenho.

A Figura 7 apresenta um fluxograma que explica as etapas de realização deste estudo.

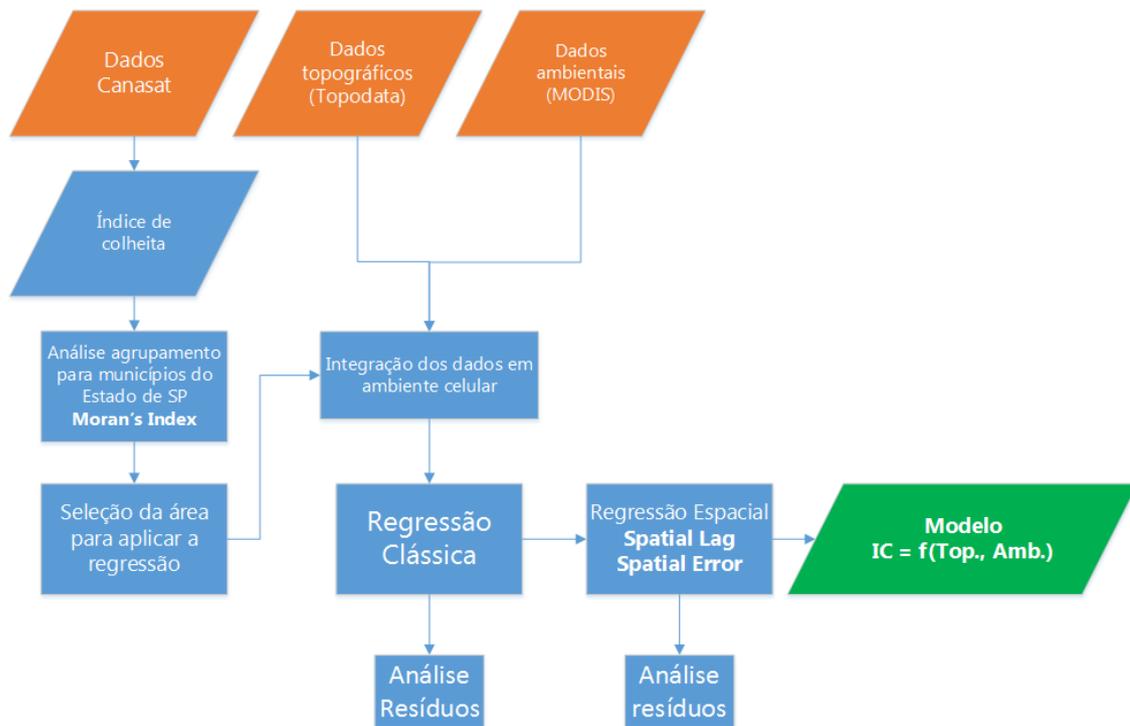


Figura 7. Fluxograma das etapas do trabalho.

3. Resultados e Discussão

3.1. Análise de agrupamento

A análise de agrupamento das áreas de cana-de-açúcar colhidas com queima da palha foi realizada para os municípios para o Estado de São Paulo. A Figura 8 apresenta os diagramas de espalhamento de Moran e os respectivos LISA Maps para os três anos avaliados.

Para 2008 o Índice de Moran foi de 0,55, testado com 5% de significância, notando-se na Figura 8 para esse ano um forte agrupamento “High-high” nas RAs de Barretos, Franca, Ribeirão Preto e Central do Estado de São Paulo, sendo essas áreas onde mais se concentra a colheita da cana-de-açúcar com queimada. Para 2010 e 2012 o Índice de Moran foi de 0,48, mostrando um padrão semelhante quanto ao agrupamento das áreas colhidas com queima, embora se perceba uma diminuição dessas áreas em cada ano. Dessa forma, com o Índice de Moran, testado a 5% de significância, verificou-se que há correlação espacial entre as áreas de cana colhida com queima, rejeitando-se a hipótese nula de que não há correlação espacial.

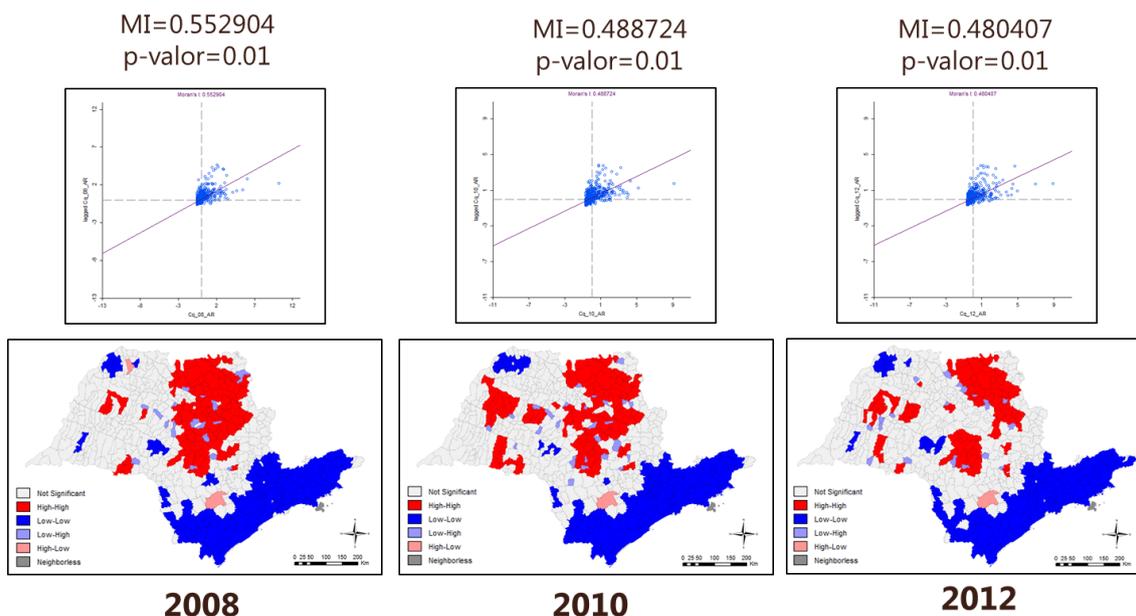


Figura 8. Diagrama de dispersão do índice Global de Moran para a colheita com queima nos municípios do Estado de São Paulo, para os anos de 2008, 2010 e 2012; LISA Map para 5% de significância.

Essa região de agrupamento “High-high” é interessante pois concentra a colheita da cana com queimada no Estado de São Paulo, além de apresentar também uma grande área de cana-de-açúcar colhida mecanicamente. Dessa forma, para aplicar os modelos de regressão selecionou-se a RA de Ribeirão Preto (Figura 1), que por apresentar-se inserida nesse cluster e ter abundância de áreas de cana-de-açúcar colhida com queima e sem queima, permite avaliar o índice de colheita em função das variáveis explicativas utilizadas nesse estudo.

3.2. Regressão Clássica

O modelo de regressão clássica relacionou a variável dependente (índice de colheita) com as variáveis explicativas (declividade, elevação, EVI2 mínimo, albedo mínimo e número de focos de queimada), sendo realizado no GeoDA. A Tabela 1 apresenta os resultados obtidos. Com relação ao coeficiente de determinação (R^2), verifica-se que o modelo explica pouco da variância dos dados, não ultrapassando o valor de 0,20. Quanto a multicolinearidade, para os três anos obteve-se um resultado similar (~20). Esse parâmetro avalia se as variáveis explicativas são muito correlacionadas, de modo que valores menores do que 30 indicam que não há problemas com relação a

multicolinearidade. Os resultados para Log Likelihood, Critério de Informação de Akaike e Critério de Schwarz indicam a qualidade do ajuste do modelo, sendo mais interessante discuti-los adiante na comparação final entre os modelos de regressão.

Tabela 1. Resultados da regressão clássica.

	R ²	Multicolinearidade	Normalidade Jarque-Bera	Heteroced. Breusch-Pagan test	Log likelihood	Akaike	Schwarz crit.
2008	0.09	19.90	2578 prob=0.0000	Valor=46 prob= 0.0000	-33915	67842	67891
2010	0.20	19.63	1633 prob=0.0000	Valor= 568 prob=0.0000	-32734	65480	65529
2012	0.09	19.63	2506 prob=0.0000	Valor= 436 prob=0.0000	-32780	65572	65621

A análise dos resíduos permitiu verificar se os mesmos apresentam distribuição normal, são homocedásticos e independentes. A normalidade foi testada pelo teste de Jarque-Bera (Tabela 1), mostrando que para os três anos avaliados os resíduos não apresentam distribuição normal (como pode ser verificado nos histogramas da Figura 9), rejeitando a hipótese de distribuição normal a um nível de significância de 5%.

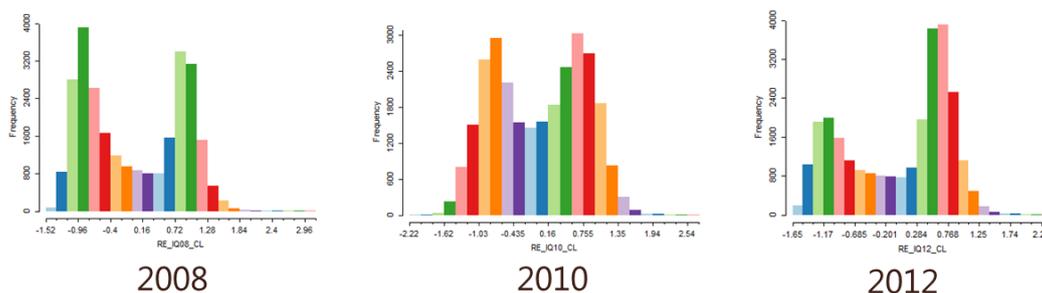


Figura 9. Histograma dos resíduos para a regressão clássica.

O teste de Breusch-Pagan (Tabela 1) indica que os resíduos da regressão clássica são heterocedásticos (variância não constante), conforme pode-se verificar visualmente na Figura 10, onde os resíduos são plotados em função das variáveis explicativas.

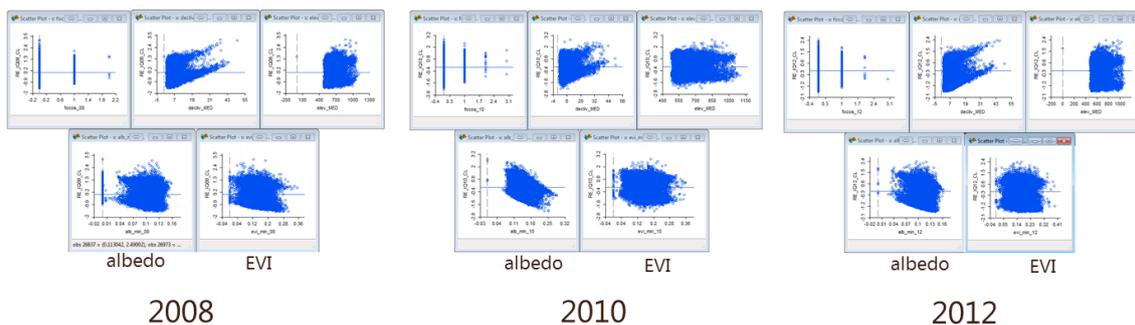


Figura 10. Diagramas de dispersão entre os resíduos da regressão clássica e as variáveis explicativas do modelo.

Verificou-se em seguida a autocorrelação espacial dos resíduos com o índice Global de Moran, conforme apresentado na Figura 11. Com índices de 0,58, 0,47 e 0,50 para os anos de 2008, 2010 e 2012, respectivamente, com nível de significância de 5%, rejeitou-se a hipótese nula de que não há autocorrelação espacial.

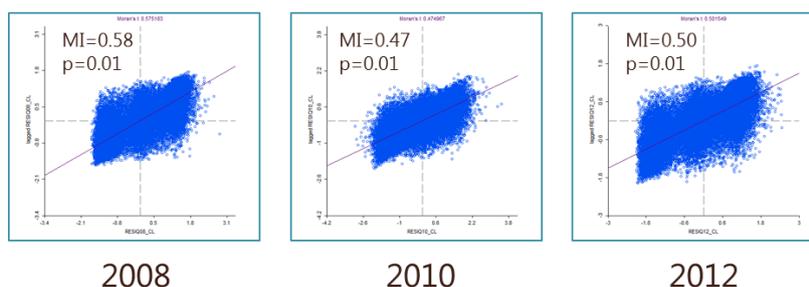


Figura 11. Índice Global de Moran para os resíduos da regressão clássicas.

Segundo os testes realizados para a regressão clássica, este modelo não atende as premissas de homocedasticidade, normalidade e independência dos resíduos, não sendo assim um modelo confiável. A dependência espacial verificada na Figura 11 mostra a necessidade de utilização de modelos de regressão que incorporam os efeitos espaciais. Para tanto, utilizou-se dois modelos disponíveis no software GeoDA, o Spatial Lag e Spatial Error.

3.3. Regressão Espacial – Spatial Lag

Aplicando-se o modelo de regressão espacial Spatial Lag no software GeoDA, foram obtidos os resultados apresentados na Tabela 2. Considerando o fator espacial na

variável dependente, os coeficientes de determinação (R^2) foram de 0,69, 0,64 e 0,62, para 2008, 2010 e 2012, respectivamente. Nota-se uma melhora nos valores do R^2 com relação ao modelo de regressão clássica, porém é importante destacar que estimadores e diagnósticos tradicionais de regressão como o R^2 não levam em conta os efeitos espaciais na sua determinação, portanto não são confiáveis para avaliar a qualidade do ajuste em modelos de regressão espacial (LOPES et al., 2006).

Para a comparação e seleção de modelos de regressão espacial, normalmente se utiliza os valores de máxima verossimilhança (Log likelihood), ponderado pela diferença no número de parâmetros estimados. Para modelos espaciais utiliza-se os critérios de informação onde a avaliação do ajuste é penalizada por uma função do número de parâmetros (LOPES et al., 2006). Assim, os parâmetros Log likelihood, Critério de informação de Akaike e Critério de Schwarz, apresentados na Tabela 2, são mais adequados para avaliar a qualidade do ajuste pelo modelo espacial.

Para o Log Likelihood, quanto maior o valor, melhor (menos negativo, nesse caso), assim, o melhor resultado foi obtido para o ano de 2008. Para o Critério de Akaike, quanto menor o valor, melhor é o ajuste, de modo que 2008 também apresentou melhores resultados, assim como indica o critério de Schwarz, que é similar ao Critério de Akaike.

Tabela 2. Resultados da regressão espacial com o modelo Spatial Lag.

	R^2	Heteroced. Breusch-Pagan test	Log likelihood	Akaike	Schwarz crit.
2008	0.69	Valor=79 prob: 0.0000	-21748	43511	43568
2010	0.64	Valor: 319 prob:0.0000	-23475	46964	47021
2012	0.62	Valor: 308 prob:0.0000	-23067	46149	46207

Os resíduos foram também analisados quanto a normalidade, homocedasticidade e independência espacial. A Figura 12 mostra que incluindo o termo espacial no modelo de regressão, a distribuição dos resíduos se aproxima mais de uma distribuição normal, para os três anos avaliados.

Já com relação a variância dos resíduos, o teste de Breusch-Pagan (Tabela 2) indica que os resíduos são heterocedásticos, conforme verifica-se visualmente na Figura 13. A autocorrelação espacial dos resíduos foi verificada utilizando o índice Global de

Moran, obtendo-se os resultados apresentados na Figura 14. Para os três anos avaliados os índices de Moran foram próximos a zero, indicando que não há correlação espacial. Porém, para o ano de 2012 o teste de pseudo-significância apresentou um valor- $p=0,3$, não podendo-se afirmar portanto que não há correlação espacial para os resíduos nesse caso. Dessa forma, o Spatial Lag também não atende a todas as premissas para que um modelo de regressão seja confiável.

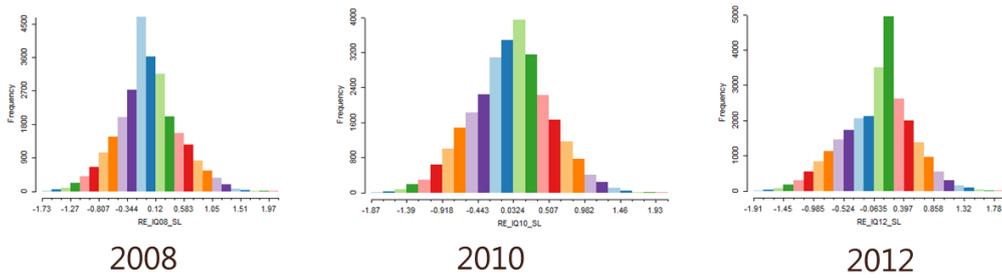


Figura 12. Histograma dos resíduos para a regressão espacial pelo modelo Spatial Lag.

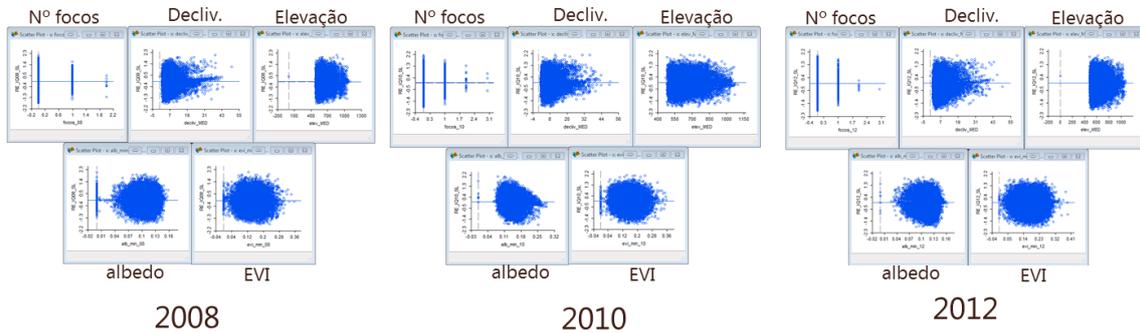


Figura 13. Diagramas de dispersão entre os resíduos da regressão espacial pelo modelo Spatial Lag e as variáveis explicativas.

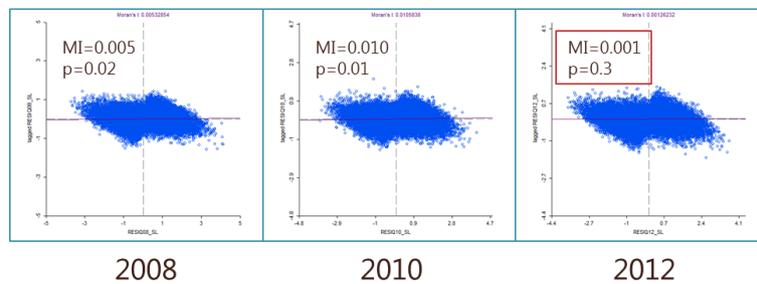


Figura 14. Índice Global de Moran para os resíduos da regressão espacial pelo modelo Spatial Lag.

3.4. Regressão Espacial – Spatial Error

O segundo modelo de regressão espacial testado no software GeoDA foi o Spatial Error, obtendo-se os resultados apresentados na Tabela 3. Considerando o fator espacial no erro, obteve valores de R^2 de 0,69, 0,64 e 0,62 para 2008, 2010 e 2012, respectivamente, assim como para o Spatial Lag. Como já discutido para o Spatial Lag, o R^2 não é um estimador confiável para avaliar a qualidade do ajuste da regressão, sendo necessário utilizar os valores de Log likelihood, Critério de Informação de Akaike e Critério de Schwarz. Segundo esses critérios, o melhor resultado foi obtido para o ano de 2008.

Tabela 3. Resultados da regressão espacial com o modelo Spatial Error.

	R^2	Heteroced. Breusch-Pagan test	Log likelihood	Akaïke	Schwarz crit.
2008	0.69	Valor=77 prob: 0.0000	-21901	43815	43864
2010	0.64	Valor: 335 prob:0.0000	-23761	47534	47584
2012	0.62	Valor: 303 prob:0.0000	-23211	46435	46484

A análise dos resíduos se deu da mesma maneira que para o modelo Spatial Lag, avaliando a normalidade, homocedasticidade e dependência espacial dos resíduos. A Figura 15 mostra que incluindo o termo espacial no erro (Spatial Error), a distribuição dos resíduos se aproxima mais de uma distribuição normal do que no caso da regressão clássica.

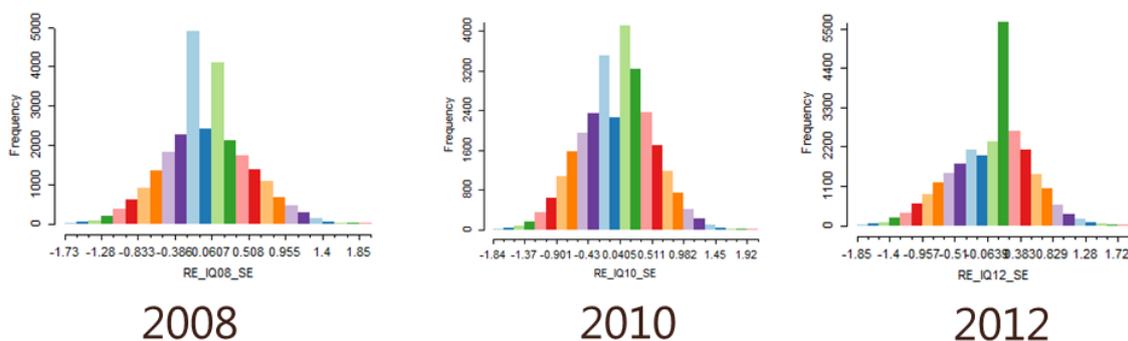


Figura 15. Histograma dos resíduos para a regressão espacial pelo modelo Spatial Error.

Com relação a variância dos resíduos, o teste de Breusch-Pagan (Tabela 3) indica que os resíduos são heterocedásticos, assim como pode-se verificar na Figura 16. Assim,

o modelo Spatial Error também não atende a premissa da homocedasticidade dos resíduos, não fornecendo um modelo de regressão confiável.

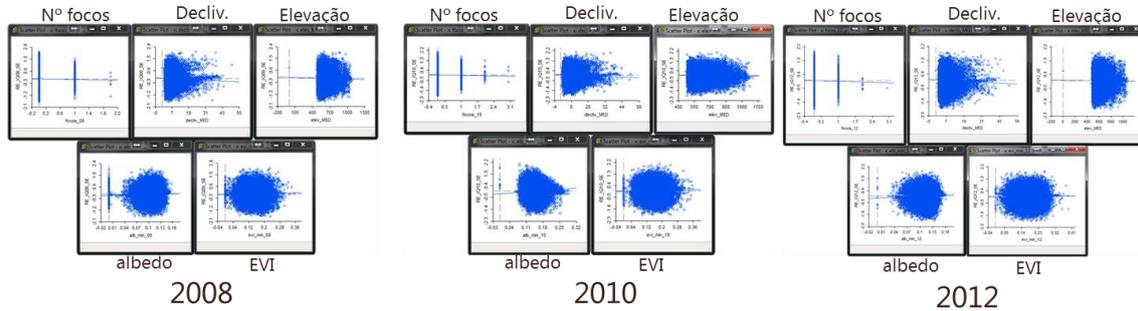


Figura 16. Diagramas de dispersão entre os resíduos da regressão espacial pelo modelo Spatial Error e as variáveis explicativas.

O teste para autocorrelação espacial é apresentado na Figura 17, com os diagramas de dispersão do índice de Moran Global dos resíduos. O índice de Moran é próximo de zero para os três anos, porém o teste de pseudo-significância indica que para os anos de 2010 e 2012 não se pode afirmar que não há autocorrelação espacial, assim como foi obtido para o ano de 2012 para o Spatial Lag.

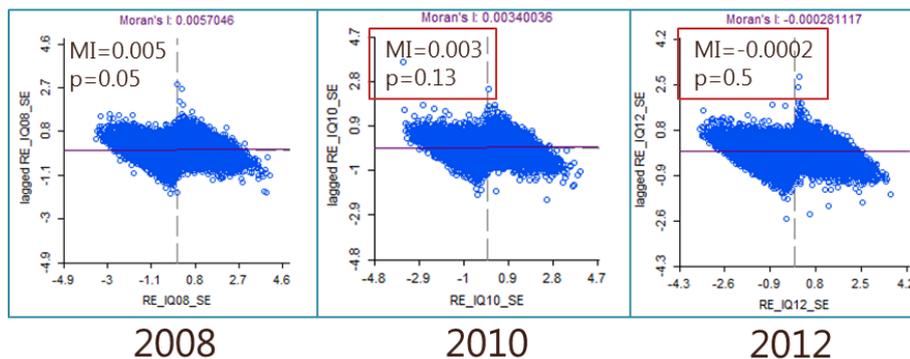


Figura 17. Índice Global de Moran para os resíduos da regressão espacial pelo modelo Spatial Error.

3.5. Comparação entre os modelos de regressão

Na tabela 4 são apresentados os resultados obtidos para os três modelos de regressão utilizados nesse estudo, para comparação dos mesmos.

Tabela 4. Comparação entre os resultados dos modelos de regressão avaliados.

	2008				2010				2012			
	R ²	Akaike	Log likel.	I.Moran Resid.	R ²	Akaike	Log likel	I.Moran Resid.	R ²	Akaike	Log likel	I.Moran Resid.
Reg. Clássica	0.09	67842	-33915	0.58 p=0.01	0.20	65480	-32734	0.47 p=0.01	0.09	65572	-32780	0.50 p=0.01
Spatial Lag	0.69	43511	-21748	0.005 p=0.02	0.64	46964	-23475	0.010 p=0.01	0.62	46149	-23067	0.001 p=0.3
Spatial Error	0.69	43815	-21901	0.005 p=0.05	0.64	47534	-23761	0.003 p=0.13	0.62	46435	-23211	0.0002 p=0.5

Conforme discutido anteriormente, os valores de R² melhoram ao se aplicar o fator espacial no modelo de regressão, porém avaliar o ajuste somente pelo R² leva a conclusões equivocadas. O modelo Spatial Lag apresentou os melhores resultados segundo o Critério de Informação de Akaike (menores valores) e Log likelihood (valores menos negativos). Apesar disso, nenhum dos modelos atende todas as premissas para um modelo de regressão confiável, não obtendo-se resíduos homocedásticos mesmo para os modelos espaciais.

A Tabela 5 apresenta os coeficientes da regressão espacial (Equação 5) obtidos para o Spatial Lag, que apresentou um desempenho superior aos demais modelos testados.

$$Ic = \rho \cdot W \cdot Ic + \beta_0 + \beta_1 \cdot n^{\circ} \text{ focos} + \beta_2 \cdot \text{declividade} + \beta_3 \cdot \text{elevação} + \beta_4 \cdot \text{albedo} + \beta_5 \cdot \text{EVI2} \quad (5)$$

Tabela 5. Coeficientes da regressão espacial pelo modelo Spatial Lag.

	2008	2010	2012
W	0.8707267	0.8180957	0.8331319
β₀	-0.2766547	-0.5039933	-0.3379003
β₁	-0.04566439	-0.01496748	0.04744849
β₂	-0.01256708	-0.009844173	-0.01356491
β₃	8.450e-005	-1.3027e-005	0.0001582646
β₄	1.632627	3.739235	2.749045
β₅	0.9549892	0.3361142	0.3934478

Os valores de β_2 (relacionados a declividade), assim como se esperava, foram negativos, mostrando que quando se aumenta a declividade, a tendência é que o Índice de Colheita diminua (o que indica predominância de colheita com queima). Os valores de β_4 e β_5 (relacionados com o albedo e o EVI2) foram positivos, indicando uma relação direta entre essas variáveis e o índice de colheita. Assim, quanto menor o valor de albedo e EVI2, menor o índice de colheita (indicando predominância de áreas colhidas com queimada).

4. Conclusões

Neste estudo, foi verificado que a inclusão dos efeitos de dependência espacial nos modelos de regressão é importante, visto que o resultado obtido para a regressão espacial foi melhor que o da regressão clássica. Assim, verificou-se que o índice de colheita apresenta correlação espacial com as variáveis explicativas, mas os modelos de regressão não foram adequados para descrever o fenômeno, visto que mesmo para os modelos espaciais a análise de resíduos mostra que estes não são homocedásticos.

Esperava-se que em cada ano o resultado fosse diferente, refletindo uma mudança nos padrões de distribuição da colheita da cana com queimada, frente às restrições impostas pelo protocolo agroambiental. Porém os resultados foram similares para os anos de 2008, 2010 e 2012.

Os dados de focos de calor, além de serem pontuais, ainda entram no modelo como valores discretos do número de focos que ocorrem em cada célula. Essa não é a maneira mais adequada de utilizar esses dados, somando ainda o fato de que zero focos em uma célula não quer dizer que não houve queimada ali. Uma alternativa pode ser utilizar não o número de focos, mas a densidade de focos (estimador de densidade de Kernel, por exemplo), que entraria como um valor médio em cada célula.

Além disso, pode-se testar outras abordagens, como verificar os resultados para outros tamanhos de células, ou mesmo trabalhar com os limites de município. Uma integração das variáveis topográficas e ambientais com outras, como as variáveis socioeconômicas utilizadas por Novaes et al. (2011), podem levar a melhores resultados.

Apesar da melhora apresentada no ajuste pelos modelos espaciais, a questão da heterocedasticidade não é resolvida. É possível que este tipo de modelo estatístico não

seja o mais adequado para o problema em questão, e nesse caso deve-se procurar outros métodos.

REFERÊNCIAS BIBLIOGRÁFICAS

AGUIAR, D.A.; WAGNER F.S.; FEITOSA, F.F.; GONÇALVES, F.G.; RIZZI, R.; RUDORFF, B.F.T. Análise espacial da colheita da cana-de-açúcar no Estado de São Paulo: a influência da precipitação. In: SBSR, 13, 2007, Florianópolis. Anais... SJC: INPE, 2007, p.2.231-2.238.

ANSELIN, L. **Spatial Econometrics: Methods and Models**. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1988.

ANSELIN, L. Local indicators of spatial association - LISA. **Geographical Analysis**, New York, v.27, n.2, p.93-115, 1995.

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

BAILEY, C.; GATRELL, A.C. **Interactive spatial data analysis**. Upper Saddle River: Prentice Hall, 1995. 432 p.

BEYER, H. L. 2004. Hawth's Analysis Tools for ArcGIS. Disponível em: <<http://www spatialecology.com/htools>> Acesso em: dez. 2014.

CAMARA, G.; MONTEIRO, A.M.; DRUCK, S.; CARVALHO, M.S. Análise espacial de áreas. In: DRUCK, S.; CARVALHO, C.; CÂMARA, G.; MONTEIRO, A.M.V. (Org.). **Análise espacial de dados geográficos**. Brasília: EMBRAPA, 2004.

CENTRO DE PESQUISA METEOROLÓGICAS E CLIMÁTICAS APLICADAS A AGRICULTURA (Cepagri). **A classificação climática de Koeppén para o Estado de São Paulo**. Disponível em: <<http://www.cepagri.unicamp.br/outras-informacoes/clima-dos-municipios-paulistas.html>>. Acesso em: dez. 2014.

DRUCK, S.; CARVALHO, C.; CÂMARA, G.; MONTEIRO, A.M.V.M. **Análise de Dados Geográficos**. Brasília, EMBRAPA, 2004.

GOVERNO DO ESTADO DE SÃO PAULO. **Caracterização socioeconômica das regiões do Estado de São Paulo**: Região Administrativa de Ribeirão Preto. Disponível em: <http://www.planejamento.sp.gov.br/noti_anexo/files/uam/trabalhos/Ribeir%C3%A3o%20Preto.pdf> Acesso em: jan. 2015.

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS (INPE). Divisão de Sensoriamento Remoto (DSR). **Canasat - Mapeamento da cana via imagens de satélite de observação da Terra**. 2013. Disponível em <<http://www.dsr.inpe.br/laf/canasat/index.html>> Acesso em: out. 2014.

JIANG, Z.; HUETE, A.R.; DIDAN, K. MIURA, T. Development of a two-band enhanced vegetation index without a blue band. **Remote Sensing of Environment**. V.112, n.10, p.3833-3845, 2008.

JUÁREZ, R. I. N. **Variabilidade climática regional e controle da vegetação no Sudeste: Um estudo de observações sobre cerrado e cana-de-açúcar e Modelagem Numérica da**

atmosfera. 2011. 185p. Tese (Doutorado em Meteorologia) - Instituto de Astronomia, Geofísica e Ciências Atmosféricas – Universidade de São Paulo, São Paulo, 2004.

KOFFLER, N. F.; DONZELI, P. L. Avaliação dos solos brasileiros para cultura da cana-de-açúcar. In: Paranhos, S. B. (coord.) **Cana-de-açúcar. Cultivo e utilização**. Campinas: Fundação Cargill, 1987. p. 19-41.

LACRUZ, M. S. **Análise de séries temporais de dados MODIS como uma nova técnica para a caracterização de paisagem e análise de lacunas de conservação**. 129 p. Tese (Doutorado em Sensoriamento Remoto) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2006.

LIANG. Narrowband to broadband conversions of land surface albedo I: Algorithms. **Remote Sensing of Environment**, v. 76 n.1, p. 213–238, 2000.

LOARIE, S. R.; LOBELL, D. B.; ASNER, G. P.; MU, Q.; FIELD, C. B. Direct impacts on local climate of sugar-cane expansion in Brazil. **Nature Climate Change**, v. 1, n. 2, p. 105-109, 2011.

LOPES, S. B.; BRONDINO, N.C.M. ; SILVA, A.N.R. . ANÁLISE DO DESEMPENHO DE MODELOS DE REGRESSÃO ESPACIAL NA PREVISÃO DE DEMANDA POR TRANSPORTES. In: XIV Congresso Panamericano de Ingeniería de Tránsito y Transporte, 2006, Las Palmas de Gran Canaria. XIV PANAM, 2006.

NASSAR, A.M.; RUDORFF, B.F.T.; ANTONIAZZI, L.B.; AGUIAR, D.A.; BACCHI, M.R.P.; ADAMI, M. **Prospects of the sugarcane expansion in Brazil**: Impacts on direct and indirect land use changes. In Sugarcane ethanol: Contributions to climate change mitigation and the environment, Zuurbier, P., van De Vooren, J., Eds.; Wageningen Academic Publishers: Wageningen, The Netherlands, 2008; pp. 63–92.

NOVAES, M. R.; RUDORFF, B.F.T.; ALMEIDA, C.M.; AGUIAR, D.A. Análise espacial da redução da queima na colheita da cana-de-açúcar: perspectivas futuras ao cumprimento do protocolo agroambiental. **Eng. Agríc. Jaboticabal**, v.31, n.3, p.572-583, 2011.

RABUS, B.; EINEDER, M.; ROTH, A.; BAMLER, R. The shuttle radar topography mission - a new class of digital elevation models acquired by spaceborne radar. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 57, n. 4, p. 241-262, 2003.

RONQUIM, C. C. **Queimada na colheita da cana-de-açúcar**: impactos ambientais, sociais e econômicos. Embrapa Monitoramento por Satélite, 2010.

RUDORFF, B. F. T.; AGUIAR, D.A.; SILVA, W.F.; SUGAWARA, L.M.; ADAMI, M.; MOREIRA, M.A. Studies on the Rapid Expansion of Sugarcane for Ethanol Production in São Paulo State (Brazil) Using Landsat Data. **Remote Sensing**, v. 2, n. 4, p. 1057-1076, 2010.

SUGAWARA, L. M. **Variação interanual da produtividade agrícola da cana-de-açúcar por meio de um modelo agrônomo**. 2007. 92 p. Tese (Doutorado em Sensoriamento Remoto). Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2010.

ÚNICA – União das Indústrias de Cana-de-açúcar, 2014. Disponível em: <<http://www.unica.com.br/>>. Acesso em: dez. 2014.

VALERIANO, M. M. Modelo digital de variáveis morfométricas com dados SRTM para o território nacional: o projeto TOPODATA. In: XII Simpósio Brasileiro de Sensoriamento Remoto, 2005, Goiânia, GO. Anais do XII Simpósio Brasileiro de Sensoriamento Remoto, 2005. p. 1-8.