

Uso de Técnicas de análise espacial de dados: estudo de distribuição da soja no Estado do Maranhão

Ulisses Denache Vieira Souza ¹
Mauricio Eduardo Salgado Rangel ¹

¹ Universidade Federal do Maranhão - UFMA
Departamento de Geociências – DEGEO
Laboratório de Geotecnologias e Análise Espacial – GEOPRO
Caixa Postal 65-080-805 - São Luís - MA Brasil
ulissesdenache@hotmail.com
mauricio.rangel@ufma.br

Abstract. This study used a spatial geographic data areas, focusing on the spatial distribution of the soybean crop in the Maranhão state in the period from 2003 to 2012, based on Municipal Agricultural Production PAM / IBGE data, revenue budget and harvested area by municipality. We used the Grimson and Scan methods, with the aim of grouping the areas of the municipalities in regions with adjacent space and similarities. We sought to identify possible spatial clusters and evaluate the spatial relationship of adjacency between the municipalities with the largest areas of soybeans. The results indicated that the southern of Maranhão state had higher probability of spatial clustering of soybean, but from the year 2006 comes a new frontier in the northeast area of the state. It was concluded that these two areas are concentrated soybean in the state, with a significance level of 5%.

Palavras-chave: Detection Conglomerates, Spatial distribution, Scan Method. Detecção de Conglomerados, Distribuição espacial, Método Scan.

1. Introdução

As técnicas de análise espacial de dados geográficos em áreas têm proporcionado a inferência sobre fenômenos de diversos ramos das ciências, auxiliando pesquisadores e estudiosos a entenderem esses fenômenos que consideram o espaço enquanto categoria de análise, facilitando a inferência tanto sobre efeitos de 1º ordem como de 2º ordem. Dentre esses ramos que se beneficiam do uso destas técnicas está o estudo da dinâmica agrícola em determinadas áreas, relevante contribuição pode ser encontrada em Salgado (2009).

O estado do Maranhão tem se destacado no cenário agrícola nacional, enquanto produtor agrícola de soja. Isto se deve a proximidade de áreas produtoras com vias de escoamento, como a Ferrovia Carajás que transporta os grãos até o Porto do Itaqui localizado no município de São Luís, capital do estado do Maranhão. A produção agrícola Maranhense que já foi destaque no período colonial por cultivar arroz e algodão, tem se desenvolvido a partir da iniciativa de grupos sulinos que adquiriram terras menos valorizadas na área sul do estado e a partir do método da calagem tem realizado a correção e o uso da terra para o cultivo da soja. Como consequência, observa-se um aumento da produtividade durante todo período desde 2003 a 2012, período analisado neste trabalho.

Segundo os dados do IBGE referentes à Produção Agrícola Municipal (PAM), a área total de soja colhida no estado do Maranhão, no ano de 2003, era estimada em 275.252 ha, passando para 512.418 ha no ano de 2012. A distribuição espacial das áreas de concentração da cultura da soja possui um papel fundamental no planejamento das atividades do setor agrícola, permitindo estabelecer prioridades na alocação de recursos.

Neste contexto, o presente trabalho objetivou analisar a distribuição espacial da cultura da soja no estado do Maranhão durante o período de 2003 a 2012, baseando-se nos dados da Produção Agrícola Municipal PAM/IBGE de área colhida, em dados orçamentários, além de considerar inferências quanto ao incremento populacional e proximidades com as vias de escoamento da cultura da soja. Avaliando-se a relação espacial de adjacência e identificando a formação de conglomerado espacial.

2. Área de Estudo

O estado do Maranhão situa-se em uma posição de transição entre três macrorregiões brasileiras: Norte, Nordeste e Centro-Oeste, apresentando características climatológicas e especialmente fitogeográficas típicas destas macrorregiões (Figura 1). Com uma extensão territorial de 328.663 km², ocupando 3,28% do território brasileiro, sendo o segundo estado do Nordeste e o sétimo do Brasil em extensão territorial Brasil (1991).

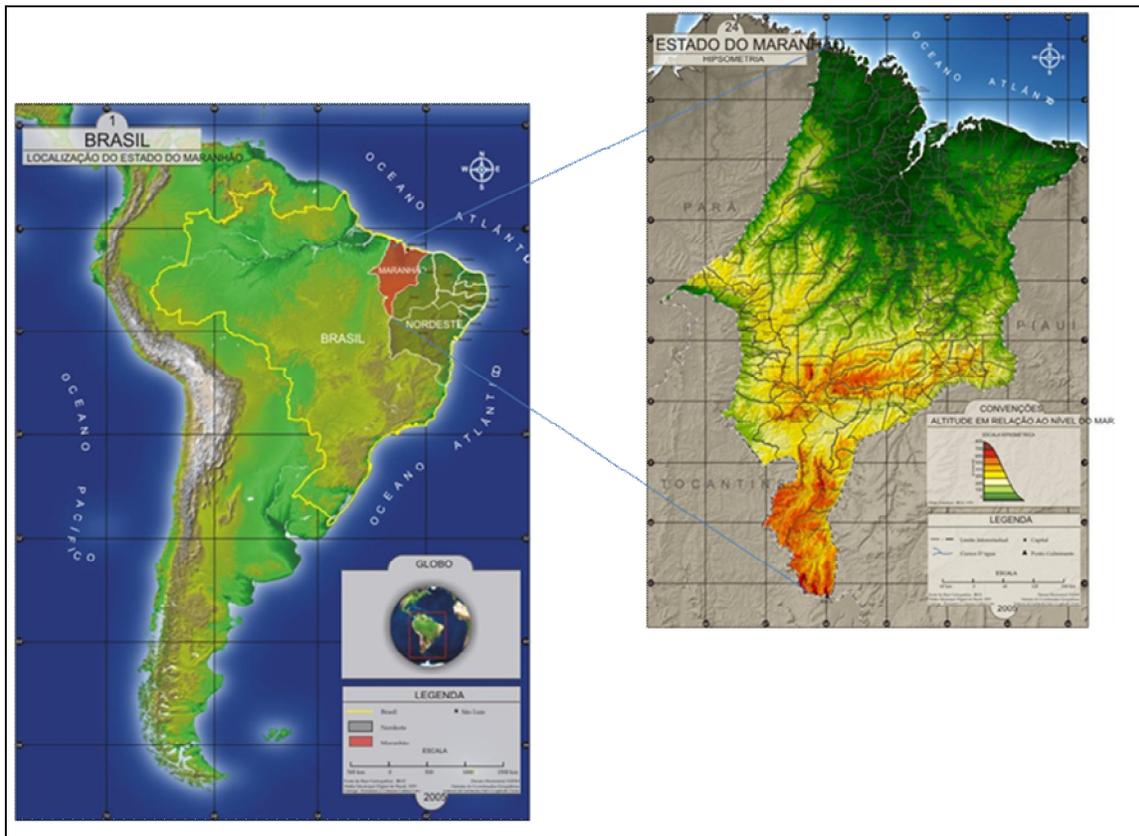


Figura 1. Situação geográfica do Maranhão
Fonte: FEITOSA e TROVÃO, 2006. (Adaptado por SOUZA, 2010)

3. Materiais e Métodos

Para realizar a análise da distribuição espacial das áreas de cultivo da soja, foram efetuadas consultas as bases de dados do governo federal, além da disponibilidade de softwares livres (*opensource*)¹ no intuito de verificar a dependência espacial entre os valores, de produção agrícola, de área colhida e de população rural observados entre os municípios maranhenses, os dados foram estruturados em um banco de dados geográficos utilizando-se o aplicativo TerraView (DPI/INPE) e o plugin TerraCluster (LESTE/UFMG). Como base fora utilizada a malha municipal digital de 2010 do estado do Maranhão disponível no item geociências do Instituto Brasileiro de Geografia e Estatística – IBGE em formato vetorial, facilitando a indexação dos dados tabulares da PAM, de receitas orçamentárias, de população rural, urbana e total às áreas geográficas (polígonos) dos municípios.

¹**opensource** – definição criada pela Open Source Initiative (OSI) identifica um programa de código aberto que deve garantir: distribuição livre, código fonte disponível, licença deve permitir trabalhos derivados e modificações em sua estrutura.

Para análise estatística dos dados utilizou-se o aplicativo R e os pacotes que estão instalados dentro deste aplicativo que possibilitaram análises diversas, em especial o pacote *spdep* Bivand et al. (2010), para estudos de regressão simples e espacial dos dados em análise neste trabalho. Para facilitar a compreensão dos procedimentos utilizados nesta abordagem elaborou-se o quadro esquemático abaixo com as etapas realizadas (Figura 2).

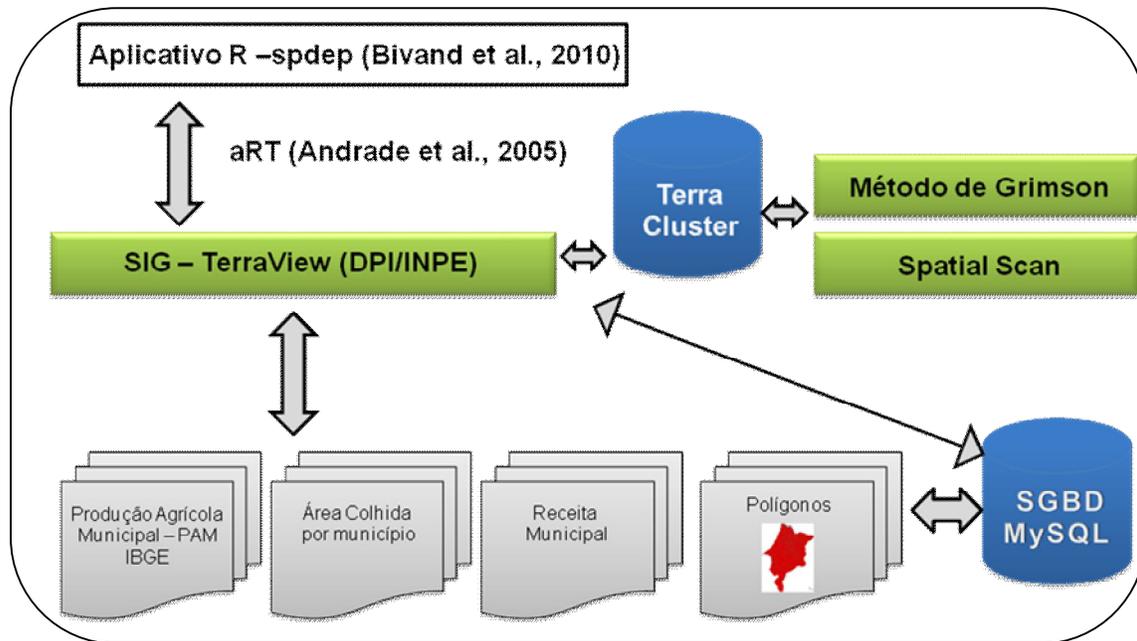


Figura 2. Procedimentos metodológicos

3.1 Análise Exploratória de dados espaciais: métodos e técnicas

Para uma visualização preliminar dos valores de produção agrícola (toneladas), área colhida (hectares) e quantidade populacional de cada município para o período de 2003 a 2012, período determinado pela disponibilidade de dados conjuntos já tabulados para análise estatística, foram utilizadas técnicas de análise exploratória de dados espaciais, elaborando-se gráficos, como o *scatterplot* entre as variáveis, utilizando o aplicativo R, tendo sido também utilizado o software TerraView para elaboração do Box map.

No que se trata da dependência espacial e verificação da correlação espacial dos valores dos atributos em estudo, seguiu-se as indicações de Bailey & Gatrell (1995), que propõem o cálculo do indicador global de autocorrelação espacial denominado Índice Global de Moran, que trata a autocorrelação espacial como uma medida de quanto o valor observado de um atributo em uma região é independente dos valores nas localizações vizinhas (Figura 3).

Como o índice global é uma medida genérica, torna-se importante examinar a autocorrelação espacial em escala local, para que isso fosse possível utilizou-se o indicador de autocorrelação espacial “Local Indicators of Spatial Association” (LISA) também denominado de Índice Local de Moran, apresentado por Anselin (1995). Ao se utilizar este indicador local em conjunto com o indicador global, tem-se a possibilidade de melhorar o conhecimento sobre os processos que dão origem a dependência espacial, refinando a análise.

Para os resultados dos indicadores global e locais utilizou-se uma simulação de distribuição por permutação aleatória (i.e. 999 permutações) nos valores dos atributos, com significância de 5% Anselin (1995).

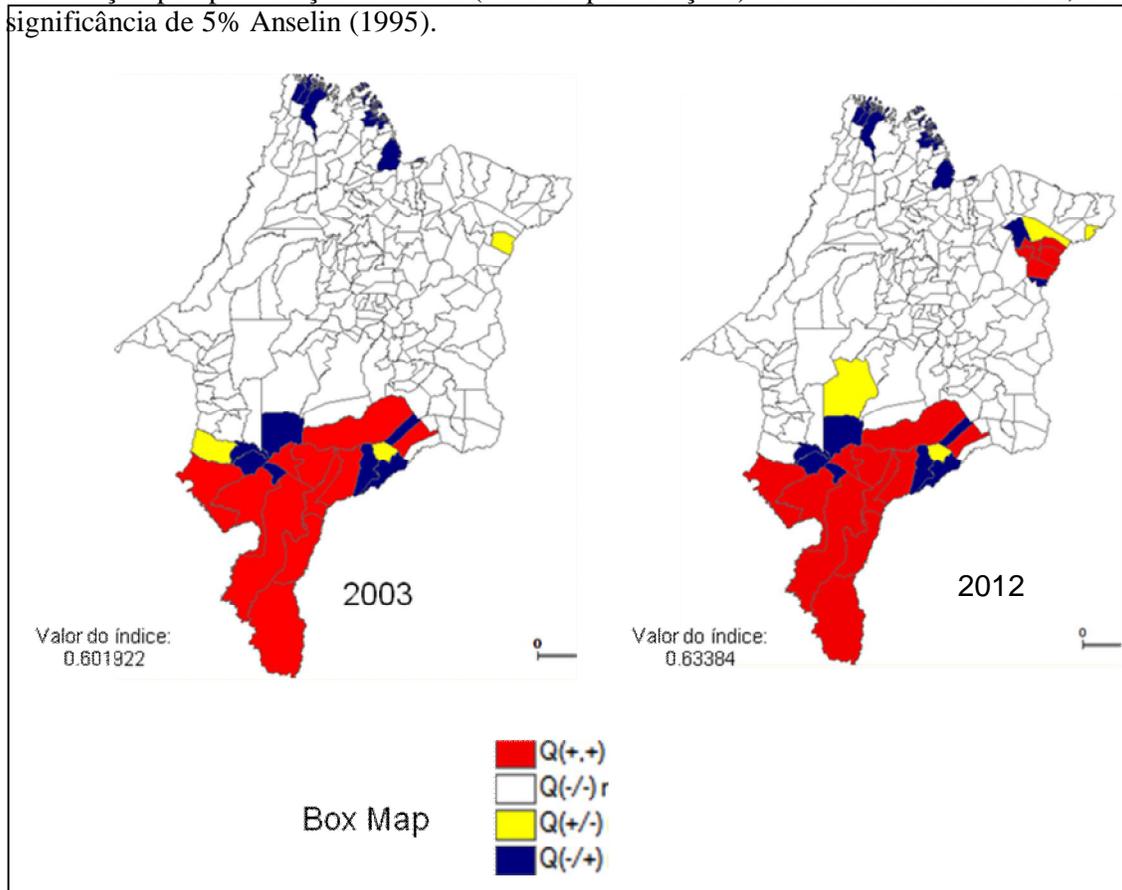


Figura 3: Produção Agrícola 2003 e 2012.

3.2 Conglomerados espaciais: Detecção e análise

Para verificar o quanto os municípios com os maiores valores dos atributos analisados neste trabalho apresentam relações de adjacência espacial não aleatória e podem ser considerados como agregados, fora empregado o método de Grimson (1989), neste método a estatística de teste corresponde à contagem do número de municípios com os maiores valores de área colhida considerados adjacente. A aleatoriedade da distribuição espacial dos municípios é a hipótese nula, ou seja, não há relação de adjacência entre os municípios de se levado em consideração o atributo escolhido. A hipótese alternativa corresponde à adjacência não aleatória destes municípios. Para avaliar a significância assume-se que sob a hipótese nula a distribuição é Poisson ou Normal Grimson (1989). A probabilidade de se obter um valor da estatística maior ou igual ao observado, sob a hipótese nula, é dada pelo p-valor calculado, no software este método está implementado e ao ser calculado apresenta uma janela conforme demonstrado abaixo com todos os parâmetros da estatística do método (Figura 4).

Visando identificar uma região com maior probabilidade de aglomeração espacial, empregou-se o método de varredura espacial Spatial Scan (Kulldorff & Nagarwalla, 1995; Kulldorff, 1997), utilizado para identificar o conglomerado espacial que provavelmente viola a hipótese nula de não aglomeração. O método avalia um conjunto Z de círculos distintos, correspondentes aos círculos posicionados nos centróides das diferentes áreas do mapa e com

raios arbitrários r variando de zero a um limite máximo especificado, cada um deles representando um potencial conglomerado.

A estatística de teste do método Spatial Scan é definida como sendo a razão de máxima verossimilhança sobre todos os possíveis conglomerados no conjunto Z , sendo $L(z)$ a máxima verossimilhança para o possível conglomerado z , e L_0 é a função de verossimilhança sob a hipótese nula. A significância do conglomerado detectado é testada através do p-valor da estatística de teste χ obtido via simulações de Monte Carlo.

Como uma etapa subsequente se propõe uma regionalização das áreas indicadas pelos métodos de Grimson e Scan, tendo a finalidade de agrupar as áreas dos municípios em regiões que apresentam adjacência espacial e semelhanças segundo os atributos propostos para análise, (Produção Agrícola Municipal, Área colhida de soja para os anos de 2003 a 2012, para tal aplicação utilizou-se a técnica de regionalização espacial via árvore geradora mínima, denominado método de Skater (Assunção et al, 2004). O método Skater considera a homogeneidade entre os grupos de regiões gerados pelo processo de agrupamento de forma a minimizar a variância interna do grupo, permitindo a geração de um mapa regionalizado de acordo com o critério de agregação das áreas baseado na relação de vizinhança entre os objetos, a aplicação deste método permitiu identificar possíveis evoluções na produção de soja e de área colhida ao longo, dos últimos anos.

4. Resultados e Discussões

Com base nos métodos e nas técnicas empregados, pode-se observar que os municípios que apresentam significativos valores quanto aos atributos, produção agrícola municipal e área colhida apresenta relação de adjacência espacial de acordo com a estatística do método de Grimson, os parâmetros do método foram adquiridos através do software TerraView.

Quanto ao método de Grimson ao analisar o p-valor obtido (1.45152×10^{-06}), pode-se verificar que não há razão para acreditar que os 29 municípios, que são os que apresentaram os maiores valores do atributo tanto de produção agrícola municipal como de área colhida, possuem uma distribuição aleatória, a estatística adquirida do método leva a considerar que há relação de adjacência entre os atributos anexados aos municípios. Ao analisar a evolução da produção agrícola municipal e de área colhida dos municípios maranhenses, percebe-se que a partir do ano de 2008 identifica-se uma área de expansão de produção de soja (Figura 6).

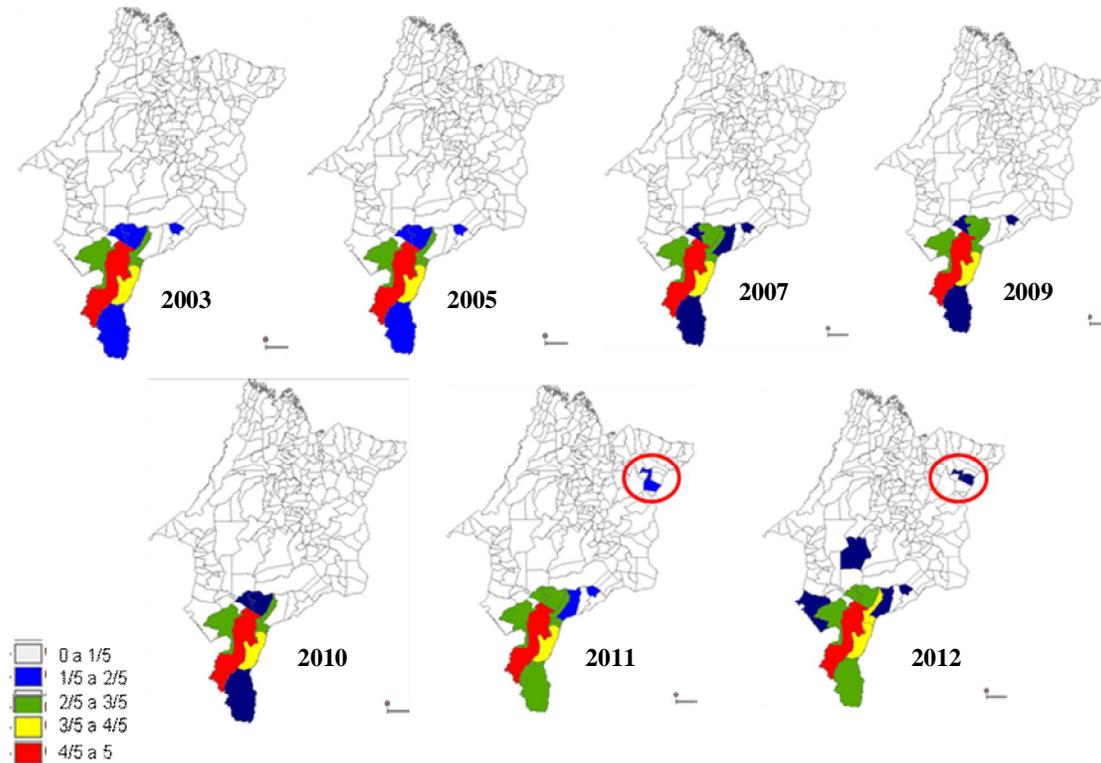


Figura 6. Evolução no período de 2003 a 2012.

Sendo assim, o método de Grimson foi capaz de indicar para essa análise que ao se considerar os municípios que possuem valores dos atributos, produção agrícola municipal e área colhida de soja, acima de um quinto do valor máximo do atributo observado (para cada ano) exemplificado na figura anterior, apresentaram relação de adjacência espacial ao longo de todo o período estudado, estando discriminados quanto a sua evolução de 2003 a 2012. Ao considerar a abordagem do método de Grimson, os municípios que foram elencados na figura anterior representam os municípios que possuem as maiores áreas de cultivo da soja no estado do Maranhão, tendo assim apresentado relação de dependência espacial com significância acima de 5%, assim estes municípios indicam regiões de concentração (não aleatória) da cultura da soja.

No intuito de fortalecer as informações referentes à relação de dependência espacial adotou-se conforme descrito no item métodos e técnicas o método de varredura Spatial Scan para identificar as regiões que apresentam aglomeração espacial tendo como base o período estudado. Ao analisar os resultados obtidos com a utilização do método de Varredura Scan, percebe-se a variação dos diferentes padrões de aglomeração durante o período em estudo, porém para facilitar à análise e tornar mais clara à importância deste resultado optou-se por demonstrar os resultados para os anos de 2003, 2006 e 2009 onde foi possível visualizar as regiões referentes aos conglomerados espaciais que provavelmente violam a hipótese nula de não aglomeração (aglomeração aleatória), estes anos também demonstram a expansão da frente agrícola da soja pelo estado do Maranhão, fornecendo um diagnóstico eficaz para o monitoramento dessa expansão, com seus benefícios e seus prejuízos dentro do contexto regional (Figura 7).

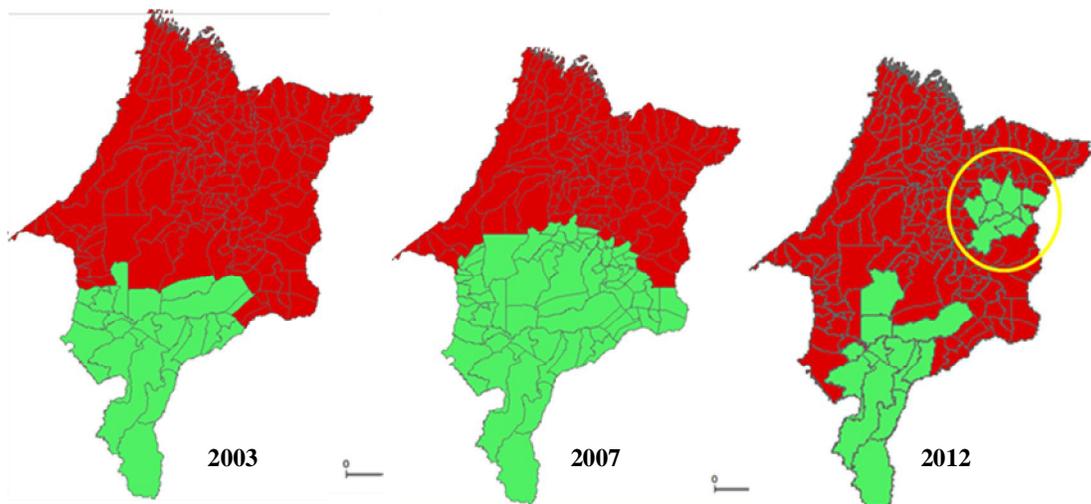


Figura 7. Padrões de aglomeração espacial em 2003, 2007 e 2012.

Analisando a Figura 7 é possível observar que o crescimento das áreas cultivadas com soja na região sul do Maranhão no ano de 2006 contribuiu para a modificação do padrão de aglomeração espacial, porém após o ano de 2006 houve outra tendência espacial que pode ser explicada pelo desgaste dos solos na região sul e pelos atrativos oferecidos pela área nordeste do estado, como subsídios agrícolas, terras mais baratas e ampla disponibilidade de mão de obra, alterando os padrões de aglomeração espacial, mostrando que municípios como Anapurus, Brejo, Nina Rodrigues e outros próximos passam a integrar a lista de municípios maranhenses com relevante produção agrícola de soja.

Para identificar os municípios com maior autocorrelação espacial para cada ano do período de estudo utilizou-se o indicador local LISA, tendo se observado que municípios que apresentaram valores de LISA acima de 0,5 com significância de 5% variaram ao longo dos anos estudados. Ao considerarmos o período de 2003 a 2012, os municípios de maior autocorrelação espacial e também mais frequentes foram: Balsas, Tasso Fragoso e Alto Parnaíba.

O indicador local de autocorrelação espacial LISA possibilitou a localização dos municípios associados às regiões de maior concentração de produção agrícola municipal e área colhida no que se refere à cultura da soja de soja ao longo do período de estudo, já no que se trata de identificar as regiões de aglomeração, essa técnica não foi eficaz, pois só conseguiu identificar municípios isolados. Sendo a utilização do método Skater uma das possibilidades adotadas por este estudo, para identificar e regionalizar esses municípios.

O método Skater permitiu gerar um mapa regionalizado em cinco grupos, possibilitando a identificação dos grupos de municípios com características próximas no que se trata de valores referentes aos atributos: Produção Agrícola Municipal e área colhida de soja para o período de 2003 a 2012. Este método consegue estabelecer esses grupos com base na relação de vizinhança entre municípios e o grau de homogeneidade interna dos grupos de municípios.

No entanto, verificou-se que ao se estudar a distribuição de soja no estado do Maranhão, esta produção está restrita a áreas que apresentam características propícias a essa produção, tais como: a área sul do estado com suas chapadas e áreas de cerrado, tabuleiros e clima favorável, assim como a área nordeste com seus subsídios e disponibilidade de terras mais baratas se comparadas ao restante do estado. Os métodos empregados na identificação dos aglomerados e das relações de adjacência auxiliaram a identificação de algumas

problemáticas, como a sobreposição das áreas de expansão da soja nos últimos dois anos pesquisados sobre as áreas de proteção ambiental do estado do Maranhão (Figura 8).

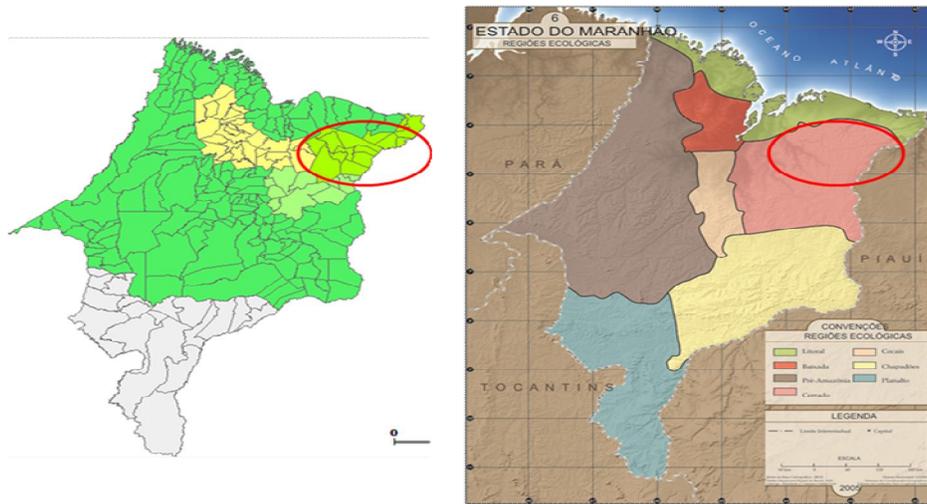


Figura 8. Comparação área de expansão e áreas de cerrado.

Ao se analisar os resultados percebe-se que estes indicam um crescente aumento da produção de soja sobre as áreas de cerrado do estado situadas na porção nordeste do estado diferente da região de concentração da cultura da soja no início da década passada que se instalara principalmente na porção sul do estado do Maranhão. Essa informação quanto à distribuição espacial da cultura da soja favorece ações de planejamento, no que diz respeito às atividades do setor agrícola, complementando planos de desenvolvimento que já estejam estruturados para essas áreas além de ações de mitigação de possíveis impactos que venham a ser causados pela cultura da soja para áreas de expansão atuais.

4. Conclusões

Neste trabalho analisou-se a distribuição espacial da cultura da soja no estado do Maranhão para o período de 2003 a 2012, tendo por referência os dados da Produção Agrícola Municipal PAM/IBGE e de área colhida para a cultura da soja, considerando ainda dados de receita orçamentária de cada município na fase de análise exploratória dos dados.

A partir dos resultados e discussões com base nos métodos e nas técnicas empregadas foi possível concluir que a soja tem ampliado seu leque de áreas no estado do Maranhão e que essa cultura se expande principalmente por áreas de cerrado, uma tendência que precisa ser analisada e monitorada, pela importância que o ecossistema cerrado tem para o estado e para o país. Com efeito, se devem realizar novas análises para poder estabelecerem de fato indicadores das condições desta distribuição pelo estado do Maranhão, considerando a influência das ferrovias e rodovias como influentes aspectos na distribuição da soja assim como na distribuição de outras culturas, tais aspectos podem ser constatados ao se observar o mapa de transportes disponibilizado pelos órgãos nacionais que cuidam do setor, que mostram que as rodovias estão cortando essas áreas.

Referências Bibliográficas

ALVES, D. S. Space-time dynamics of deforestation in Brazilian Amazonia. **International Journal of Remote Sensing**, v. 23, n. 14, p. 2903-2908, Jul 2002.

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

ASSUNÇÃO R.M., LAGE J.P. E REIS E.A. Análise de Conglomerados Espaciais via Árvore Geradora Mínima. **Revista Brasileira de Estatística**, v. 63, n. 220, p. 7-24, 2004.

BAILEY, T. C.; A. C. GATRELL. **Interactive Spatial Data Analysis**. New York: Wiley. 413 p. 1995.

BIVAND, R. et al. spdep: **Spatial dependence: weighting schemes, statistics and models. R package version 0.5-24**. Disponível em: <http://CRAN.R-project.org/package=spdep>.

BRASIL. Instituto Brasileiro de Meio Ambiente e dos Recursos Naturais. Secretaria de Meio Ambiente e Turismo do Estado do Maranhão. **Diagnóstico dos problemas ambientais do Estado do Maranhão**. São Luís, LITHOGRAPH, 1991.

FEITOSA, Antonio C.; TROVÃO, José R. **Atlas Escolar do Maranhão: Espaço histórico e cultural**. João Pessoa: Editora Grafset, 2006.

GRIMSON, R.C. Assessing patterns of epidemiologic events in space-time. In: PUBLIC HEALTH CONFERENCE ON RECORDS AND STATISTICS, 1989, Hyattsville, **Proceedings...** Maryland: National Center for Health Statistics, 1989.

KULLDORFF, M., NAGARWALLA, N. Spatial disease clusters: detection and inference. **Statistics in Medicine**, v. 14, p. 799–810, 1995.

KULLDORFF, M.A. A spatial scan statistic. **Communications in Statistics: Theory and Methods**. v. 26, p. 1481–1496, 1997.

SALGADO, M. P. G. **Análise da Distribuição da Soja no Estado do Mato Grosso**. Trabalho de conclusão da disciplina de análise espacial. Turma 2009. Disponível em: <<http://wiki.dpi.inpe.br/doku.php?id=ser301>> Acesso em: 08 de out. 2012.

TERRAVIEW . Terraview 3.5.0. São José dos Campos, SP: INPE, 2010. Disponível em: www.dpi.inpe.br/terraview>. Acesso em: 10 nov. 2012.