

Análise Espacial de Indicadores Sócio-Econômicos na Compreensão do Manejo da Colheita da Cana-de-Açúcar

Maikon R. de Novaes¹

Daniel Aguiar¹

Bernardo F. T. Rudorff¹

¹Divisão de Sensoriamento Remoto - DSR
Instituto Nacional de Pesquisas Espaciais - INPE
Caixa Postal 515 - 12201-970 - São José dos Campos - SP, Brasil

[maikon;daniel;bernardo]@dsr.inpe.br

1. Introdução

Alguns países têm adotado como medida para redução da emissão de gases do efeito estufa a produção e o uso de biocombustíveis. No Brasil, desde a década de 70, utiliza-se o etanol produzido a partir da cana-de-açúcar como combustível automotivo em ampla escala (Moreira e Goldemberg, 1999). Recentemente, com o desenvolvimento dos automóveis bicombustíveis, a produção do etanol alcançou números nunca antes vistos e é a segunda fonte da matriz energética brasileira, responsável por 16% de toda a energia consumida no país (Veiga Filho, 2008).

Novas áreas de cana-de-açúcar têm sido cultivadas em toda a região Centro-Sul, principalmente no Estado de São Paulo, onde, apenas nos dois últimos anos, foram plantados 1,3 milhões de ha (www.dsr.inpe.br/canasat). Em contrapartida, diversos autores apontam para problemas sociais e ambientais relacionados ao cultivo da cana-de-açúcar (Roseiro e Nakayanagui, 2004; Goldemberg et al., 2008; Smeets et al., 2008), dos quais pode-se destacar a poluição atmosférica e os efeitos nocivos à saúde humana, causados pela queima pré-colheita da palha da cana-de-açúcar (Ribeiro e Assunção, 2002; Lopes e Ribeiro, 2006; Ribeiro, 2008).

Em 2007, a União da Indústria de Cana-de-açúcar (UNICA), a Organização de Plantadores de Cana da Região Centro-Sul do Brasil (ORPLANA) e a Secretaria do Meio Ambiente do Estado de São Paulo assinaram um protocolo agroambiental que, dentre outras coisas, estabelece novos prazos para extinção da queima da palha da cana-de-açúcar nos canaviais paulistas, sejam eles: 2014 para áreas mecanizáveis e 2017 para áreas não mecanizáveis.

A queima da palha da cana-de-açúcar pré-colheita é realizada para tornar menos cansativo e perigoso o corte manual, pois elimina a folhagem e a incidência de animais peçonhentos, e para aumentar a eficiência do trabalhador. Em geral a queima da palha da cana-de-açúcar está relacionada ao corte manual enquanto que a colheita sem o uso do fogo - cana crua - está

associada à utilização de máquinas colheitadeiras. Número de trabalhadores, qualificação profissional, custo de operação e logística da colheita, carga tributária, variáveis ambientais e produtividade são fatores relacionados ao tipo de colheita adotado pelo produtor e que dependem, em parte, das condições socioeconômicas dos municípios.

Em síntese, pressupõe-se que relacionado ao tipo de colheita estão presentes condições sócio-econômicas diversas, as quais podem apontar para uma proporção de tipo de colheita maior ou menor dependendo da realidade municipal. Logo, em uma tentativa de entender quais fatores influenciam o manejo da colheita nos municípios do Estado de São Paulo, este trabalho visa correlacionar as condições sócio-econômicas aos tipos de colheita por meio de técnicas de estatística, análise espacial e geoprocessamento.

2. Área de Estudo e Material Utilizado

A área de estudo compreende toda extensão territorial deste estado o qual localiza-se, aproximadamente, entre os paralelos 19° 05' e 25° 00' sul e os meridianos 44° 00' e 53° 05' oeste e está inserido na região sudeste do Brasil, com uma área total de 248.209,426 km², contendo 645 municípios e um total de população estimado, em 2007, em torno de 39.827.570 de pessoas, ou seja, concentra mais de 20% de toda a população do país (IBGE, 2009). Além de ser o estado responsável por grande parte do PIB brasileiro, este concentra, atualmente, cerca de 60% do total de área plantada de cana-de-açúcar. Por razões ambientais, a plantação de cana-de-açúcar não se encontra presente na região sul do estado (Figura 1).

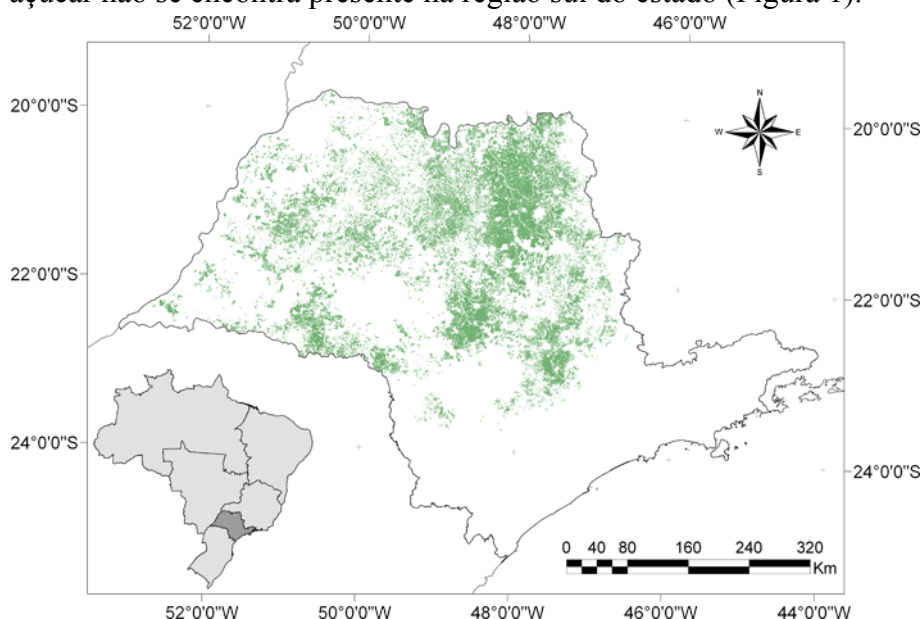


Figura 01 – Localização das áreas canavieiras no estado de São Paulo em 2008.

Para este trabalho foram utilizados dados relativos à estimativa de área de cana-de-açúcar para o estado de São Paulo realizado pelo projeto CANASAT (2008), além de dados do Censo 2000 (IBGE) e do SEADE (2005).

Foram utilizados como auxílio ferramental os seguintes softwares: TerraView 3.2.0 e SPRING 4.3.3 para processamento do Banco de dados e para extração de índices de autocorrelação espacial. Além de pacotes estatísticos como S-PLUS 4.5 e o GeoDa 0.95i para a realização de regressão espacial.

3. Metodologia

O banco de dados e as análises foram dispostos para dois períodos de colheita: os anos de 2006 e 2007, em todos os municípios do Estado de São Paulo que contam com plantação de cana.

3.1. Construção das variáveis

Para a área colhida com o uso do fogo, dá-se o nome de CANA-QUEIMA e, sem o uso desse, considera-se CANA-CRUA. Os dados de área de colheita de cana-de-açúcar são dados em hectares. Portanto, para facilitar a manipulação dos mesmos, foi necessário construir um índice que represente uma proporção de colheita por fogo ou sem o uso deste, a saber:

$$I_c = \frac{(\%A) - (\%B)}{100} \quad (1)$$

onde:

IC = Indicador de Colheita ($-1 \leq IC \leq 1$)

A = Quantidade de cana-crua

B = Quantidade de cana-queima

O indicador de colheita é tratado aqui como a variável resposta ou dependente. As variáveis explicativas ou independentes foram selecionadas a fim de relacioná-las a primeira. A princípio foram escolhidas as seguintes:

- PIB per capita em reais (SEADE, 2005);
- % de empregados na agricultura (SEADE, 2006 e 2007);
- % de empregados na indústria (SEADE 2006 e 2007);
- Rendimento médio na agricultura em reais (SEADE 2006);
- Rendimento médio na indústria em reais (SEADE 2006);
- Índice de Desenvolvimento Humano municipal (IDHm) – 0 a 1 - (IBGE, 2000);
- IPRS, variável discreta que corresponde a níveis sintéticos de riqueza, longevidade e instrução, dividida em 5 níveis, sendo o nível cinco, de maior desenvolvimento. (SEADE, 2005);
- Riqueza, Longevidade e Instrução em índices de valores contínuos (SEADE, 2005).

3.2. Análise exploratória

A fim de entender como o fenômeno se distribui no espaço é necessário realizar uma análise exploratória da variável resposta, isto é, o índice de colheita, para detectar se há presença de aglomerados. Algumas técnicas de análise espacial mostram-se eficientes para este objetivo como, por exemplo, o uso da média móvel espacial. Esta é uma forma simples de explorar a

variação da tendência espacial dos dados calculando a média dos valores dos vizinhos, reduzindo a variabilidade espacial calculado a partir da matriz normalizada de proximidade espacial W como expressa a fórmula a seguir:

$$\hat{\mu}_i = \frac{\sum_{j=1}^n W_{ij} y_j}{\sum_{j=1}^n W_{ij}} \quad i = 1, 2, \dots, n \quad (2)$$

onde:

W_{ij} é a matriz de proximidade.

y_i é o valor do atributo em cada área.

n é o número de polígonos (áreas).

São apresentadas nas figuras 02 e 03 os mapas de colheita para o período de 2006 e 2007 aplicando a técnica da média móvel espacial.

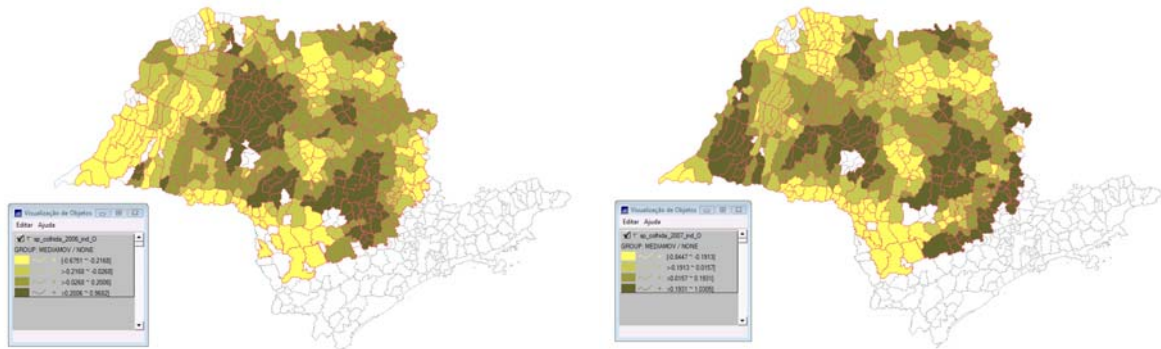


Figura 02 – Média Móvel Espacial – Colheita 2006

Figura 03 – Média Móvel Espacial – Colheita 2007

Ao remover flutuações aleatórias do índice através de sua vizinhança pode se detectar valores mais altos (maior proporção de cana-crua) no eixo centro-sul do estado onde já são áreas mais consolidadas de plantação de cana-de-açúcar e índices mais baixos (maior proporção de cana-queima), principalmente, no oeste de São Paulo no período de 2006. Nesta mesma área verifica-se uma parcial mudança em 2007 evidenciando algum impacto da adoção das diretrizes do protocolo ambiental citado anteriormente.

A verificação de dependência espacial foi realizada através do Índice de Moran, o qual é um indicador global de autocorrelação espacial e mostra como os valores estão correlacionados no espaço (Druck et al., 2005)

O Índice de Moran apresenta valores variando de -1 a 1, sendo que um índice negativo significa correlação inversa, ou seja, o objeto analisado tem valor baixo e seus vizinhos, altos e vice-versa. Valores positivos significam que o objeto analisado tem valor alto assim como seus vizinhos. Valores próximos de zero (0) indicam não-correlação espacial. Este Índice é calculado segundo a fórmula a seguir:

$$I = \frac{n}{\sum_{i=1}^n \sum_{j=1}^n w_{ij}} \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (3)$$

Outra forma de compreender a distribuição dos objetos analisados se dá através do diagrama de Moran como é mostrado nas figuras 04 e 05:

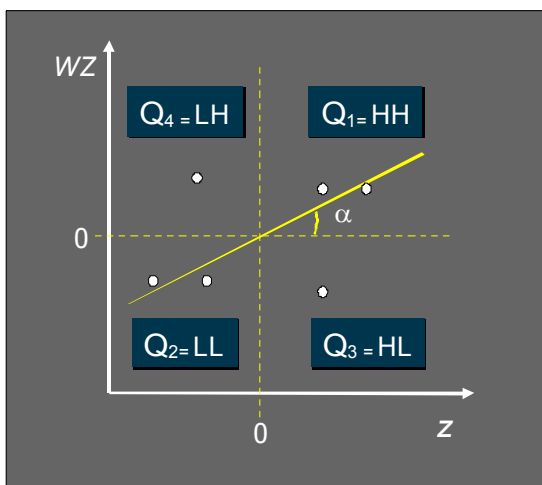


Figura 04 – Representação do Digrama de Moran.

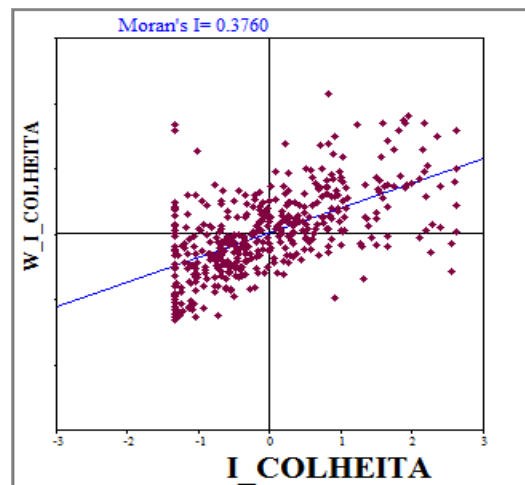


Figura 05 – Exemplo de Diagrama de Moran: Índice de colheita no período de 2006.

Câmara et al. (2004b) alertam que, após o cálculo do índice de Moran, é necessário sua validação. Isto pode ser realizado através do teste de pseudo-significância, no qual são geradas diferentes permutações dos valores de atributos associados às regiões; cada permutação produz um novo arranjo espacial, em que os valores são redistribuídos entre as áreas. A hipótese implícita do cálculo do índice de Moran é que os dados são estacionários de primeira e segunda ordem.

Para os períodos analisados, o Índice de Moran confirmou autocorrelação espacial para o fenômeno cujos valores foram 0,3760 e 0,3687 para 2006 e 2007, respectivamente. A partir da localização dos objetos analisados em cada quadrante, pode-se gerar um mapa (BoxMap) correspondente ao número destes como é mostrado nas figuras 06 e 07.

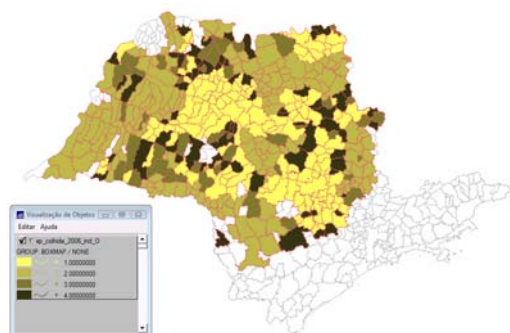


Figura 06 – Box Map – representação do Índice de Moran (2006).

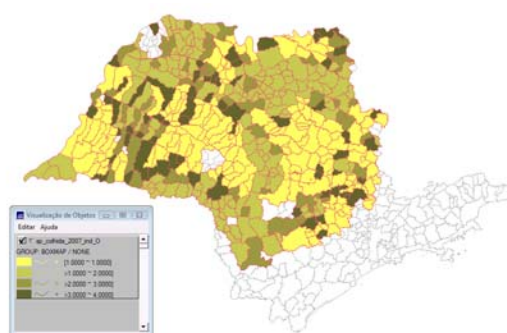


Figura 07 – Box Map – representação do Índice de Moran (2007).

Para o período de 2006 percebe-se segundo o BoxMap assim como a Média móvel local uma concentração de mais altos índices de colheita na região centro-sul do estado. Índices correlacionados de baixos-índices de colheita – maior percentual de colheita pelo uso do fogo – na região norte e oeste do estado. Em 2007, nas regiões central e sul mantém índices mais alto, apresentando-se como áreas mais consolidadas de plantação de cana e uma maior presença de colheita sem o uso do fogo. Áreas do norte e meio oeste com agrupamentos menores de uso do fogo em relação ao ano anterior. Uma inversão de situação do modo de colheita do extremo oeste em relação ao ano anterior evidenciado pelo Boxmap possivelmente é explicado por uma frente de expansão em processo de transição. Estas áreas podem ser identificadas como aquelas de quadrante 3 e 4.

3.3. Regressão Linear

Uma forma de se relacionar o índice de colheita com uma variável sócio-econômica é através do método de regressão. Segundo Neter (1996), a regressão é uma técnica que descreve a relação entre duas ou mais variáveis através de uma função matemática. Antes de selecionar as melhores variáveis explicativas foi preciso eliminar aquelas redundantes. Este processo foi realizado através de uma matriz de correlação entre todas as variáveis para os dois períodos (Figuras 08 e 09).

	PIB. PC	AGRO. PERC	AGRO. REND	IDHM	RIQUEZA
PIB. PC	1.000000000	-0.08437037	0.20019694	0.009894945	0.40581817
AGRO. PERC	-0.084370369	1.00000000	0.16751998	-0.072462386	-0.21836453
AGRO. REND	0.200196938	0.16751998	1.00000000	-0.030866130	0.40670329
IDHM	0.009894945	-0.07246239	-0.03086613	1.00000000	0.09788768
RIQUEZA	0.405818173	-0.21836453	0.40670329	0.097887681	1.00000000
LONGEVIDAD	0.135272899	-0.13578961	0.10325213	0.065461361	0.17422423
INSTRUCAO	0.046440872	-0.08200250	0.05340731	0.083990828	-0.06832013
IND. COLHEI	0.024594136	0.10228609	0.05137683	0.060970594	0.04466456
	PIB. PC	AGRO. PERC	AGRO. REND	IDHM	RIQUEZA
PIB. PC	1.000000000	-0.08437037	0.20019694	0.009894945	0.40581817
AGRO. PERC	-0.084370369	1.00000000	0.16751998	-0.072462386	-0.21836453
AGRO. REND	0.200196938	0.16751998	1.00000000	-0.030866130	0.40670329
IDHM	0.009894945	-0.07246239	-0.03086613	1.00000000	0.09788768
RIQUEZA	0.405818173	-0.21836453	0.40670329	0.097887681	1.00000000
LONGEVIDAD	0.135272899	-0.13578961	0.10325213	0.065461361	0.17422423
INSTRUCAO	0.046440872	-0.08200250	0.05340731	0.083990828	-0.06832013
IND. COLHEI	0.024594136	0.10228609	0.05137683	0.060970594	0.04466456

Figura 08 – Matriz de correlação entre variáveis - 2006

	PIB.PC	AGRO.REND	IDHM	RIQUEZA	LONGEVIDAD
PIB.PC	1.000000000	0.193172849	0.004393316	0.39273779	0.11870073
AGRO.REND	0.193172849	1.000000000	-0.037995371	0.40547281	0.10509648
IDHM	0.004393316	-0.037995371	1.000000000	0.08333660	0.05292766
RIQUEZA	0.392737792	0.405472815	0.083336597	1.00000000	0.16760600
LONGEVIDAD	0.118700729	0.105096476	0.052927664	0.16760600	1.00000000
AGRO.PERC	-0.089902513	0.170854259	-0.065025038	-0.21096346	-0.14493323
INSTRUCAO	0.051138486	0.052612600	0.078868638	-0.06349051	0.19791180
IND.COLHEI	0.033034785	-0.004921475	0.013558089	0.05431418	-0.03281345
	AGRO.PERC	INSTRUCAO	IND.COLHEI		
PIB.PC	-0.08990251	0.05113849	0.033034785		
AGRO.REND	0.17085426	0.05261260	-0.004921475		
IDHM	-0.06502504	0.07886864	0.013558089		
RIQUEZA	-0.21096346	-0.06349051	0.054314179		
LONGEVIDAD	-0.14493323	0.19791180	-0.032813454		
AGRO.PERC	1.00000000	-0.09074434	0.010446317		
INSTRUCAO	-0.09074434	1.00000000	0.061436516		
IND.COLHEI	0.01044632	0.06143652	1.000000000		

Figura 09 – Matriz de correlação entre variáveis - 2007

Diante de bastante número de variáveis foram selecionadas 7 das 10 iniciais. Estas foram PIB per capita, percentual de empregados na agricultura, rendimento médio na agricultura, IDH municipal, longevidade e instrução. Assim, foi realizado a regressão linear simples uma a uma a fim de verificar quais indicadores poderiam se mostrar significativos para o modelo.

Neste tipo de regressão nenhum componente da função matemática leva em conta a estrutura espacial dos dados, supondo pelo modelo que são independentes espacialmente (Bailey & Gatrell, 1996). A equação que caracteriza o modelo é:

$$y = \beta_0 + \beta_1 x + \varepsilon \quad \text{sendo que } \varepsilon \sim N(0; \sigma^2) \quad)4)$$

onde:

β_0 : intercepto

β_1 : inclinação

Y = Variável independente;

X = Variável dependente;

ε = erro aleatório

Como condição para realizar a regressão a variável dependente ou resposta (Y) deve ter distribuição normal. Desta forma verificou-se através do gráfico de aderência e o teste de Kolmogorov-Smirnov. Como a distribuição não é normal, à mesma foi realizada uma transformação logarítmica. Os resultados são apresentados nas Figura 10 e 11:

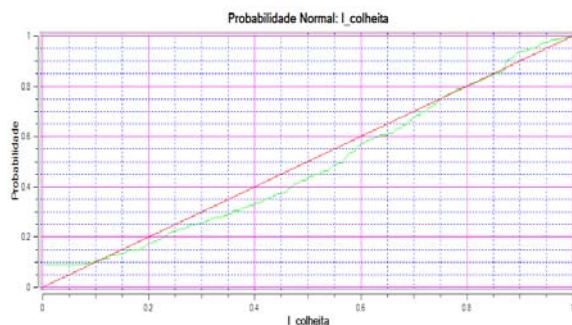


Figura 10 – Curva de aderência após transformação em Log. Teste Kolmogorov-Smirnov = 0,09 (0,05 de significância)

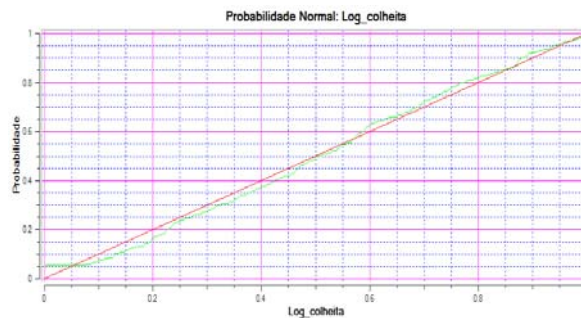


Figura 11 – Curva de aderência após transformação em Log. Teste Kolmogorov-Smirnov = 0,0029 (0,05 de significância)

Atendida às condições da Regressão as variáveis foram ordenadas segundo suas significâncias e, assim foram selecionadas as variáveis IDH municipal e renda média na agricultura e então feita uma regressão linear múltipla, isto é, leva em consideração mais que uma variável explicativa como a fórmula a seguir:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \epsilon \quad \text{sendo que } \epsilon \sim N(0; \sigma^2) \quad)5)$$

O resultado da regressão múltipla com estas variáveis para os dois períodos estão apresentadas nas Figuras 12 e 13:

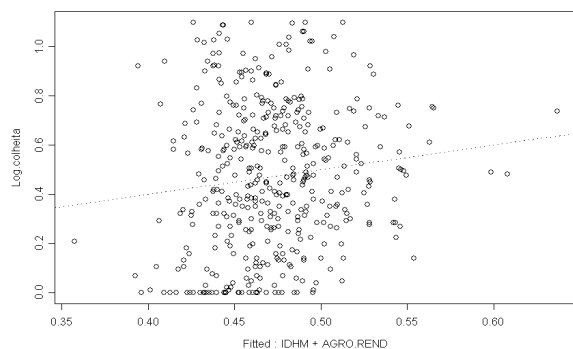


Figura 12 – Regressão linear múltipla Y=Índice de Colheita; X= IDH municipal + Rendimento médio na agricultura em 2006. $R^2 = 0,013669$; Log likelihood = -83,612; Akaike = 173,224

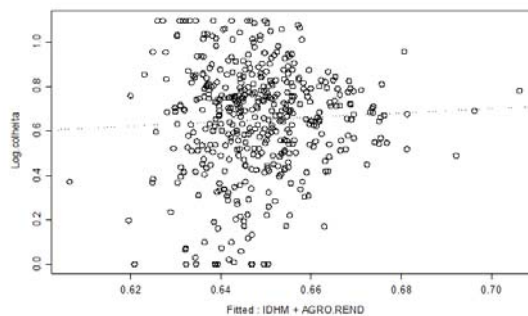


Figura 13 – Regressão linear múltipla Y=Índice de Colheita; X= IDH municipal + Rendimento médio na agricultura em 2006. $R^2 = 0,002433$; Log likelihood = -7,75115; Akaike = 21,5023

A qualidade da regressão é pode ser avaliada pelo coeficiente de determinação (R^2), que corresponde a percentagem que as variáveis independentes (IDHm e renda na agricultura) podem explicar a variável dependente ou resposta (Índice de colheita). Segundo os resultados deste coeficiente para os dois anos, neste modelo de regressão a variável independente explica muito pouco o fenômeno analisado.

Outra suposição do modelo de regressão é inferir os valores dos resíduos e constatar que os mesmos tenham distribuição normal e que a variância seja constante, bem como verificar a presença de outliers que estejam prejudicando o modelo (Neter, 1996). Assim foram feitas análises de presença de outliers (Cook Distance), teste de heterocedasticidade (teste de Breusch- Pagan) e normalidade (teste Jarque-Bera e teste Kolmogorov-Smirnov) para os resíduos. Os resultados são apresentados nas Figuras 14 e 15:

TEST	VALUE	P-valor
Jarque-Bera	22.60708	0.0000123
Kolmogorov-Smirnov	0.0714	0
Breusch-Pagan	28.25303	0.0000007

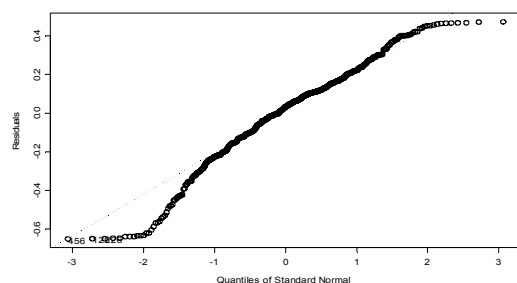


Figura 14a – Testes de normalidade e de variância para 2006

TEST	VALUE	p-valor
Jarque-Bera	15.43109	0.0004458
Kolmogorov-Smirnov	0.0489	0.0149
Breusch-Pagan	12.35233	0.0020784

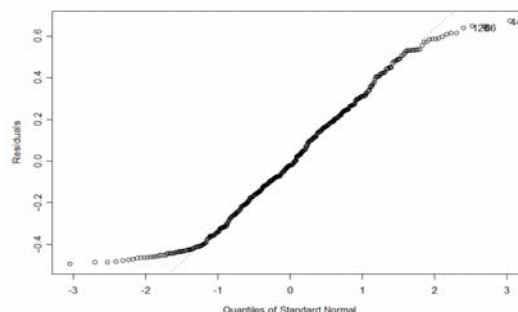


Figura 14b – Testes de normalidade e de variância para 2007

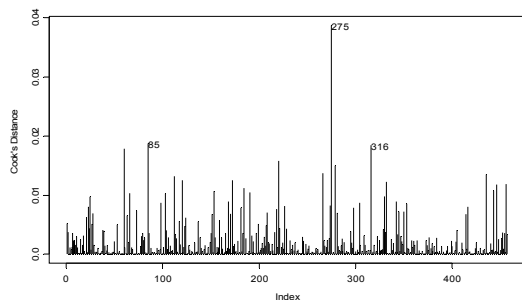


Figura 15a – Gráfico de Cook Distance para 2006

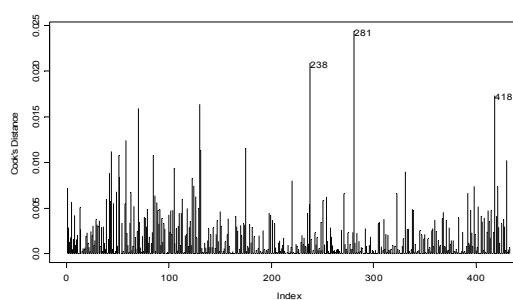


Figura 15b – Gráfico de Cook Distance para 2007

Ao analisar os gráficos e resultados dos testes pode-se afirmar que há evidências de correlação entre as variáveis, porém, as mesmas não estão sendo bem explicadas pelo modelo de regressão linear. Uma possível causa para isto é que as variáveis explicativas podem estar tendo interferências devido à uma dependência espacial. Para verificar isto, foi feito análises de agrupamentos através do Índice Global de Moran para os resíduos. Os resultados para os anos de 2006 e 2007 apresentaram, respectivamente, 0,360212 e 0,379860. Estes valores comprovam a evidência de autocorrelação espacial que, logo, necessitaria de um modelo mais adequado que levasse a dependência espacial em consideração. Assim foi realizado uma Regressão Espacial com Efeitos Globais.

3.4. Regressão Espacial com Efeitos Globais – *Spatial Lag*

O modelo de regressão com efeitos espaciais globais supõe que é possível capturar a estrutura de correlação espacial em um único parâmetro adicionando ao modelo de regressão clássico (Druck et al., 2005). Considerando que a autocorrelação espacial é atribuído à variável resposta, como foi verificado na análise exploratória inicialmente, o modelo adotado foi o *spatial lag* cujo é expresso pela seguinte fórmula:

$$Y = \beta_0 + \beta_1 X + \rho WY + \varepsilon \quad (6)$$

onde:

W = matriz de correlação espacial;

WY = expressa a dependência espacial;

ρ = coeficiente espacial autoregressivo

Após a aplicação deste método verificou-se que os índices apontaram para um modelo mais ajustado o qual incorporou a dependência espacial. Nas Tabelas 1 e 2 abaixo é apresentado os resultados em comparação aos do modelo linear múltiplo.

Tabela 1 – Índices de Regressão para o período de 2006

	Regressão Linear Múltipla	Regressão Espacial <i>Spatial lag</i>
R ²	0,013669	0,369569
Log Likelihood	-83,612	4,84555
Akaike (AIC)	173,224	-1,69111

Tabela 2 – Índices de Regressão para o período de 2007

	Regressão Linear Múltipla	Regressão Espacial <i>Spatial lag</i>
R ²	0,002433	0,308261
Log Likelihood	-7,75115	55,4867
Akaike (AIC)	21,5023	-102,973

Segundo Druck et al. (2004), análise feitas a partir do coeficiente de determinação (R²) seriam incorretas portanto, deve-se verificar os valores de máxima verossimilhança (Log Likelihood) e critérios de informação (Akaike). Usualmente a comparação de modelos é feita utilizando o logartimo da máxima verossimilhança, que é o que possui o melhor ajuste para os dados observados. Já o critério de informação de Akaike (AIC) é expresso por:

$$AIC = 2 \times LIK + 2k \quad (7)$$

onde LIK é o log de máxima verossimilhança e k é o número de coeficientes de regressão. Segundo este critério o melhor modelo é aquele que possui um menor valor de Akaike e, por conseguinte um maior valor de LIK (Druck et al. 2004).

Verificando as Tabelas 1 e 2, pode-se confirmar tal suposição que o modelo de regressão espacial *spatial lag* foi que melhor explicou o fenômeno. Resta então analisar os resíduos a fim de confirmar que este não possuem correlação espacial. Para isto foi realizado, novamente, o teste de pseudo-significância Moran para os resíduos, apresentados nas Figuras 16 e 17.

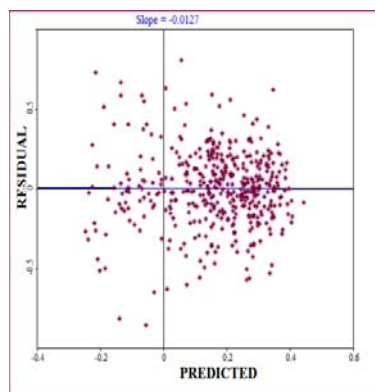


Figura 16a – Gráfico dos resultados preditos X resíduos mostrando aleatoriedade - 2006

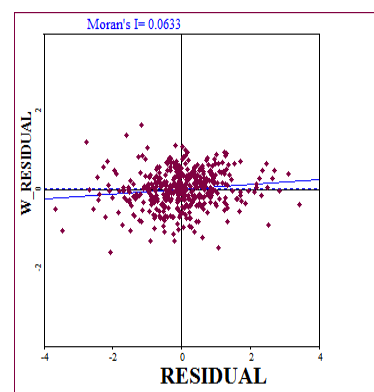


Figura 16b – Índice de Moran dos Resíduos (2006) = 0,0633

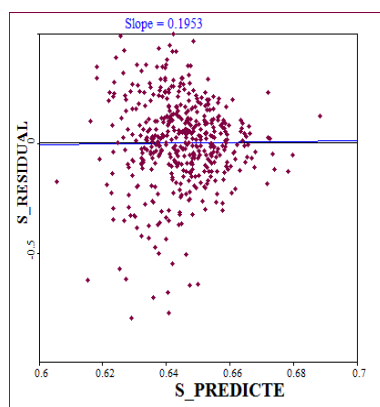


Figura 17a – Gráfico dos resultados preditos X resíduos mostrando aleatoriedade - 2006

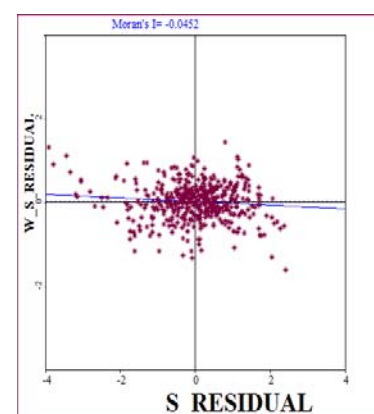


Figura 17b – Índice de Moran dos Resíduos (2006) = 0,0633

4. Resultados e Discussão

Os resíduos se apresentam, segundo o teste de Moran, sem dependência espacial e posicionados de forma aleatória confirmando que o modelo não possui tendência e se ajusta adequadamente ao tipo de regressão espacial. Aplicando a técnica de LisaMap para os

resíduos, com significância de 0,05%, verifica-se a ausência de agrupamentos espaciais por região (Figura 18 e 19).

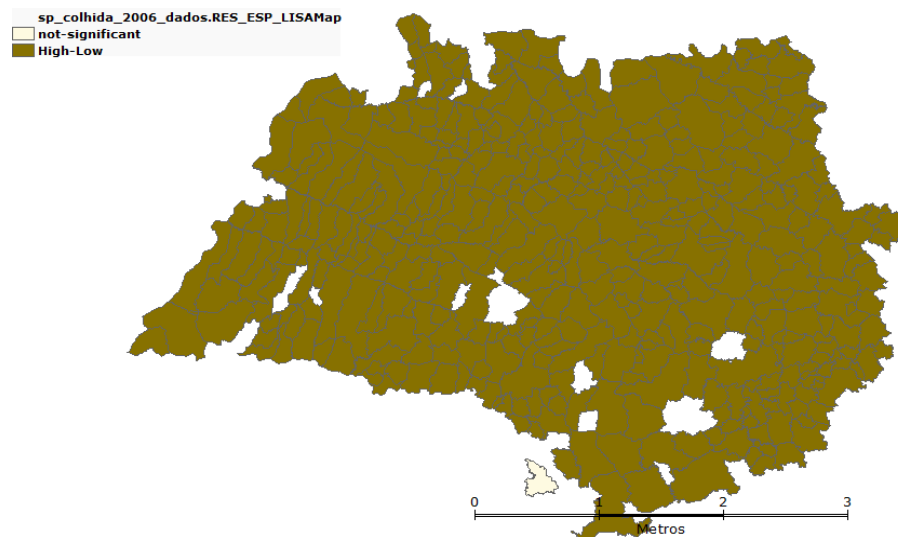


Figura 18 – Aplicação da técnica de LisaMap para os resíduos - 2006

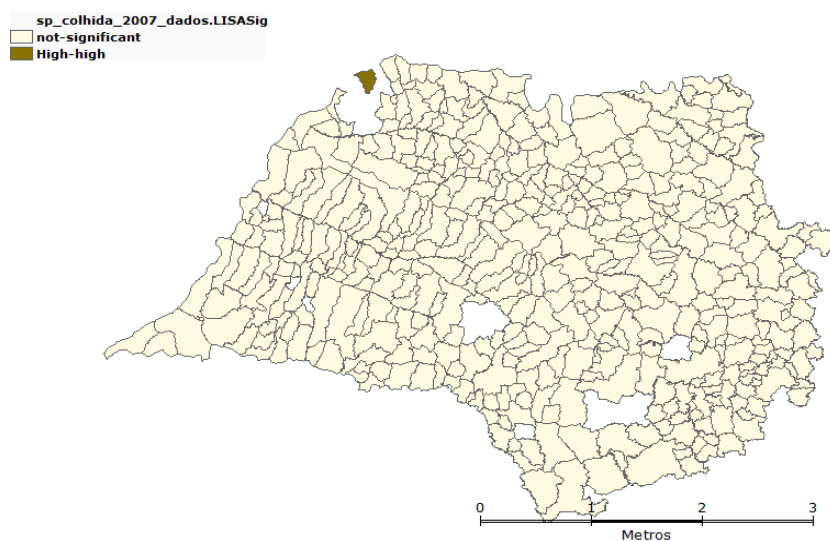


Figura 19 - Aplicação da técnica de LisaMap para os resíduos – 2007

Em relação às variáveis utilizadas, com a aplicação deste modelo, percebe-se que a entrada da variável “renda média na agricultura” deixou de ser significativa em 2006, deste modo podendo ser retirada pois não contribui para explicar o fenômeno enquanto que o Índice de desenvolvimento humano (IDH), torna-se mais relevante neste mesmo período.

Por outro lado, no período de 2007, ambas variáveis não se mostraram capazes de explicar o tipo da colheita com significância. Outras duas variáveis – IPRS e Instrução - foram utilizadas em seus lugares a fim que confirmasse a hipótese de que fatores sócio-econômicos contribuíssem para a explicação do tipo de colheita majoritária no município. Como é

apresentado na Tabela 3, ambas variáveis contribuem de maneira significativa para o modelo no período de 2007.

Tabela 3 – Modelo de Regressão Espacial – *Spatial Lag* - 2007

Variable	Coefficient	Std.Error	z-value	Probability
W_LOG_HEITA	0.6410678	0.0445408	14.39282	0.0000000
CONSTANT	0.2991687	0.04231146	7.070632	0.0000000
IPRS	-0.02086697	0.008561489	-2.437306	0.0147971
INSTRUCAO	0.002643359	0.001388178	1.904193	0.0568849

5. Considerações Finais

O modelo por regressão espacial demonstrou que a variável resposta possui forte dependência espacial. O uso desta técnica permitiu evidenciar que o manejo da colheita da cana sem o uso do fogo possa ocorrer em municípios com níveis sócio-econômico mais elevados em parte retratados através do IDH municipal.

Todavia, os indicadores sócio-econômicos não são o bastante explicar o tipo de colheita. Outras variáveis como oscilações no mercado especulativo, variáveis topográficas e ambientais servem para complementar o estudo. A prática da queima em um estado de grande extensão territorial como São Paulo é promovida por fatores complexos que não ocorrem de maneira homogênea em todo o espaço.

Contudo, este trabalho evidenciou que a prática ou não da queima possui padrões espaciais de aglomeração no espaço. Trabalhos futuros são motivados em observar essas ocorrências na interação espaço-tempo. Sugestões são indicadas em analisar a taxa de aderência ao protocolo ambiental por parte dos municípios ao longo dos últimos anos.

Apesar de algumas diferenças de significância das variáveis explicativas entre os dois períodos analisados, no ano de 2007 torna-se visualmente evidente a tendência de adequação dos fazendeiros aos novos protocolos ambientais.

6. Referências

- Bailey, C. & Gatrell, A. C. **Interactive spatial data analysis**. Longman. England. 1996
- CANASAT, 2008. Disponível em: www.inpe.br/canasat (acesso em novembro, 2008)
- Druck, S.; Carvalho, C.; Câmara, G.; Monteiro, A.M.V.M. **Análise de Dados Geográficos**. Brasília, EMBRAPA, 2004.
- Fundação Sistema de Análise de Dados (SEADE). Disponível em: <http://www.seade.gov.br/>. Acesso em novembro de 2008.
- Instituto Brasileiro de Geografia e Estatística, (IBGE), 2008. **Censo 2000**. Disponível em: www.sidra.ibge.gov.br (acesso em novembro, 2008).

_____. **Levantamento Sistemático da Produção Agrícola**. Disponível em: www.sidra.ibge.gov.br (acesso em novembro, 2008)

_____. **Malha Municipal do Brasil: Situação 2007**. Rio de Janeiro: IBGE, 2008.

Lopes, F. S.; Ribeiro, H. Mapeamento de internações hospitalares por problemas respiratórios e possíveis associações à exposição humana aos produtos da queima da palha de cana-de-açúcar no estado de São Paulo. **Revista Brasileira de Epidemiologia**, 2006, 9(2) p. 215-25.

Roseiro1, M. N. V. e Takayanagui, A. M. M. Meio ambiente e poluição atmosférica: o caso da cana-de-açúcar. **Saúde**, Vol. 30 (1-2). p. 76-83, 2004

Nassar, A. M.; Rudorff, B. F. T.; Antoniazzi, L. B.; Aguiar, D. A.; Bacchi, M. R. P.; Adami, M. Prospect of the sugarcane expansion in Brazil: impacts on direct and indirect land use changes. In: Zuurbier, P.; Vooren, J. **Sugarcane Ethanol: Contributions to climate change mitigation and the environment**. Wageningen Academic Publishers. Netherlands. 2008, 255 p.

Neter, J.; Kutner, M. H.; Nachtsheim, C.J.; Wasserman, W. **Applied Linear Statistical Models**. 3a. Ed. 1996.

Ribeiro, H. Queimadas de cana-de-açúcar no Brasil: efeitos à saúde respiratória. **Rev. Saúde Pública**, 2008, 42(2). p. 370-376.

Ribeiro, H.; Assunção, J. V. Efeitos das queimadas na saúde humana. **Estud Av.**, 2002,16(44). p. 125-48.

Secretaria do Meio Ambiente do Estado de São Paulo (SMASP). **Protocolo agroambiental**. São Paulo, 2008. Disponível em: <http://www.ambiente.sp.gov.br/cana/protocolo.pdf> Acesso em: novembro de 2008.

Smeetsa, E; Jungingera, M.; Faaija, A; Walterb, A; Dolzanb, P. Turkenburga, W. The sustainability of Brazilian ethanol - An assessment of the possibilities of certified production. **Biomass and Bioenergy**. 32 (2008). p. 781 – 813.

Veiga Filho, L. Sustainable energy ignites Brazil's economy. **Valor Econômico**, 2008. p. 6 – 11.