

Análise da expansão agrícola na bacia do rio Curuá-Una, no período de 2004 à 2014.

Lidiane Cristina Oliveira Costa¹

¹ Instituto Nacional de Pesquisas Espaciais - INPE
Caixa Postal 515 - 12227-010 - São José dos Campos - SP, Brasil
lidiane.costa@inpe.com.br

1. Introdução

A partir de meados de 1960 ocorre uma transformação no setor agrícola do país, uma modernização incentivada pela expansão do crédito rural subsidiado destinado à aquisição de insumos modernos, gerando aumento nas taxas de crescimento da produção. Na década de 70, além da consolidação dessa modernização na área agrícola, foram criadas instituições de ensino, pesquisa e extensão rural, com destaque para a fundação da Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA) (CONCEIÇÃO & CONCEIÇÃO, 2014).

Nessa mesma época na Amazônia, estratégias de ocupação começam a serem implementadas, primeiramente com a inserção das rodovias Belém-Brasília e Cuiabá-Porto Velho, e posteriormente com o processo de desenvolvimento agrícola por meio da criação da Superintendência do Desenvolvimento da Amazônia (SUDAM) em 1966, que a partir de 1968 promove investimentos na região através de incentivos fiscais e créditos especiais direcionados a agropecuária. Tais estratégias de ocupação da Amazônia resultaram em um desmatamento acelerado e uma rápida apropriação de terras.

Nas últimas duas décadas o destaque é para os agronegócios, sendo que as duas principais empresas da agroindústria, segundo NUNES, 2007, Cargil e Bunge possuem sede em Santarém, um dos sete municípios que compõe a bacia do Curuá-Una.

Com isso, o objetivo deste trabalho é fazer uma análise de correlação espacial entre a expansão agrícola, utilizando técnica de regressão espacial e variáveis, tais como: relevo, tipo de solo, possibilidade de mecanização agrícola, distância de estradas, distância de Portos, proximidade ao corpo d'água, vegetação secundária, pasto, desmatamento e distância dos municípios de Santarém e Itaituba.

2. Área de estudo

A área de estudo deste trabalho é a bacia do Curuá-Una, afluente do Rio Amazonas, localizada na região Sudoeste do estado do Pará (Figura 1). A bacia compreende uma área de aproximadamente 31 mil km² e abrange parte dos municípios de Belterra, Uruará, Placas, Prainha, Medicilândia, Santarém e Mojuí dos Campos. Este último município, foi criado em 2013, e sua área foi desmembrada do município de Santarém. Na porção Leste da bacia localiza-se Floresta Nacional do Tapajós (Flona Tapajós) e, ao extremo Sul, compreendendo uma pequena parte de duas áreas de terra Indígena, Arara e Cachoeira Seca.

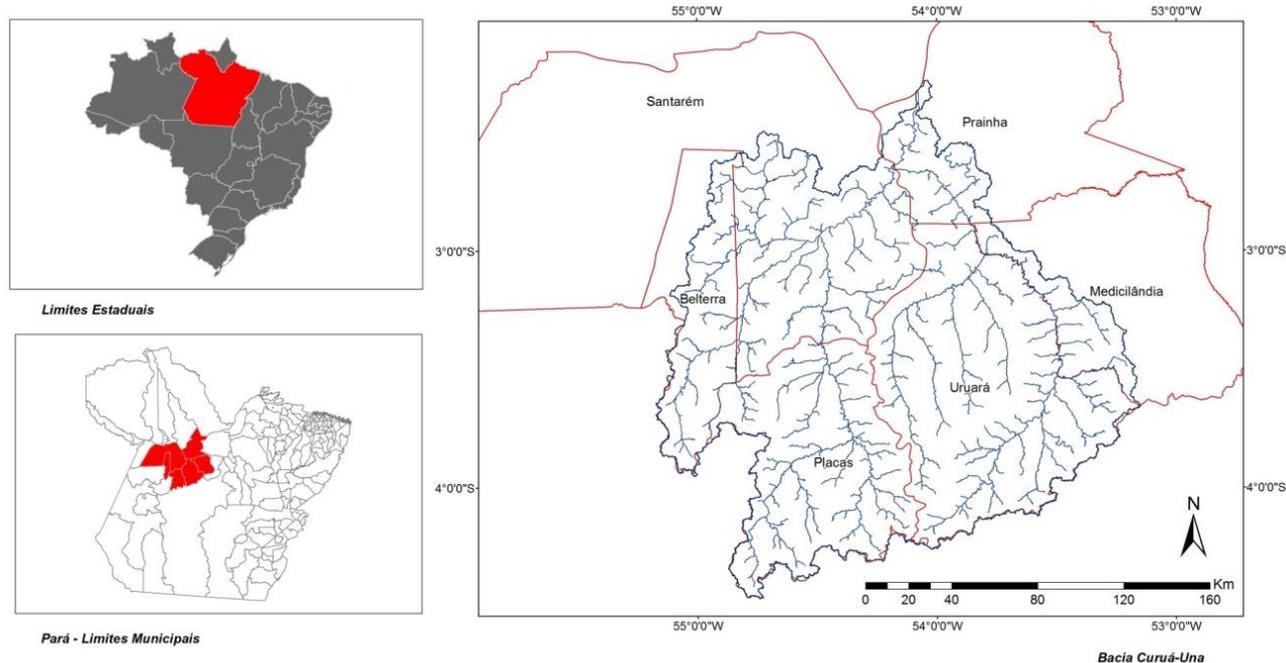


Figura 1 - Área de estudo.

3. Materiais

Foram utilizados dados de declividade, desmatamento, rodovias, hidrografia, uso e cobertura da terra, solo, portos e cidade polo, que para esse trabalho é composta de Santarém e Itaituba.

3.1 Variáveis

A variável dependente é o incremento da agricultura anual (I_{aa}), para extraí-la foi utilizado a operação de conjuntos, aplicando a diferença, tal que:

$$I_{aa} = x : x \in A_{aa}^{2014} \text{ e } x \notin A_{aa}^{2004} \quad (1)$$

Onde, A_{aa}^{2014} é a área de agricultura anual do dado TerraClass 2014 e A_{aa}^{2004} .

Quanto as variáveis independentes, com exceção de porto, hidrografia, rodovia, e agricultura anual (2004) todas tiveram um processo de preparação. As cidades polos escolhidas foram Santarém e Itaituba, que no caso seriam as possibilidades para o escoamento da produção agrícola.

Para as variáveis vegetação secundária e pasto, foi realizada a operação de união dos dados para os anos de 2004, 2008, 2010 e 2012 do TerraClass. Para o desmatamento foi aplicada a mesma metodologia com os dados de incremento de desmatamento do PRODES dos anos de 2004 à 2013.

Para as variáveis que indicam a aptidão potencial natural agrícola para a produção de grãos, soja e milho, foram utilizados como base para elaboração das Instruções Normativas (EMBRAPA, 2008) e Ramalho (1995). Para declividade foi realizado um fatiamento a fim de obter as áreas com declividades inferiores a 5%. Para os solos, foram excluídos os apontados como inaptos, que seriam solos de perfil com profundidade inferior a 50cm, muito pedregosos ou alagável.

Foi realizado também um mapa quanto ao grau de impedimento à mecanização. Dados de solo e declividade compõem esse índice, para o caso foi considerando que onde o nível de impedimento à mecanização é nulo ou ligeiro há maior chances de ocorrer a expansão da agricultura anual. Portanto todos os demais níveis foram considerados inaptos a mecanização.

As variáveis independentes e a dependente foram construídas a partir de um plano celular de 3 x 3 km. A priori foi utilizado um retângulo envolvente que abrangesse as cidades polo e os portos para o cálculo das distâncias, e posteriormente realizado o recorte para a área da bacia do Curuá-Una. Para a extração das variáveis para o plano celular foi utilizado o preenchimento de células e operadores de porcentagem de classe por célula e distância mínima. A Tabela 1 apresenta as variáveis independentes, premissa, operação utilizada e informações sobre os dados.

Tabela 1- Apresentação das variáveis.

	Fator	Premissa	Variável	Unidade	Dado	Fonte
Uso e Cobertura da terra	Agricultura 2004	Células com maior área de agricultura anteriores, possuem maior potencial para expansão agrícola	p_agri2004	% célula	TerraClass	EMBRAPA/INPE
	Desmatamento	Células com maior área de desmatamento, possuem maior potencial para expansão agrícola	p_incdes	% célula	PRODES	INPE
	Pasto	Células com maior área de pasto, possuem maior potencial para expansão agrícola	p_pasto	% célula	TerraClass	EMBRAPA/INPE
	Vegetação Secundária	Células com maior área de vegetação secundária, possuem maior potencial para expansão agrícola	p_vegsec	% célula	TerraClass	EMBRAPA/INPE
Ambientais	Hidrografia	Células mais próximas de cursos d'água, tem maior potencial para expansão agrícola.	d_hidro	km	Hidrografia, 1:100000	ANA
	Declividade	Células com maior área de declividade inferior a 5%, tem maior potencial para expansão agrícola.	p_decliv_5	% célula	Declividade (%)	Topodata
Acesso ao mercado	Solo	Células com menor área de solos impróprios para produção de grãos, possuem maior potencial para expansão agrícola	p_soloin	% célula	Mapa de Solos	EMBRAPA, ZEE BR-163
	Porto	Células mais próximas dos portos, tem maior potencial para expansão agrícola.	d_porto	km	Porto	IBGE
	Cidade Polo	Células mais próximas das cidades polo, tem maior potencial para expansão agrícola.	d_sede	km	Sedes Municipais	IBGE
	Rodovia	Células mais próximas as rodovias, tem maior potencial para a expansão agrícola.	d_rodovia	km	Estradas	DNIT+IBGE
	Mecanização Agrícola	Células com maior área mecanizável, possuem maior potencial para expansão agrícola	Mec_agri0	% célula	Mapa de Solos + Declividade	EMBRAPA, ZEE BR-163, TOPODATA

3.2 Métodos

A metodologia desse trabalho foi dividida em x etapas: 1) Análise exploratória, através de regressão linear simples; 2) Regressão múltipla, observando um melhor modelo que explique a variável resposta; 3) Índice de Moran I, a fim de verificar se a regressão apresenta dependência espacial; 4) Regressão espacial.

3.2.1 Regressão Linear

A regressão linear simples tem o objetivo de verificar a existência de uma relação entre a variável dependente (Y) e a variável independente (X). E obter o melhor modelo baseado nas observações, segundo a Equação 1 (Peternelli, 2013).

$$Y_i = \beta_0 + \beta_1 \cdot X_i + \varepsilon \quad (1)$$

Onde: Y_i é a variável dependente; β_0 constante de regressão, que representa o intercepto da reta com o eixo Y; β_1 coeficiente de regressão, representa a variação de Y em função da variação de uma unidade da variável X; X_i representa a i-ésima observação da variável independente, ε é uma constante que representa os erros ou resíduos (Peternelli, 2013).

Neste presente trabalho a regressão linear foi utilizada para comparação dos melhores R^2 e p-valor. O valor de R^2 refere-se ao ajuste do modelo, p-valor ao seu nível de significância e.

A regressão múltipla diferencia-se da linear pelo número de variáveis dependentes, seguindo os mesmos pressupostos, mas compondo-se de três ou mais variáveis. Para avaliação do modelo, além dos índices citados anteriormente para a regressão linear simples é utilizado o índice de variância inflada (VIF), que diz sobre a multicorlinearidade do dado.

A construção do modelo de regressão múltipla pode se dá pelo método hierárquico, com base em trabalhos anteriores. Entrada forçada, onde todas as variáveis dependentes são “forçadas” simultaneamente no modelo. E por sim, a stepwise, com duas diferentes lógicas: backward e forward. Na primeira é inserido todas as variáveis independentes e posteriormente sendo retirada conforme o nível do contribuição e na segunda as variáveis são inclusas até que o nível da contribuição pare de subir significativamente conforme a adição de variáveis. (FEITOSA, 2014)

3.2.2 Índice de Moran I

Após aplicado o melhor modelo de regressão múltipla ao dados realiza-se o teste de Moran I no resíduos da regressão múltipla para verificação da dependência espacial. O índice varia de -1 a 1, sendo a correlação espacial negativa para valores de Moran (-), correlação alta, com índices acima de zero e a não existência de correlação espacial para valores zero.

O Índice de Moran indica a presença ou não da correlação espacial, mas não aponta qual o melhor modelo espacial a ser utilizado. Uma forma de identificar qual o tipo de modelo espacial aplicar é através do Teste de Lagrange. O teste através dos valores apresentados em seus resultados de Lmerr, Lmlag, RLMerr, RMLag e suas respectivas significâncias indica se o melhor modelo é o Spatial Error ou Spatial Lag.

3.2.3 Regressão espacial

Os modelos de regressão espacial acrescentam em seu formato duas variáveis que tratam das limitações existentes no outro modelo. O mesmo tem como premissa a existência de uma correlação espacial. O modelo Spatial Lag, assim como o Spatial Error apresentam índices globais para todo o objeto de estudo.

A Equação 2 apresenta o modelo do Spatial Lag Model:

$$Y_i = \beta \cdot X_i + \rho \cdot W \cdot Y_i + \varepsilon \quad (2)$$

Onde, W é a matriz de proximidade espacial, e o produto $W \cdot Y_i$ expressa a dependência espacial em Y_i e ρ é o coeficiente espacial autorregressivo. A hipótese nula para a não existência de autocorrelação é que $\rho = 0$.

4. Resultados e Discussões

O resultado da regressão linear simples para todas as variáveis independentes (Tabela 2) e o conhecimento prévio de correlação entre as variáveis Figura 2, norteou a escolha de quais delas entrariam para o modelo múltiplo. Também foi utilizado o VIF como parâmetro para exclusão de variáveis adicionadas ao modelo.

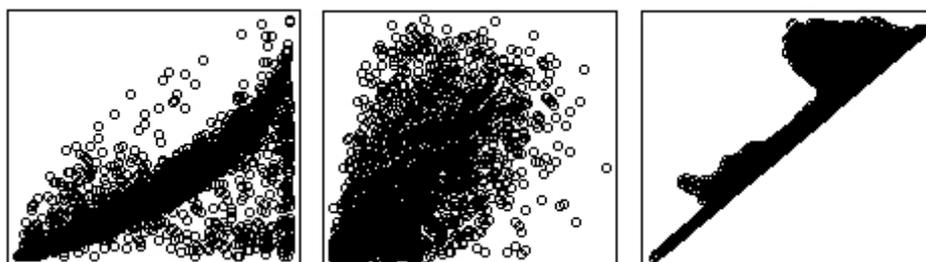


Figura 2 - Da esq. para dir.. Correlação de Declividade e Agricultura Mecanizada; Pasto e Vegetação Secundária; Cidade Polo e Porto.

Para composição do modelo linear múltiplo, seguiu-se a metodologia *stepwise, forward*. Iniciando com a $p_agri2004$ e acrescentando na ordem as demais variáveis: d_sede , p_decliv_5 , p_pasto e p_soloin . Algumas variáveis foram eliminadas por serem correlacionadas com outras que já estavam no modelo. A decisão de qual variável usar nos casos nos quais havia correlação foi verificar qual retornava um melhor ajuste, R^2 .

Tabela 2 - Resultados de R^2 e suas respectivas significâncias.

Variável	R^2	P-valor
Pasto	0,0427	<2,2e-16
Vegetação secundária	0,0620	<2,2e-16
Hidrografia	0,0010	0,0470
Porto	0,1318	<2,2e-16
Sede	0,1287	<2,2e-16
Rodovia	0,0396	<2,2e-16
Declividade (5%)	0,0590	<2,2e-16

Agricultura 2004	0,2456	<2,2e-16
Desmatamento	0,0038	0,0038
Solo Impróprio	0,0034	0,0003
Mecanização Agrícola	0,0487	<2,2e-16

Após o cálculo do modelo múltiplo (Figura 3) foi realizado o teste de Moran I, para verificar a possível existência da dependência espacial. O valor encontrado no Moran I foi de 0.3, o que indica a presença da dependência espacial. A Figura X, mostra o resultado do Índice de Moran I e seu respectivo teste de significância.

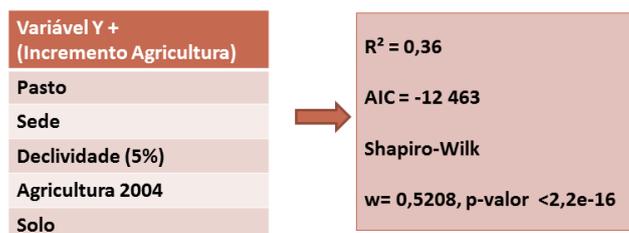


Figura 3 - Apresentação do resultado obtido no modelo linear múltiplo.

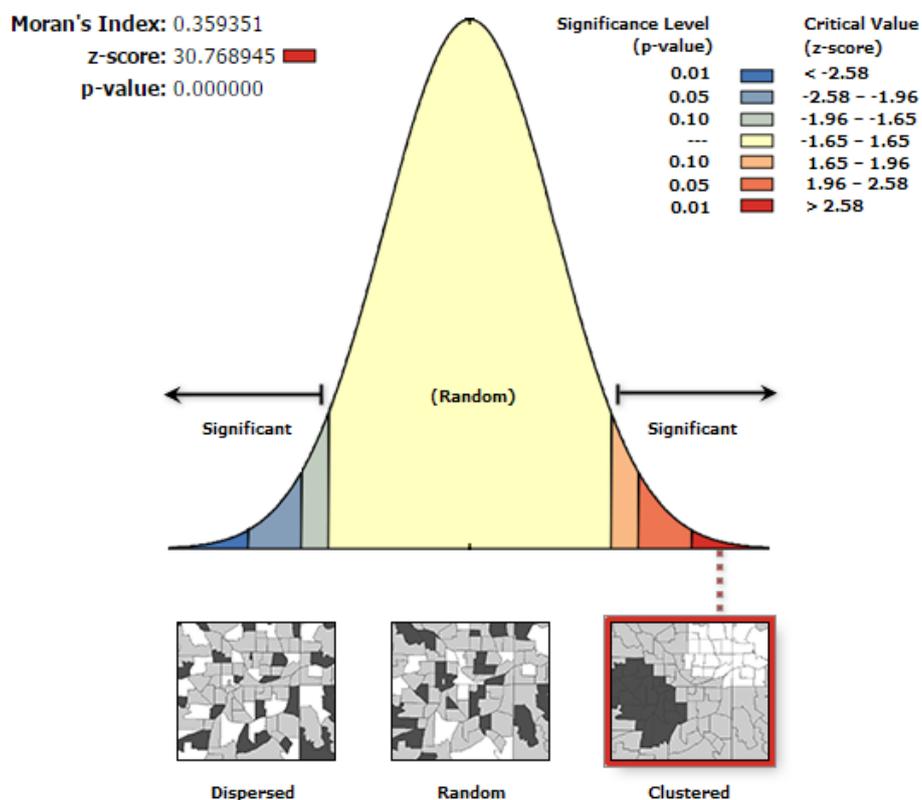


Figure 4- Estatística de Moran I.

Com a dependência espacial constatada, foi realizado o teste de Lagrange para indicação do melhor modelo espacial a ser utilizado. Conforme o resultado, Tabela 3, o modelo indicado é o Spatial Lag, a fim diminuir os resíduos e obter um melhor ajuste.

Tabela 3 - Teste de Lagrange

Teste LM		p-valor
Lmerr	1246,9	$2,2 \times 10^{-16}$
RLMerr	0,72607	0,3442
Lmlag	1615,2	$2,2 \times 10^{-16}$
RMlag	369,06	$2,2 \times 10^{-16}$

5. Considerações finais

O modelo obtido pela regressão múltipla linear apresentou um poder de explicação menor do que o esperado. Entretanto as variáveis que entraram para compô-lo são coerentes.

Ao aplicar o modelo de Spatial Lag, observou-se uma melhora em relação ao modelo linear que pode ser vista pelo AIC que foi de respectivamente de -13.443 para o -12.463. E literatura aponta que quanto menor o AIC melhor o ajuste.

Para outros estudos fica a sugestão de utilizar-se outros tipos de normalização e a inclusão de variáveis ainda não testadas.

Referência

ANSELIN, L. **An Introduction to Spatial Regression Analysis in R**. Illinois: University of Illinois, 2003. 25 p.

Becker, B. K. **Amazônia**. Ed. Ática, 1990.

Conceição, J. C. P. R.; Conceição, P. H. Z.. **Agricultura: Evolução E Importância Para A Balança Comercial Brasileira**. IPEA. Brasília, 2014.

MELLO, M. P.; PETERNELLI, L. A. **Conhecendo o R: uma visão mais que Estatística**. Viçosa, MG: Editora UFV, 2013. 222 p. ISBN 9788572694957. Disponível em: <http://www.dsr.inpe.br/~mello/Conhecendo_o_R.html>. Acesso em: 13 dez. 2016.

RAMALHO-FILHO, A.; BEEK, K. J. **Sistema de avaliação da aptidão agrícola das terras**. 3. ed. Rio de Janeiro: EMBRAPA-CNPS, 1995.