

Caracterização espaço-temporal da evolução das áreas municipais semeadas com soja nos municípios do Norte no Brasil

Victor Hugo Rohden Prudente (139181)
Pós-Graduação em Sensoriamento Remoto
Instituto Nacional de Pesquisa Espaciais, São José dos Campos – SP

Resumo: Para este estudo pretende-se caracterizar a nível municipal a da soja nos municípios do Norte do Brasil. Nesta busca, foram utilizados dados de área semeada e municipal da base PAM-IBGE (Produção Agrícola Municipal) entre os anos de 1990 a 2016. A espacialização se deu por meio de indicadores locais de autocorrelação espacial em cada ano e por meio da aplicação da estatística espaço-temporal SatScan. Encontrou-se agrupamentos em municípios ao oeste do estado do Tocantins, a leste do estado de Rondônia e ao nordeste do estado do Pará. As proximidades destes com regiões de maior tradição na sojicultura pode ter beneficiado ao desenvolvimento agrícola dessas regiões, porém mais estudos são necessários para qualquer análise conclusivas dos motivos da formação destes padrões. De maneira geral, as técnicas utilizadas apresentaram bons resultados para identificar os padrões de distribuição espacial para a soja nos municípios do Norte do Brasil.

Palavras-chaves: Satscan, padrões espaciais, agricultura, Moran

1 Introdução

A soja, inicialmente auxiliada pelo trigo, foi a grande responsável pelo surgimento da agricultura moderna no Brasil (Embrapa Soja, 2004; FAOSTAT, 2016). Este fato auxiliou para a criação de novas fronteiras agrícolas. A dinâmica agrícola são processos espacialmente condicionantes, onde o resultado em um lugar é afetado pelos eventos em outro lugar (Clemente et al., 2017).

As estatísticas espaciais abordam os efeitos da dependência e heterogeneidade espacial, incluindo métodos que incorporam informações de localizações, tais como coordenadas geográficas de lugares ou o polígono da região de estudo (Clemente et al., 2017). Alguns estudos usaram de estatística espacial de área para estudar padrões de agrupamentos com dados municipais de diversas culturas agrícolas (Clemente et al., 2017; Marconato, Larocca, & Quintanilha, 2012; Wrublack, Prudente, Sílvia, & Coelho, 2013)

A estatística espacial de área é uma das técnicas ligadas à análise exploratória de dados espaciais (ESDA - *Exploratory Spatial Data Analysis*). A análise ESDA têm por objetivo descrever os padrões de associação espacial (*clusters*), verificar a existência de diferentes regimes espaciais (não-estacionariedade) e identificar observações atípicas (i.e. *outliers*) (Almeida, Perobelli, & Ferreira, 2008; Maranduba Júnior & Almeida, 2009)

Existem vários métodos para se proceder às análises espaciais de áreas, entretanto o Índice Global, ou *I* de Moran, e o Índice Local de Associação Espacial (*Local Index of Spatial Association* - LISA) são os mais difundidos e amplamente usados em diversos campos. Estes, por meio da matriz de peso (*W*) utilizada para definir vizinhos, determinam se há autocorrelação espacial entre os dados (Anselin, 1995). Para o desenvolvimento destes índices, há três elementos básicos que descrevem as associações espaciais: matriz de proximidade espacial (ou de vizinhança) (*W*), vetor de desvios (*Z*) e vetor de médias (*Wz*) (Marconato et al., 2012).

O I de Moran é um índice que expressa autocorrelação espacial, indicando a existência ou não de áreas agrupadas (*clusters*), cujos atributos se apresentam de maneira semelhante no espaço. Ou seja, é uma medida cujo resultado indica se os dados se distribuem ou não de forma aleatória no espaço (Anselin & Lozano-Gracia, 2008; Novaes, Rudorff, Almeida, & Aguiar, 2011).

I de Moran fornece um único valor de associação como medida para todo o conjunto de dados, podendo esconder três distintos padrões locais de autocorrelação espacial. Com valor I de Moran global insignificante, pode ocorrer indicações de autocorrelação espacial local insignificante, positiva ou negativa. Um I de Moran global positivo (negativo), pode ocultar autocorrelação espacial local negativa (positiva) e insignificante (Perobelli, Almeida, Alvim, & Ferreira, 2007).

Para evitar uma conclusão equivocada sobre a autocorrelação espacial existente tendo apenas como base o I de Moran, pode-se desagregar tal índice global de acordo com seus constituintes locais, índice *LISA* e diagrama de dispersão de Moran. O índice de *LISA*, uma decomposição da medida global, produz valor específico para cada área geográfica indicando significância estatística de formação de conglomerados de áreas geográficas semelhantes ou isoladas em relação a determinado atributo (Perobelli et al., 2007; Predebon, Mathias, Aidar, & Rodrigues, 2010).

Os índices anteriormente descritos tem por função encontrar padrões espaciais de distribuição das variáveis. Entretanto, não possuem a capacidade de encontrar padrões espaço-temporais. Neste sentido se destaca a técnica SatScan, capaz de encontrar padrões espaço-temporais, ou somente temporais ou somente espaciais das variáveis (Block, 2007; Kulldorff, 2015).

Isto é feito por meio de um escaneamento gradual, usando uma janela no tempo e/ou espaço, notando o número de observações e/ou o valor associados a essas observações com os valores esperados dentro do círculo. A técnica SatScan forma um intervalo no tempo, um círculo ou uma elipse no espaço, um cilindro com base circular ou elíptica no espaço-tempo. Múltiplos tamanhos de janela são utilizados, sendo que a janela com a máxima probabilidade é o cluster mais provável, ou seja, o cluster menos provável de ser devido ao acaso. Um valor p é atribuído a este cluster (Kulldorff, 2015).

Esta ferramenta é muito utilizada para estudos que envolvem distribuição espacial de casos de doenças (Selemani et al., 2015; Sugumaran, Larson, & DeGroot, 2009; Wu, Guo, Guan, Sun, & Zhou, 2011; Zhao et al., 2013), sendo que pouco difundido na aplicação para estudos envolvendo variáveis do meio agrícola.

2 Objetivo

Diante do exposto, este trabalho tem por objetivo analisar as evoluções espaço e espaço-temporal dos valores de Área Municipal Semeada com Soja (AMSS) para os municípios do Norte do Brasil entre os anos de 1990 a 2016.

3 Material e Métodos

3.1 Área de estudo

À área de estudo compreende os 450 municípios da região norte do Brasil. Estes dividem-se em 7 estados (Figura 1). De maneira geral, os municípios possuem grande extensão territorial, atingindo em alguns casos extensões maiores que estados brasileiros e até mesmo outros países.

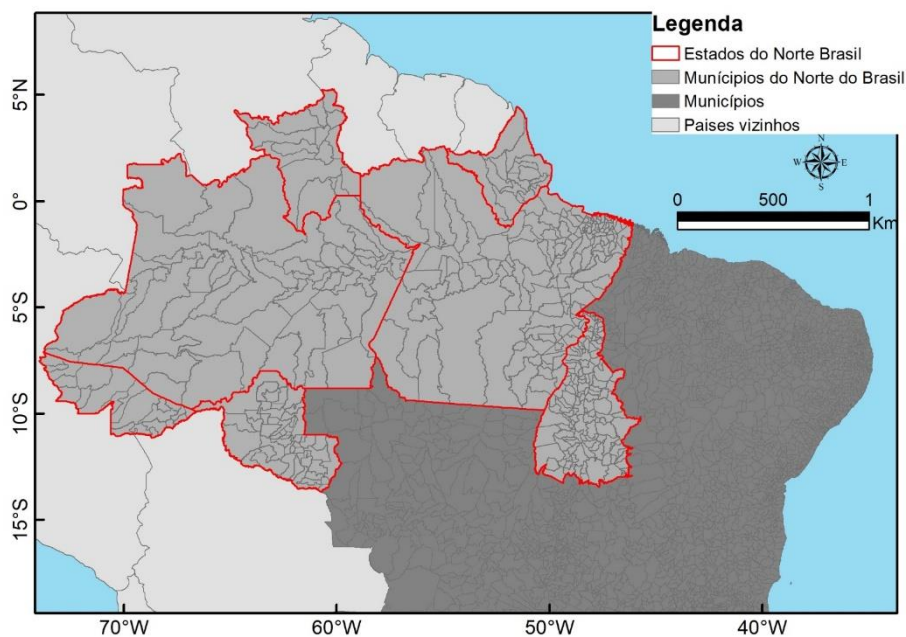


Figura 1. Localização da área de estudo

3.2 Origem e pré-processamento dos dados

A variável de estudo foi as áreas municipais semeada com soja (AMSS) (expressas em hectares) para o intervalo de anos de 1990 a 2016. Os dados de foram obtidos da Pesquisa Agrícola Municipal – PAM, do Instituto Brasileiro de Geografia e Estatística – IBGE. Utilizou-se o Sistema IBGE de Recuperação Automática – SIDRA para download dos dados (IBGE, 2017).

Os dados foram tabulados em uma planilha eletrônica e posteriormente integrados com a representação espacial em formato *shapefile*. Assim, tem-se a representação geográfica dos dados e se torna possível as análises espaciais de áreas.

3.3 Estatística descritiva

Para cada ano em estudo, procedeu-se uma análise estatística descritiva, obtendo-se valores mínimos, máximos, mediana, média, desvio padrão, coeficiente de variação, quartis, variância, etc. Todas essas análises foram realizadas no *software* R.

A normalidade dos dados foi aferida por meio do teste Shapiro Wilk, a 5% de significância. Testou-se as seguintes hipóteses:

- H_0 : A amostra provém de uma população Normal;
- H_1 : A amostra não provem de uma população Normal.

Caso o p-valor seja menor que o nível de 5%, rejeita-se a hipótese nula de normalidade dos dados. Este teste foi realizado para os 27 anos em estudo, por meio do *software* R.

Também se gerou, no *software* R, gráficos de histogramas e boxplot para cada ano a fim de representar a distribuição dos dados.

3.4 Taxa e aceleração de crescimento da AMSS

A expansão e/ou retração dos valores de AMSS foi representado por meio da Taxa de Crescimento relativo (TC) e Crescimento Acelerado (CA). TC é um índice de diferença normalizada que trabalha como um importante operador de séries temporais, representando a porcentagem de mudança entre em um período de tempo (Clemente et al., 2017), sendo expressa pela seguinte equação:

$$TC = \frac{SO_f - SO_i}{SO_i} \quad \text{Equação 1}$$

onde, SO_i e SO_f são os valores de área semeada para cada ano, iniciais e finais do estudo. Foram realizadas as diferenças para cada ano, ou seja, comparou-se o ano 1991 com 1990 (TC19911990), o ano 1992 com 1991 (TC19921991) e assim sucessivamente até o ano 2016 com 2015 (TC20162015).

Os valores de CA representam a variação entre TCs subsequentes. A CA é expressa pela seguinte equação:

$$CA = TC_f - TC_i \quad \text{Equação 2}$$

onde TC_f e TC_i representam as TC iniciais e finais comparadas. Foram realizados cálculos das CA para cada intervalo das TC, logo, comparou-se a TC19921991 com a TC19911990 gerando a CA19921990, a TC19931992 com a TC19921991 gerando a CA19931991, e assim subsequentemente até a TC20162015 com a TC20152014 gerando a CA20162014.

Todo o processamento destes dados foi realizado no *software* R. Os mapas finais de espacialização das variáveis foram confeccionados no *software* ArcGis 10

3.5 Critérios de vizinhança

Matriz de proximidade espacial é uma ferramenta útil para descrever o arranjo espacial dos municípios (áreas). Dado um conjunto de n áreas ($A_1, A_2, A_3, \dots, A_n$) constrói-se a matriz de pesos $W_{n \times n}$ onde cada um dos elementos w_{ij} representa uma medida de proximidade espacial entre A_i e A_j . A célula w_{ij} é nula no caso das regiões não serem vizinhas ($w_{ij} = 0$; para $i \neq j = 1, 2, \dots, n$), e caso a região A_i seja vizinha da região A_j , o valor passa a ser 1 ($w_{ij} = 1$ para $i \neq j = 1, 2, \dots, n$).

Ao utilizar matrizes de proximidade espacial que considerem a contiguidade, é importante convencionar as formas de vizinhança. Para tanto, os critérios são baseados em movimentos de algumas peças do jogo de xadrez (Figura 2), como rainha, torre e bispo.

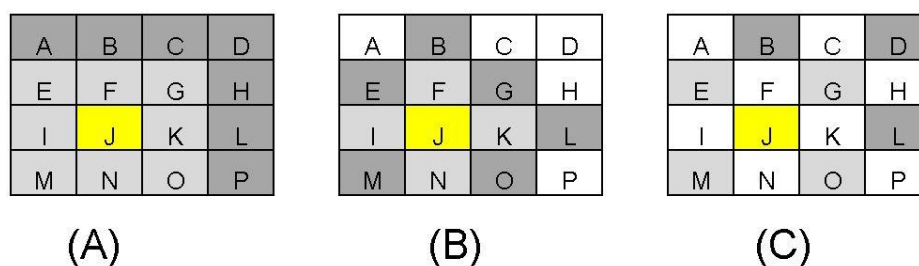


Figura 2. Critérios de Vizinhança para (A) Rainha, (B) Torre e (C) Bispo.

Para o critério Rainha (Figura 2A), as células E, F, G, I, K, O, N e M são vizinhas de primeira ordem em relação à J . Ao considerar o critério Torre (Figura 2B), as células I, N, K e F são vizinhas de primeira ordem em relação a J , por fim, ao se considerar o

critério Bispo (Figura 2C), tem-se que os vizinhos de primeira ordem da célula J são as células E , G , O e M .

O vetor dos desvios Z (Equação 3), é um vetor de dimensão n , cujos valores dos elementos são os desvios dos atributos associados aos objetos em relação à média dos n atributos (μ) (Perobelli et al., 2007). Para o cálculo do vetor de desvios, Z , é calculada, primeiramente, a média (μ) dos valores do atributo em estudo ($y_1 \dots y_n$), considerando os n objetos. Cada elemento i de Z , z_i , é obtido subtraindo-se o valor da média, do valor do atributo correspondente ($z_i = y_i - \mu$).

$$Z = (z_i - \mu)1^T; \quad \text{Equação 3}$$

em que: z_i representa o elemento i do vetor de desvios Z ; μ corresponde à média dos valores e; T representa que o vetor se apresenta na sua forma transposta.

O vetor das médias ponderadas (W_z) resulta da multiplicação do vetor transposto do vetor dos desvios Z^T pela matriz de proximidade normalizada W' (Marconato et al., 2012). Obtendo-se desta multiplicação uma matriz W_z (Equação 4), onde cada elemento W_z contém um valor correspondente à média dos desvios dos vizinhos da área i , caracterizando uma média móvel espacial.

$$W_z = Z^T \cdot W' \quad \text{Equação 4}$$

em que: Z^T representa o vetor de desvios na sua forma transposta e W' corresponde a matriz de proximidade espacial na sua forma normalizada.

Neste estudo utilizou-se o critério rainha de vizinhança, considerando apenas os vizinhos de primeira ordem, para estudar os padrões de autocorrelação espacial de dos valores de AMSS. Esta matriz foi construída nos *softwares* Geoda e R.

3.6 Estatística espacial de área

O coeficiente I de Moran tem um valor esperado $E(I)$, o qual é definido pela equação $E(I) = -[1/(n-1)] = -0,0022$, onde n indica o número de polígonos (no caso, 450 municípios) da área em estudo. Deve-se notar que, ao contrário de um coeficiente de correlação ordinário, essa estatística não é centrada em zero e, portanto, tal estatística varia entre -1 e $+1$ (Betarelli Junior & Almeida, 2009; Montenegro & Betarelli Junior, 2008). Valores do I de Moran maiores ou menores do que o valor esperado $E(I)$ revelam autocorrelação espacial positiva ou negativa, respectivamente.

Para testar a significância do I de Moran, gera-se uma distribuição de referência usando valores aleatórios de AMSS nas áreas e calcula-se um novo valor de AMSS para cada permutação feita, permitindo a realização do seguinte teste de hipóteses:

- $H_0: I = 0$, que representa a hipótese nula de ausência de correlação espacial versus a hipótese alternativa
- $H_1: I \neq 0$.

Se o nível descritivo (p-valor) for menor que o nível de significância estipulado (usualmente 0,05 de probabilidade) rejeita-se a hipótese nula de ausência de correlação espacial. Caso contrário, se o nível descritivo for maior ou igual ao nível de significância, não se rejeita a hipótese nula

Considerando todos os critérios anteriormente mencionados, o I de Moran pode ser apresentado conforme a Equação 5:

$$I = \frac{Z^T \cdot W_z}{Z^T \cdot Z} \quad \text{Equação 5}$$

em que: Z corresponde ao vetor de desvios; Z^T representa o vetor de desvios na sua forma transposta e; W_z corresponde ao vetor de médias ponderadas.

O I de Moran retorna apenas um valor numérico, que pode induzir facilmente a erros de interpretação. Uma alternativa para auxiliar a interpretar o valor do índice em questão é construir os diagramas de dispersão de Moran (*Moran Scatter Plots*). Estes revelam a defasagem espacial da variável de interesse (ou seja, a média do atributo nos vizinhos) no eixo vertical (W_z) e o valor da variável de interesse no eixo horizontal (Z). Neste gráfico o valor I de Moran, representa o coeficiente angular da reta de regressão, ou seja, esse valor indica a inclinação da reta de W_z para Z (Marconato et al., 2012; Perobelli et al., 2007; Predebon et al., 2010).

Além da medida global de associação linear espacial, esse diagrama está dividido em quatro quadrantes, conforme a Figura 3, sendo eles: Alto-Alto (AA), Baixo-Baixo (BB), Alto-Baixo (AB) e Baixo-Alto (BA). Os quatro quadrantes representam o relacionamento entre os valores do vetor Z e os valores das médias ponderadas W_z .

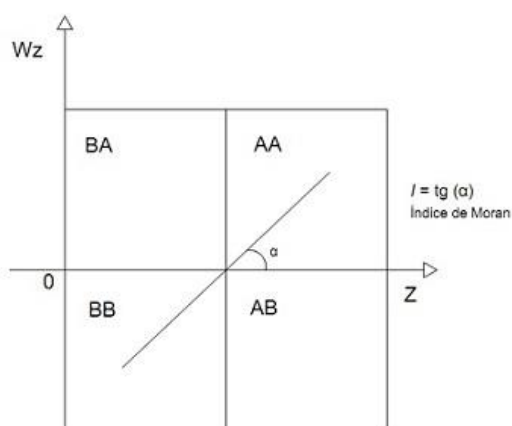


Figura 3. Gráfico de Espalhamento de Moran.

O índice $LISA$ (Equação 6) visa atender a dois objetivos: permitir a identificação de padrões de associação espacial significativos, e também ser uma decomposição do índice global de associação espacial (Anselin, 1995). O índice $LISA$ é representado pela,

$$I_i = (Z_i \cdot W_{z_i}) / \sigma^2 \quad \text{Equação 6}$$

em que: σ^2 representa a variância da distribuição dos valores dos desvios (Marconato et al., 2012)

Uma vez determinada a significância estatística do índice local de Moran, é útil gerar um mapa indicando as regiões que apresentam correlação local significativamente diferente do resto dos dados.

Totalizou-se 27 análises de AMSS (uma para cada ano) com o I de Moran, gráficos de espalhamento e $LISA$. Os critérios de vizinhança foram adotados de acordo com o item anterior e foram utilizados os softwares Geoda e R, sendo os resultados posteriormente exportados em *shapefiles* e planilhas eletrônicas para tabulações e espacializações nos softwares Excel e ArcGis 10.

3.7 Estatística espaço-temporal de áreas

Para identificar os padrões de variação do espaço em diferentes tempos, analisou-se os índices I de Moran das diferenças de área semeada entre os anos 1995-1990, 2000-1995, 2005-2000, 2010-2005, 2015-2010. Para estes mesmos anos foram

analisados os gráficos de espalhamento e os índices *LISA* das diferenças. Esta técnica de análise espacial das diferenças está implementada dentro do *software* Geoda.

Os padrões espaço-temporais dos dados de AMSS foram analisados por meio do emprego da técnica *SatScan* (dentro do *software SatScan* (Block, 2007; Kulldorff, 2015), considerando modelo de probabilidade Normal (Kulldorff, Huang, & Konty, 2009), em uma análise retrospectiva espaço-temporal. Tempo de agregação foi de um ano, os limites de formação de clusters, tanto do ponto de vista espacial, quanto temporal foram de 50% do total dos valores.

A utilização do modelo de probabilidade se deu devido ao interesse de detectar clusters espaciais de municípios com alto ou baixos valores da variável contínua de AMSS. Este modelo pode ser aplicado a dados não normais (Kulldorff et al., 2009).

A estatística escaneamento circular espacial é definida por meio de um grande número de círculos sobrepostos. Para cada círculo z , uma razão de log de probabilidade é calculada (*log likelihood ratio* – $LLR(z)$), e o teste estatístico é definido para o máximo LLR de todos os círculos. A janela de escaneamento irá depender da aplicação, mas é tipicamente definida a janela com todos os círculos centrados em uma observação e com o raio variando continuamente do zero até um limite superior (geralmente 50% de todas as observações) (Kulldorff et al., 2009).

Para a técnica acima os dados foram necessários pequenos ajustes no banco de dados. Foram transpostos de colunas (uma para cada ano) para linhas, criando-se assim duas colunas, uma para os anos e outra para os valores de áreas semeadas de soja.

4 Resultados e Discussões

A média de AMSS para os municípios do Norte é baixa, entretanto destaca-se uma evolução considerável dos valores de AMSS principalmente no período de 2000 a 2016 (Tabela 1 e Figura 6). Muitos municípios não apresentaram valores de AMSS (Tabela 1, Figura 4 e Figura 5), o que culminou refletindo em medianas iguais a zero para todos os anos sob estudo (Tabela 1).

Os gráficos de boxplot e histogramas (Figura 4 e Figura 5) demonstram a concentração de valores próximos e igual a zero. Este fato faz com que praticamente qualquer valor acima de zero seja apontado com um ponto de *outlier* (Figura 4). Além disto, essa concentração nos valores próximos a zero não caracteriza uma distribuição normal dos dados (Figura 5), fato este detectado pelo teste de Shapiro Wilk (Tabela 1) a 5% de significância.

Entre os anos de 1990 a 1995 haviam dados de AMSS para alguns municípios dos estados de Rondônia e Tocantins, sendo que no restante dos estados os valores de AMSS foram iguais a zero (Figura 6).

Dos anos 1996 a 2002 alguns municípios no estado do Amazonas apresentaram valores de até 2000 hectares de AMSS (Figura 6). Estes possivelmente sejam referentes a uma (ou umas) fazenda(s) que foram implementadas nessas regiões neste período de tempo. A partir dos anos 1998, intensificados após os anos 2005, surgem municípios com valores maiores de 20000 hectares de AMSS (Figura 6). Os estados de Tocantins e Rondônia são os que mais possuem crescimento dos valores de AMSS (Figura 6).

Tabela 1. Estatística descritiva para os valores de AMSS nos municípios do Norte do Brasil entre os anos 1990-2016

Anos	Mínimo	1º Quartil	Mediana	Média	3º Quartil	Máximo	Variância	Shapiro Wilk		DP	CV
								Estatística W	p-valor		
1990	0,00	0,00	0,00	77,24	0,00	9500,00	335736,27	0,12	3,18E-27	579,43	750,12
1991	0,00	0,00	0,00	12,09	0,00	1000,00	7306,54	0,13	4,63E-27	85,48	707,08
1992	0,00	0,00	0,00	24,84	0,00	4000,00	51293,04	0,09	1,10E-28	226,48	911,59
1993	0,00	0,00	0,00	46,86	0,00	6930,00	192378,39	0,08	9,93E-28	438,61	936,09
1994	0,00	0,00	0,00	79,44	0,00	9400,00	325013,39	0,12	4,13E-27	570,10	717,61
1995	0,00	0,00	0,00	54,97	0,00	4500,00	149374,05	0,13	4,58E-27	386,49	703,08
1996	0,00	0,00	0,00	16,99	0,00	1440,00	15314,56	0,12	3,75E-27	123,75	728,52
1997	0,00	0,00	0,00	79,43	0,00	16400,00	841605,74	0,06	4,27E-28	917,39	1155,02
1998	0,00	0,00	0,00	150,62	0,00	20000,00	1295822,41	0,11	2,76E-27	1138,34	755,77
1999	0,00	0,00	0,00	127,01	0,00	20360,00	1218236,21	0,09	1,28E-27	1103,74	869,04
2000	0,00	0,00	0,00	162,23	0,00	24870,00	2043683,90	0,09	1,26E-27	1429,57	881,20
2001	0,00	0,00	0,00	235,86	0,00	28800,00	3463748,02	0,11	2,15E-27	1861,11	789,08
2002	0,00	0,00	0,00	312,57	0,00	30300,00	4933352,26	0,12	3,98E-27	2221,12	710,60
2003	0,00	0,00	0,00	471,70	0,00	33626,00	7215570,96	0,17	1,95E-26	2686,18	569,47
2004	0,00	0,00	0,00	798,74	0,00	37000,00	11784648,93	0,24	4,36E-25	3432,88	429,79
2005	0,00	0,00	0,00	1142,88	92,50	45000,00	19712946,93	0,27	1,53E-24	4439,93	388,49
2006	0,00	0,00	0,00	1150,98	0,00	48000,00	20141900,06	0,27	1,32E-24	4487,97	389,92
2007	0,00	0,00	0,00	1011,28	0,00	49000,00	16147008,57	0,27	1,09E-24	4018,33	397,35
2008	0,00	0,00	0,00	1128,94	0,00	45000,00	17898339,78	0,29	3,33E-24	4230,64	374,74
2009	0,00	0,00	0,00	1111,22	0,00	45000,00	18380233,47	0,28	2,44E-24	4287,22	385,81
2010	0,00	0,00	0,00	1250,55	0,00	48000,00	21696985,36	0,30	4,48E-24	4658,00	372,48
2011	0,00	0,00	0,00	1418,48	0,00	53000,00	28605847,39	0,29	3,71E-24	5348,44	377,06
2012	0,00	0,00	0,00	1525,58	0,00	53500,00	29080616,73	0,32	1,09E-23	5392,64	353,48
2013	0,00	0,00	0,00	2057,13	230,00	67800,00	46842945,98	0,34	2,98E-23	6844,19	332,71
2014	0,00	0,00	0,00	2647,39	500,00	73000,00	65830384,51	0,37	1,51E-22	8113,59	306,47
2015	0,00	0,00	0,00	3190,84	692,50	95000,00	94269572,92	0,37	1,37E-22	9709,25	304,29
2016	0,00	0,00	0,00	3478,46	1000,00	102163,00	108030987,81	0,38	2,02E-22	10393,80	298,80

DP = Desvio Padrão; CV = Coeficiente de Variação; *a 5% de significância

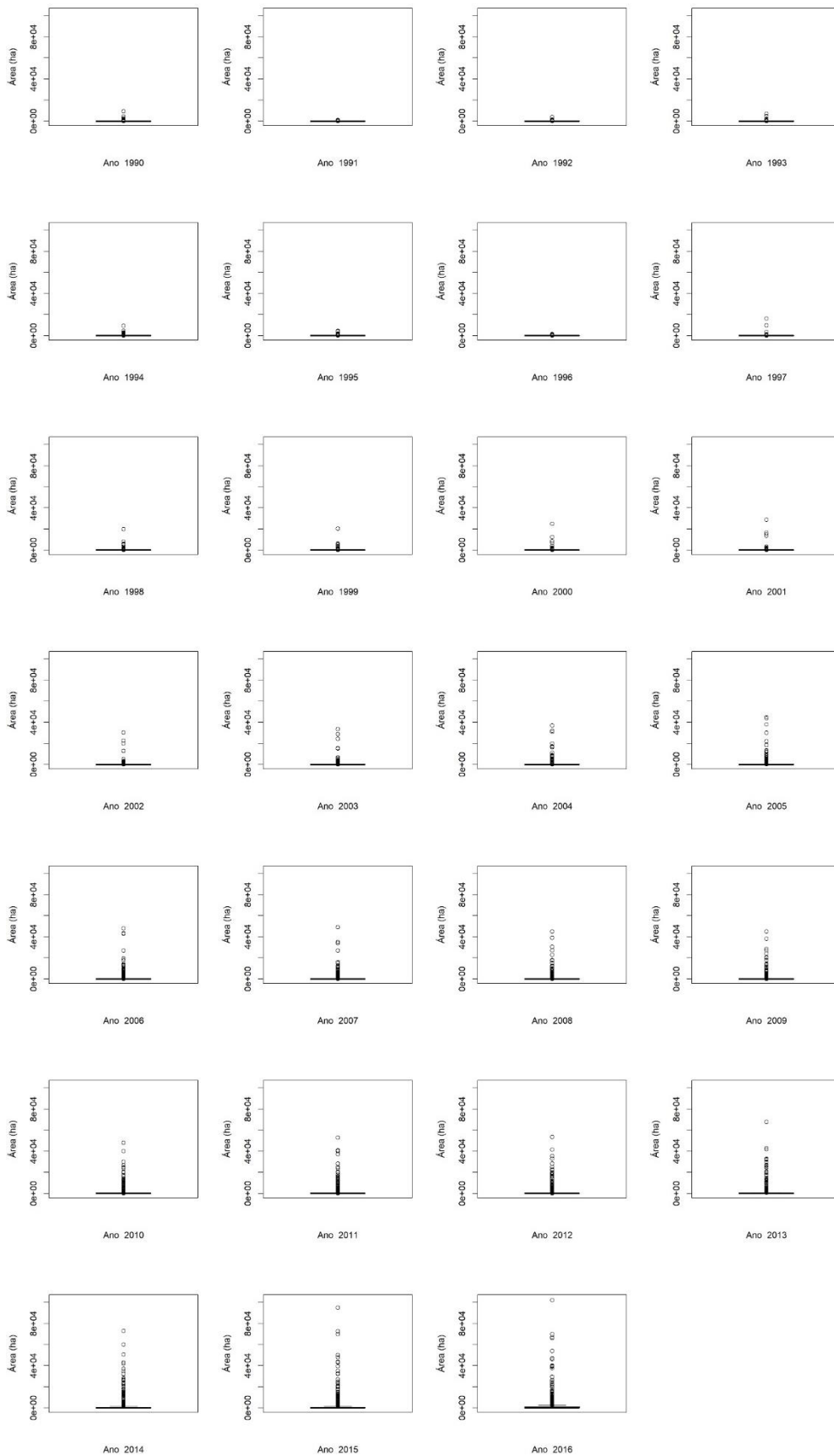


Figura 4. Boxplot para AMSS para a região norte do Brasil entre os anos de 1990 a 2016

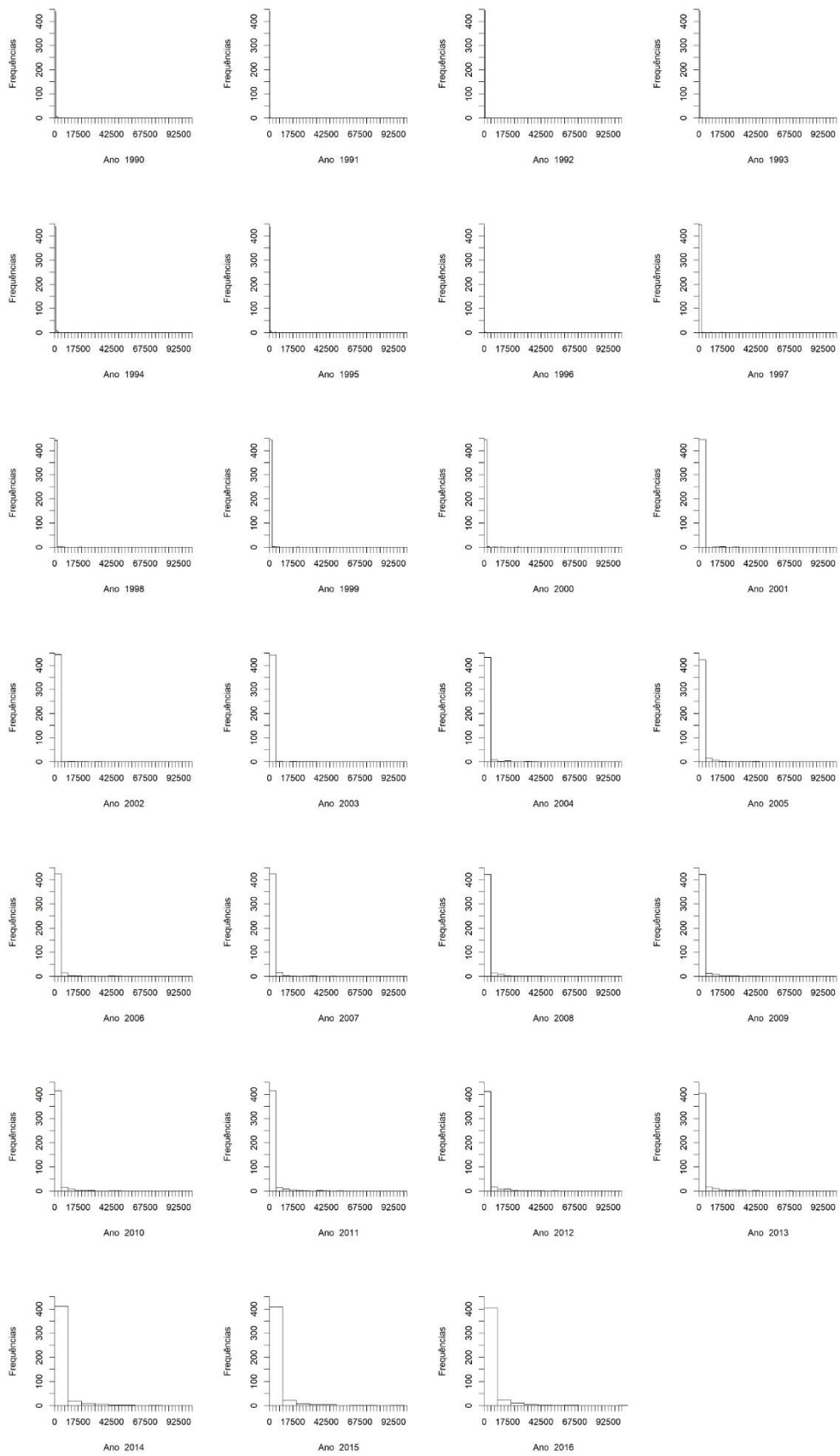


Figura 5. Histogramas para AMSS para a região norte do Brasil entre os anos de 1990 a 2016

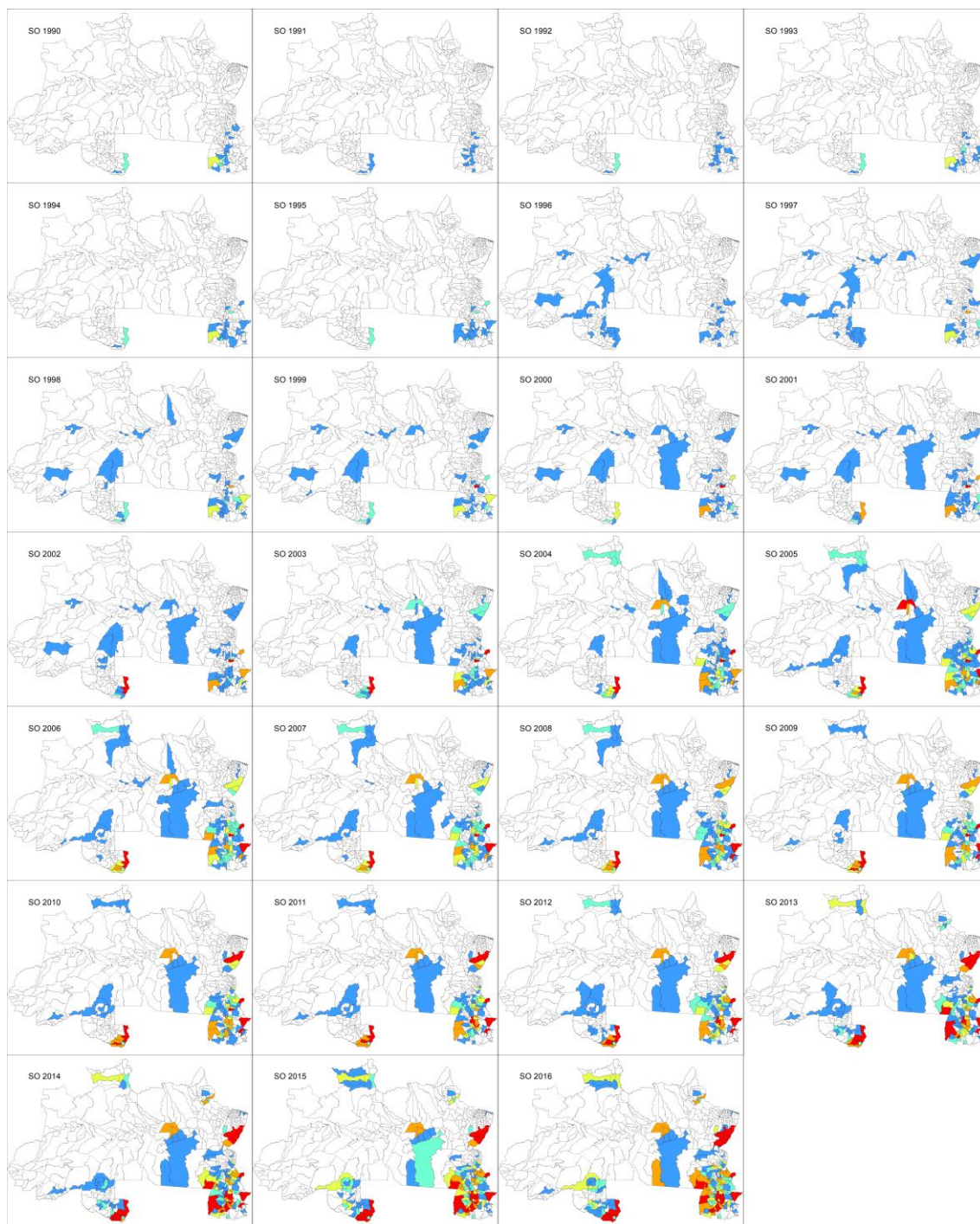


Figura 6. AMSS para a região norte do Brasil entre os anos de 1990 a 2016

O surgimento de novas áreas agrícolas culminou em TC de crescimentos exorbitantes em alguns municípios (Figura 7). Até os anos 2000, os maiores crescimentos haviam ocorrido em municípios do Tocantins (Figueirópolis TC19921991 = 16,86 vezes, Porto Nacional TC19931992 = 13,17 vezes), de Roraima (Vilhena TC19981997 = 9,73 vezes, Colorado TC19991998 = 49 vezes) e em alguns poucos casos do Pará (Redenção e Paragominas TC19981997 = 9 e 1,83 respectivamente) e do Amazonas (Itacoatiara TC19981997 = 0,55 vezes). O maior número de municípios que reduziram as TC de AMSS ocorreram entre os TC20062005 a TC20092008, onde o máximo de 76 municípios (TC20072006) reduziram, de alguma maneira, os valores de AMSS.

Até os anos 1996 houve CA apenas nos municípios de Rondônia e Tocantins (Figura 8). Após os anos 2000 os valores de CA começaram a se intensificar em municípios dos demais estados. Após o ano de 2010 houve CA maiores para alguns municípios de Roraima, Amapá e Acre, além dos demais estados.

Uma maior flutuação entre os valores de CAs (Figura 8) ao longo dos anos ocorre devido ao baixo valor de AMSS ao longo dos anos (Figura 6). Algumas vezes os valores de AMSS correspondem a extensão de poucas áreas, fazendo que qualquer área que tenha uso diferente de soja em determinado ano, culmine em uma um valor de CA extremamente baixo, ou qualquer nova área destinada a soja (de extensão não tão exorbitante) pode culminar em valores de CA extremamente altos.

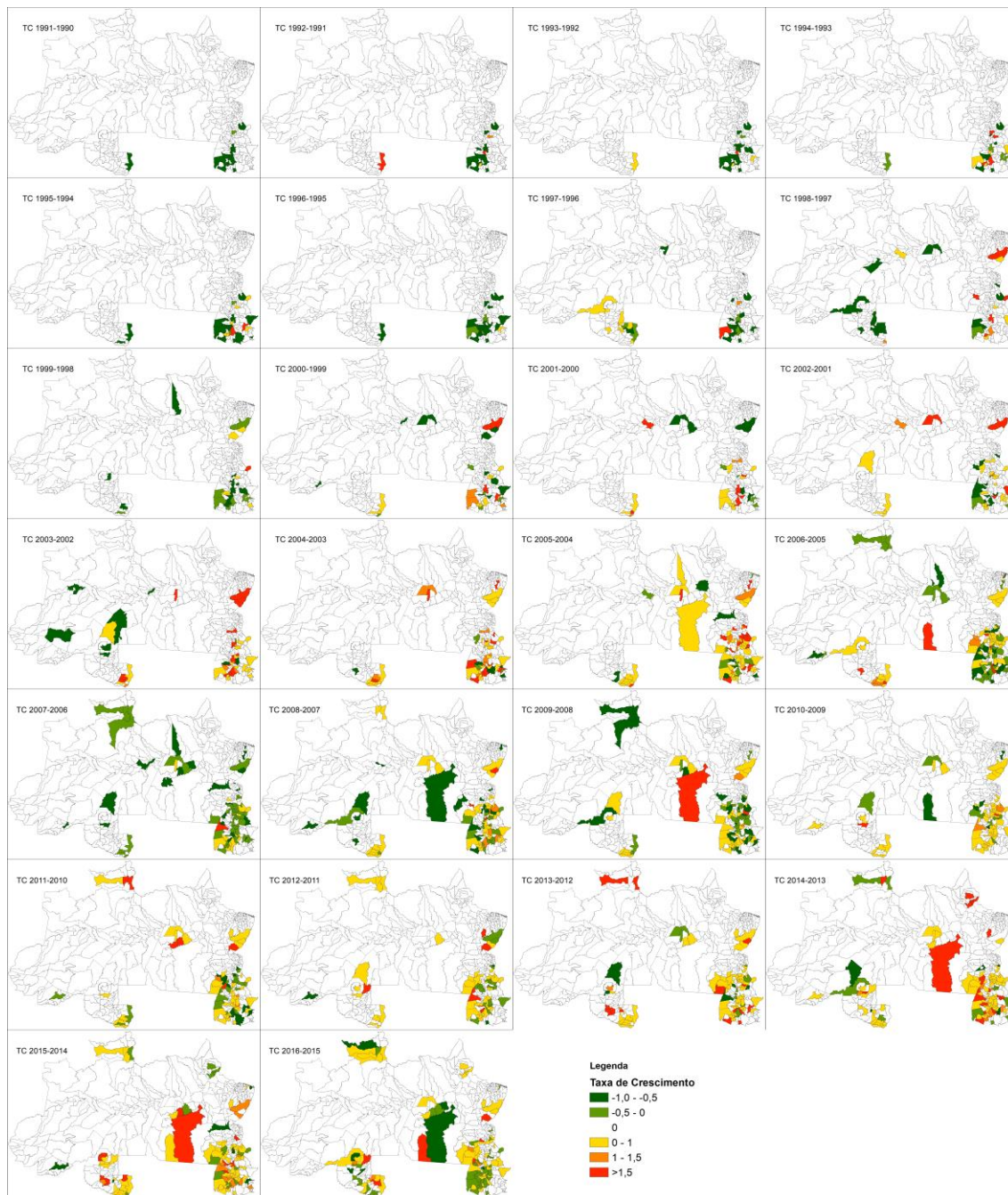


Figura 7. Taxa de Crescimento para os valores de AMSS nos municípios do Norte do Brasil entre os anos 1990-2016

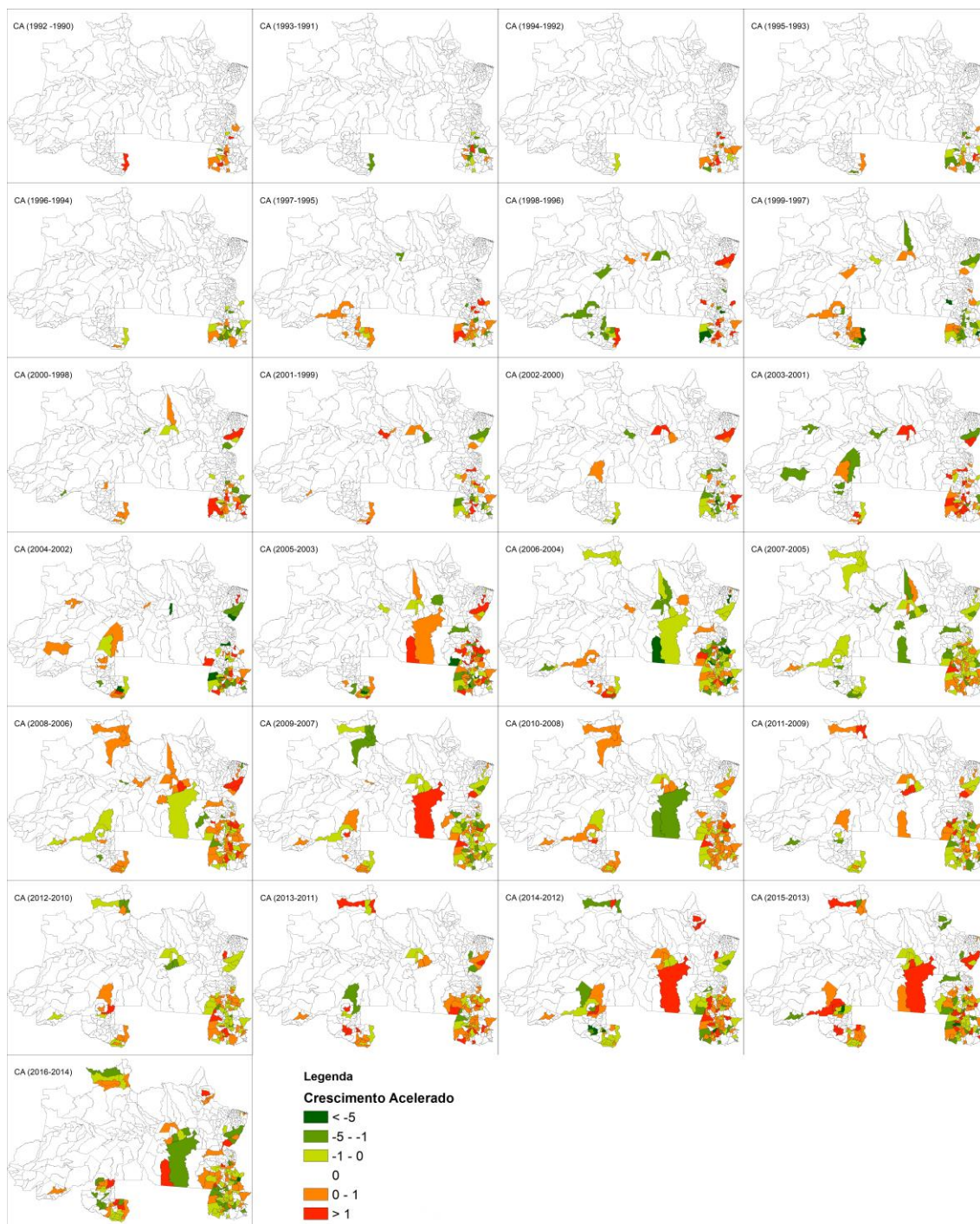


Figura 8. Crescimento Acelerado para os valores de AMSS nos municípios do Norte do Brasil entre os anos 1990-2016

A partir dos anos 2000 o valor do I de Moran (Tabela 2 e Figura 9) apresentou constantes crescimentos, representando uma maior tendência de autocorrelação espacial para os valores de AMSS.

Entre os anos de 1991 a 2003 os valores do I de Moran foram menores, demonstrando uma menor autocorrelação global para os valores de AMSS. Este fato pode ser observado nos gráficos de espalhamento de Moran (Figura 10), onde entre os anos de 1991 a 2003 os valores de Z e Wz estão sobre os eixos, destacando a baixa

autocorrelação espacial. Não houve valores do I de Moran negativos, sendo fato notado pela maior concentração de pontos/municípios ocorrerem no quadrante AA dos diagramas de espalhamento de Moran.

O maior valor do I de Moran foi justamente para o último ano de estudo, caracterizando uma possível consolidação da soja nas áreas de agricultura para os municípios sob estudo.

Tabela 2. Índice I de Moran para os valores de AMSS nos municípios do Norte do Brasil entre os anos 1990-2016

Anos	I Moran	Anos	I Moran
1990	0,15	2004	0,12
1991	0,07	2005	0,14
1992	0,03	2006	0,15
1993	0,01	2007	0,16
1994	0,08	2008	0,20
1995	0,08	2009	0,18
1996	0,07	2010	0,22
1997	0,01	2011	0,22
1998	0,02	2012	0,26
1999	0,00	2013	0,29
2000	0,01	2014	0,31
2001	0,03	2015	0,37
2002	0,03	2016	0,40
2003	0,08		

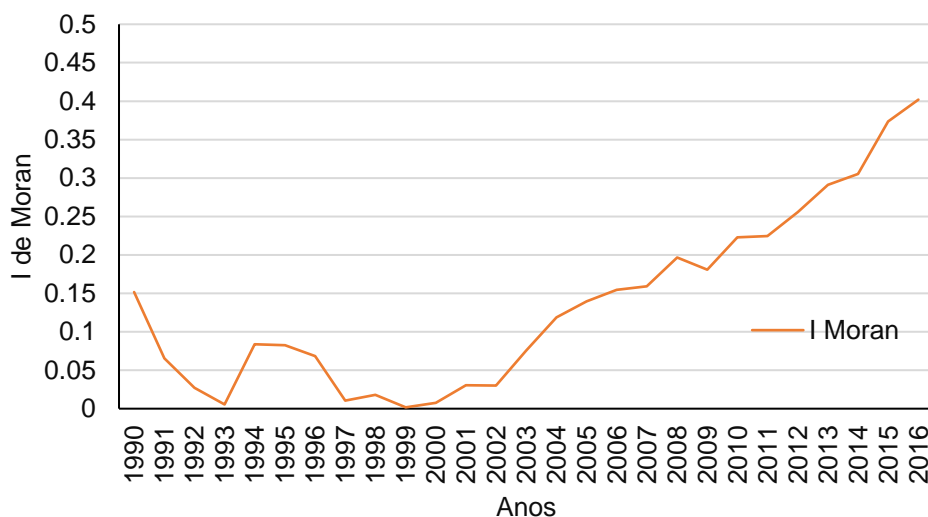


Figura 9. Valores do I de Moran para os dados de AMSS nos municípios do Norte do Brasil entre os anos 1990-2016

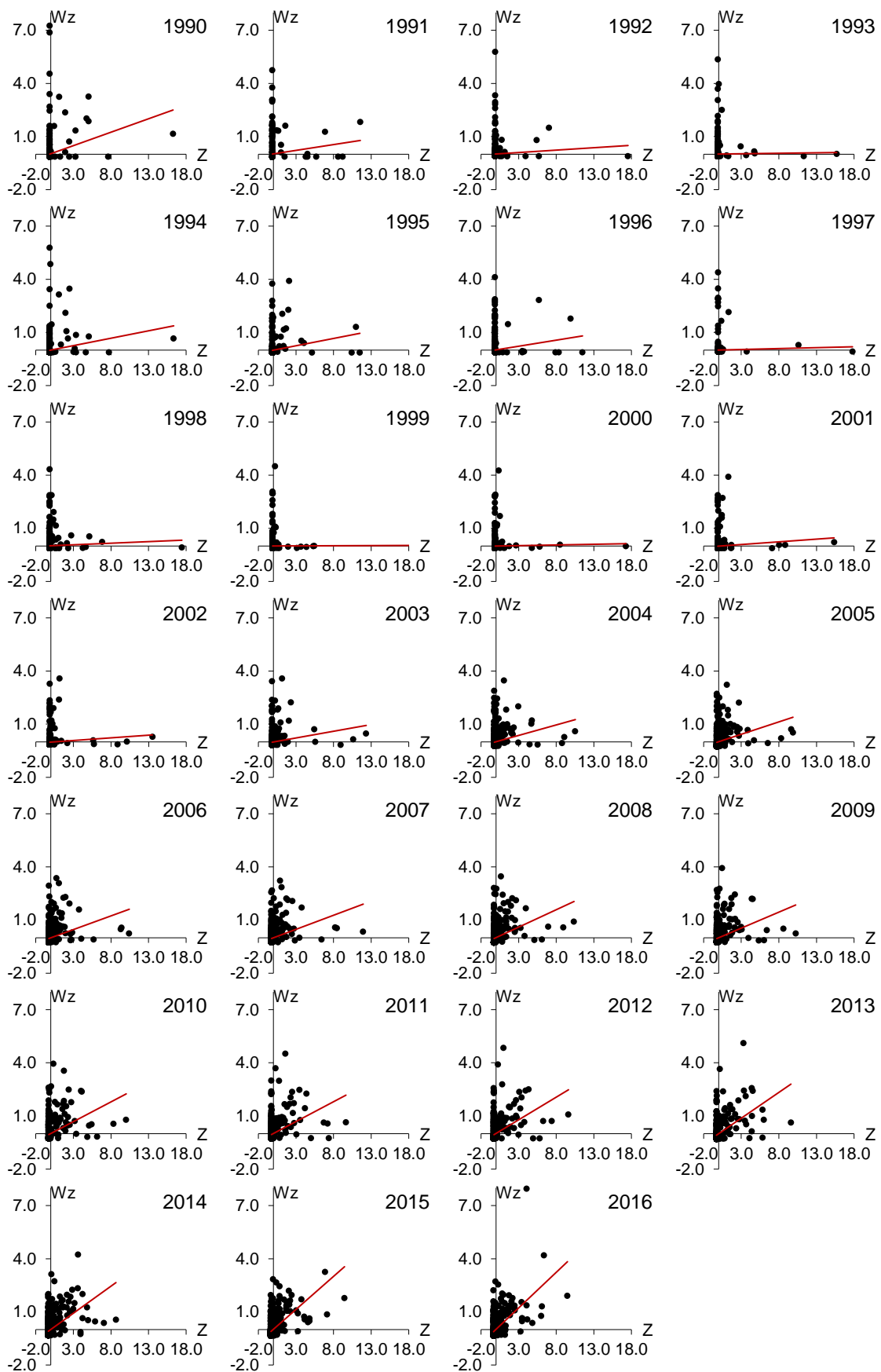


Figura 10. Gráficos de espalhamento de Moran para os dados de AMSS nos municípios do Norte do Brasil entre os anos 1990-2016

Inicialmente houve poucos *clusters* formados, destacando apenas os municípios da região sudoeste do estado do Tocantins, que formaram um *cluster do tipo Alto-Alto* no ano de 1990 (Figura 11).

Em todos os anos de estudo houve a presença de um *cluster* de altos valores de AMSS próximo ao município de Peixe, na região centro sul do estado de Tocantins (Figura 11). Após os anos 2000, iniciou-se a formação de outro *cluster* do tipo Alto-Alto para os municípios próximos a Corumbiara, ao leste do estado de Rondônia. Fato semelhante aconteceu para o município de Ulianópolis, a centro leste do estado do Pará.

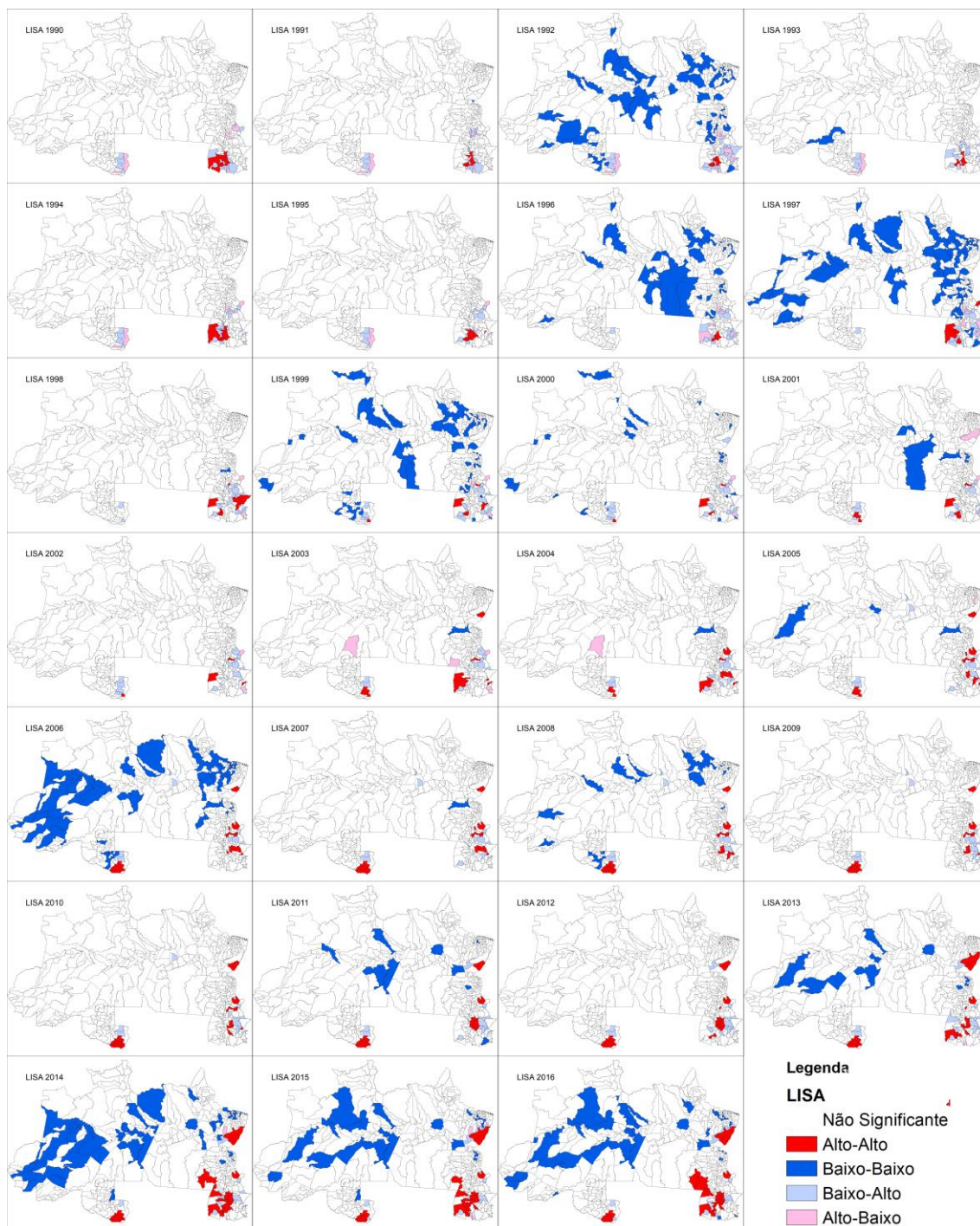


Figura 11. LISA para os dados de AMSS nos municípios do Norte do Brasil entre os anos 1990-2016

Para o último ano sob estudo houve a formação de 4 *clusters* de altos valores de AMSS. Sendo que além dos três anteriormente citados, houve a formação de um cluster envolvendo municípios ao sudeste do estado do Pará (Santana do Araguaia) e ao noroeste do estado do Tocantins (Pium). O município de Goiatins (ao nordeste do estado do Tocantins) se caracteriza como um pequeno *cluster* do tipo Alto-Alto desde o ano 2004 até o ano de 2016.

Considerando as análises da diferença, tem-se os maiores ganhos foram comparando os anos 2005 com 2000 e os anos 2015 com 2010 (Figura 12). Nestes anos destacam-se os crescimentos de AMSS para muitos dos municípios do estado do Tocantins, e em alguns municípios dos estados do Pará e Roraima. Os municípios do estado de Rondônia possuem, de maneira geral, maior aumento de AMSS comparando o ano de 2015 com 2010.

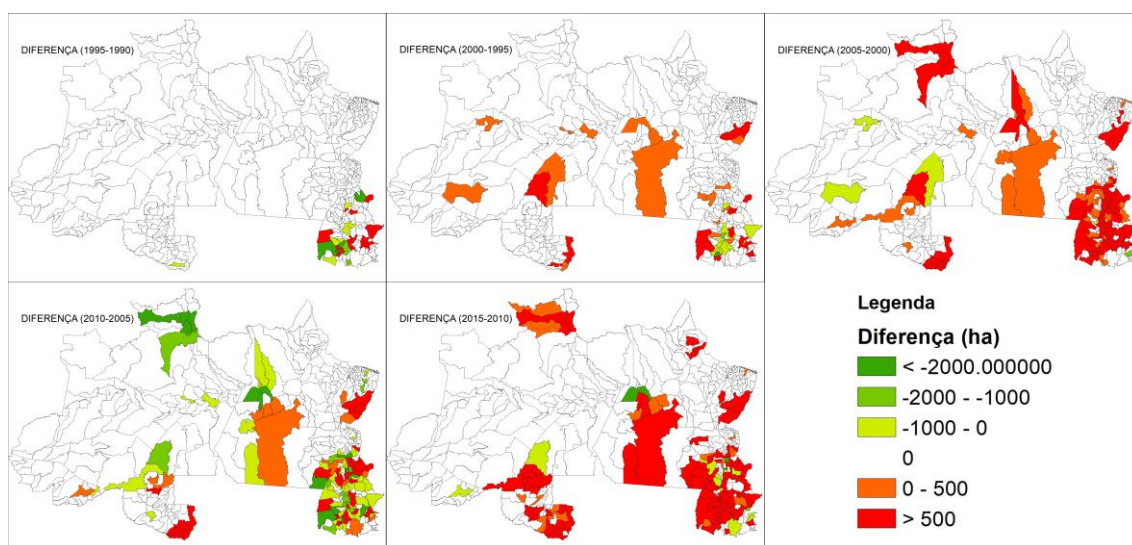


Figura 12. Diferenças de AMSS nos municípios do Norte do Brasil entre os anos 1990-2016

Há uma tendência de aumento dos valores do *I* de Moran a medida que se estudam as diferenças de AMSS para anos mais recentes (Tabela 3 e Figura 13). Este fato pode ser explicado principalmente pelo maior agrupamento dos dados no quadrante Alto-Alto (Figura 14), indicando que as diferenças de AMSS se concentram em municípios próximos.

Tabela 3. Índice *I* de Moran para os valores de diferenças de AMSS nos municípios do Norte do Brasil entre os anos 1990-2016

Intervalo em anos	<i>I</i> Moran
1995 - 1990	0,043473
2000 - 1995	0,010929
2005 - 2000	0,156331
2010 - 2005	0,268364
2015 - 2010	0,380714

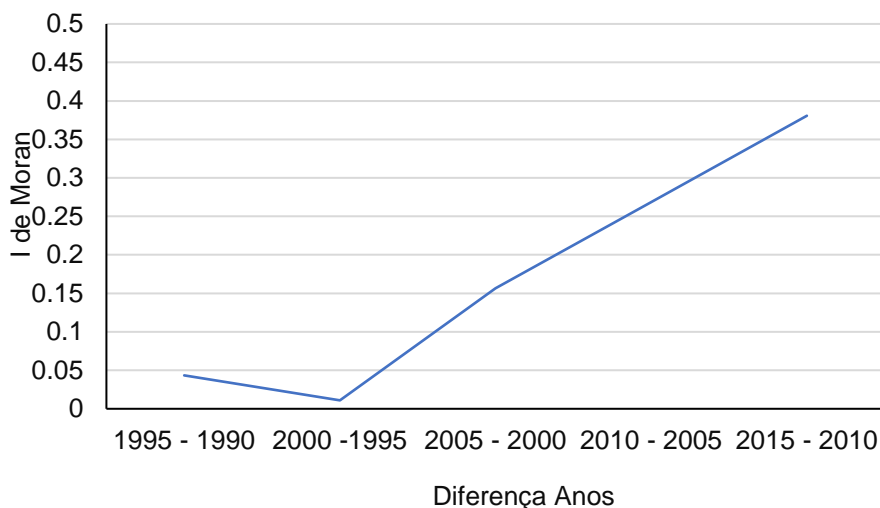


Figura 13. Evolução do índice I de Moran para os valores de diferenças de AMSS nos municípios do Norte do Brasil entre os anos 1990-2016

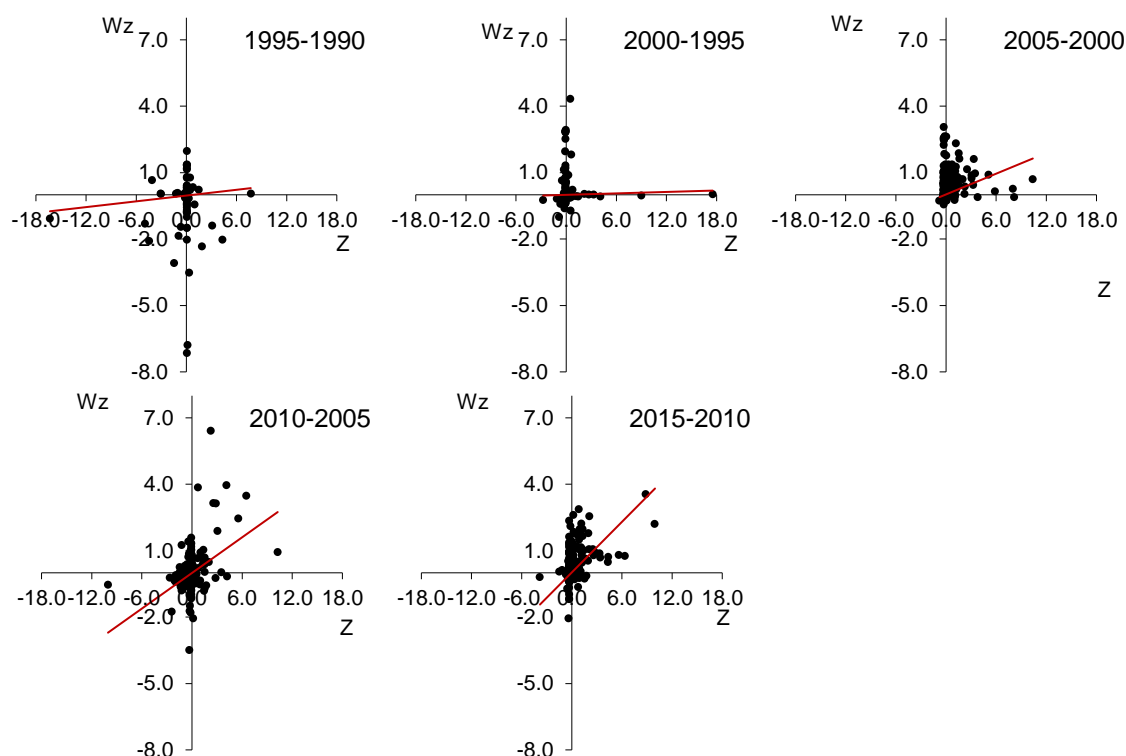


Figura 14. Gráficos de espalhamento de Moran para os valores de diferenças de AMSS nos municípios do Norte do Brasil entre os anos 1990-2016

Nas análises de autocorrelação local das diferenças temporais de AMSS, houve a formação de dois *clusters*, do tipo Alto-Alto, maiores para considerando a diferença de 2010 – 2005, sendo um situado a leste do estado de Rondônia e outro a nordeste do estado do Pará (Figura 15). Estes mesmos *clusters* já haviam sido mencionados nas análises anteriores (Figura 11). Considerando o ano as diferenças entre 2015-2010, manteve o *cluster* do Pará, e surgiu outro que ocupa praticamente toda a porção oeste do estado de Tocantins.

Isto vem evidenciar um maior uso do solo nesses municípios para a soja, porém não há como caracterizar, nesse estudo, se esses aumentos provêm de uma intensificação de uso, ou de uma extensificação de área.

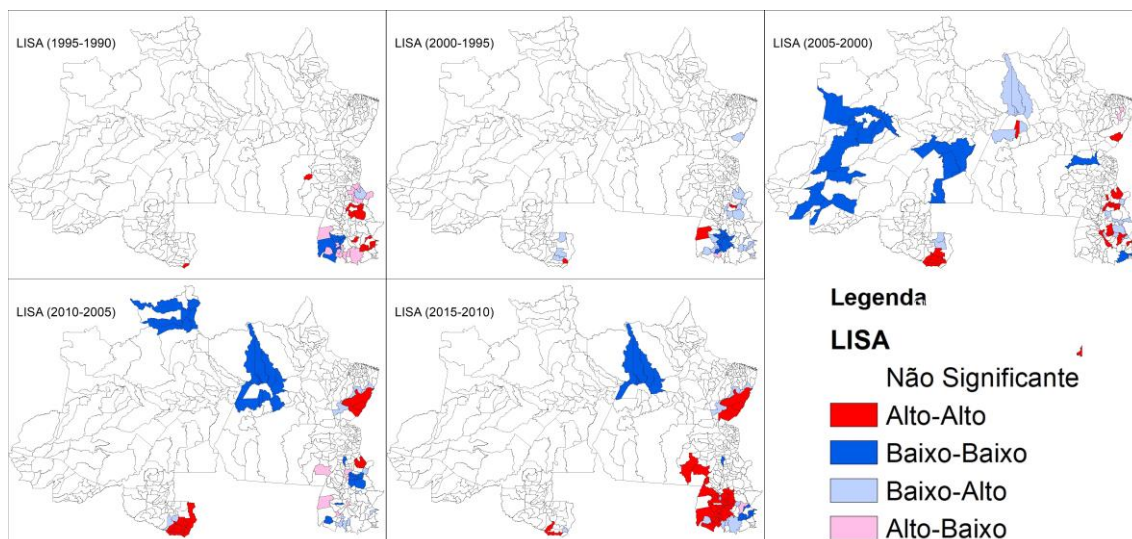


Figura 15. LISA para os valores de diferenças de áreas municipais semeadas com soja nos municípios do Norte do Brasil entre os anos 1990-2016

A aplicação da técnica SatScan evidenciou quatro aglomerados do tipo Alto para os dados de AMSS durante o período sob estudo (Figura 12). Os intervalos de 2013-2016 formaram dois aglomerados, um para os municípios ao oeste do estado do Tocantins e o município de Santana do Araguaia (PA) e outro para os municípios a nordeste do estado do Pará (Dom Eliseu, Paragominas e Ulianópolis).

Os municípios a leste do estado de Rondônia (Vilhena, Cabixi, Corumbiara, etc.) formaram um *cluster* que engloba dos anos de 2005-2016. O *cluster* de maior intervalo temporal (2004-2016) ocorreu para o município de Campos Lindos (TO). Este mesmo município já havia apresentado a formação de clusters para o índice *LISA* nesse mesmo intervalo de tempo.

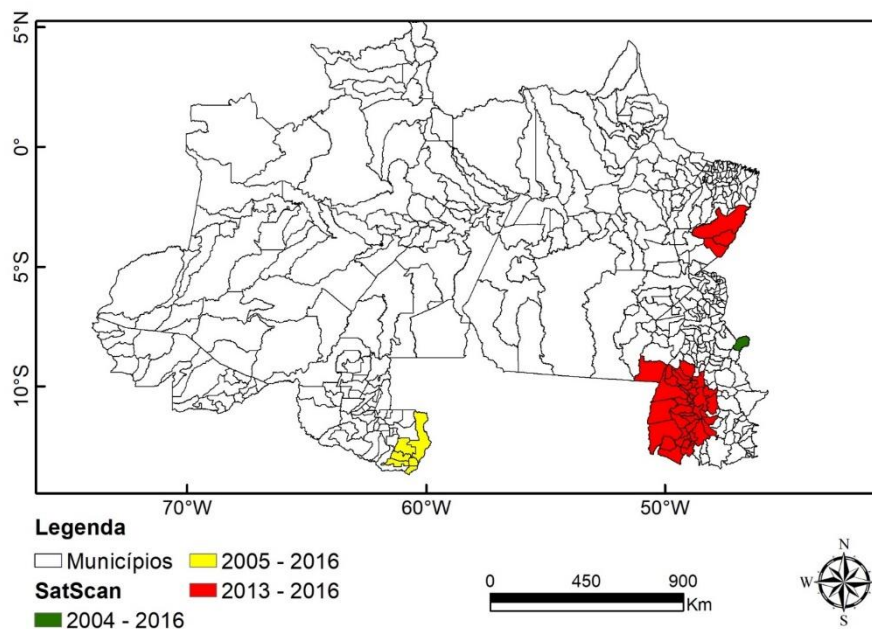


Figura 16. Análise espaço-temporal dos valores de AMSS nos municípios do Norte do Brasil entre os anos 1990-2016

Há maioria dos municípios não apresentaram formação de aglomerados para a técnica SatScan, assim como ocorreu para a análise espacial de cada ano. Este fato já

era esperado, uma vez que a maioria destes municípios se encontram dentro da floresta amazônica e conseqüentemente não se espera um uso agrícola nessas áreas.

A escolha de parâmetros diferentes para a técnica SaTScan (critérios de limites espaciais e temporais) podem culminar no surgimento de novos *clusters*, porém não foi abordado nesse trabalho.

De uma maneira geral, observa-se que houve um aumento dos *clusters* com o passar dos anos, o que identifica uma maior intensificação ou extensificação das AMSS nesta região. As regiões leste de Rondônia e oeste de Tocantins, que se destacam em ambas abordagens (SaTScan e LISA) como um *cluster* de altos valores de AMSS, fazem fronteira com o estado do Mato Grosso, maior produtor de soja do Brasil. A expansão da agricultura para fora das fronteiras do estado do Mato Grosso pode ser um dos motivos que levou a maior exploração dessas áreas e conseqüentemente a formação de aglomerados nessas regiões.

Os *clusters* identificados próximos a Paragominas (PA) podem ter surgido devido as proximidades de áreas agrícolas do estado do Maranhão e da proximidade da capital Belém (PA), onde há demanda pelo produto e porto para exportações.

Embora não configurados como *clusters* pelas análises, os municípios de Altamira e Novo Progresso no Pará, e os municípios de Roraima possuírem consideráveis aumentos nos valores de AMSS. No caso dos municípios do Pará, a proximidade com o estado do Mato Grosso bem como a existência da BR163 são fatores que possivelmente auxiliaram neste crescimento. No caso dos municípios de Roraima, a existência de uma área com vegetação rasteira (semelhante ao cerrado), aliada a rotas de exportações a países da América do Sul (Venezuela, Guiana, etc) são fatores que podem ter impulsionados tais crescimentos.

O acordo comercial denominado Moratória da Soja, assinado em 2006, prevê a não comercialização da produção da soja provinda de áreas desmatadas dentro do bioma amazônico. Este acordo pode ter pressionado produtores e políticos, sendo possivelmente um dos motivos para a expansão da cultura ficasse relativamente estagnada entre os anos de 2006 a 2010.

Identificar fatores políticos e econômicos que tenham possibilitado essa expansão é uma tarefa árdua devido à alta dinâmica da agricultura, sendo por este motivo não contemplados inteiramente nessa pesquisa.

Além disto, a soja pode se beneficiar de práticas de manejos adotadas por outras culturas, não sendo possível associar diretamente (não são associações diretas) seu cultivo com qualquer fator de caráter ambiental, político e/ou econômico. Um exemplo disto é a presença de áreas com feijão (Clemente et al., 2017) em áreas onde não foi encontradas AMSS.

O estudo envolvendo mais culturas e todos os municípios brasileiros possivelmente trará significativas contribuições no entendimento do comportamento espaço-temporal da soja, uma vez que a agricultura é dinâmica e altamente dependente de diversos fatores.

5 Conclusões

Os padrões de distribuição espacial dos dados de AMSS começaram a se intensificar após os anos 2000, sendo que posteriormente ao ano 2010 houve a formação de *clusters* mais bem definidos no espaço.

As regiões Oeste de Tocantins e Leste de Rondônia possuem os municípios com maior destaque para a formação de aglomerados com altos valores de AMSS. Isto provavelmente se deve a proximidade com o estado do Mato Grosso, maior produtor brasileiro de soja.

A consolidação dos *clusters* de 2013 a 2016 sugere uma expansão das áreas do estado do Mato Grosso para novas áreas no Tocantins e no Pará.

O uso das técnicas de agrupamentos espaciais e espaços temporais se mostraram promissoras para identificar os aglomerados agrícolas de AMSS, necessitando pesquisas com mais culturas.

6 Referências bibliográficas

- Almeida, E. S. de, Perobelli, F. S., & Ferreira, P. G. C. (2008). Existe convergência espacial da produtividade agrícola no Brasil? *Revista de Economia E Sociologia Rural*, 46(1), 31–52. <https://doi.org/10.1590/S0103-20032008000100002>
- Anselin, L. (1995). Local Indicators of Spatial Association-LISA. *Geographical Analysis*, 27(2), 93–115. <https://doi.org/10.1111/j.1538-4632.1995.tb00338.x>
- Anselin, L., & Lozano-Gracia, N. (2008). Errors in variables and spatial effects in hedonic house price models of ambient air quality. *Empirical Economics*, 34(1), 5–34. <https://doi.org/10.1007/s00181-007-0152-3>
- Betarelli Junior, A. A., & Almeida, E. (2009). Os principais fatores internos e as exportações microrregionais brasileiras. *Revista de Economia Contemporânea*, 13(2), 201–227. <https://doi.org/10.1590/S1415-98482009000200002>
- Block, R. (2007). Software Review: Scanning for Clusters in Space and Time. *Social Science Computer Review*, 25(2), 272–278. <https://doi.org/10.1177/0894439307298562>
- Clemente, A. M., Júnior, O. A. de C., Guimarães, R. F., McManus, C., Turazi, C. M. V., & Hermuche, P. M. (2017). Spatial-Temporal Patterns of Bean Crop in Brazil over the Period 1990–2013. *ISPRS International Journal of Geo-Information*, 6(4), 107. <https://doi.org/10.3390/ijgi6040107>
- Embrapa Soja. (2004). *Tecnologia de produção de soja - Paraná 2005. Sistema de Produção* (1st ed.). Londrina, PR, Brasil: Embrapa.
- FAOSTAT. (2016). Production quantities by country - Average 1993-2014.
- IBGE. (2017). Pesquisa Agrícola Municipal – PAM. Sistema IBGE de Recuperação Automática – SIDRA. Retrieved from <https://sidra.ibge.gov.br/tabela/1612>
- Kulldorff, M. (2015). SaTScan User Guide V9.4. Boston. Retrieved from <http://www.satscan.org/>
- Kulldorff, M., Huang, L., & Konty, K. (2009). A scan statistic for continuous data based on the normal probability model. *International Journal of Health Geographics*, 8(1), 58. <https://doi.org/10.1186/1476-072X-8-58>
- Maranduba Júnior, N. G., & Almeida, E. S. De. (2009). Análise de convergência espacial dos repasses da Lei Robin Hood. *Economia E Sociedade*, 18(3), 583–601. <https://doi.org/10.1590/S0104-06182009000300007>
- Marconato, R., Larocca, A. P. C., & Quintanilha, J. A. (2012). Análise do uso de tecnologias em estabelecimentos agropecuários por meio dos índices de Moran

global e local. *Revista de Política Agrícola*, 21(1), 5–21. Retrieved from <https://seer.sede.embrapa.br/index.php/RPA/article/view/68>

- Montenegro, R. L. G., & Betarelli Junior, A. A. (2008). Análise e investigação dos fatores determinantes da inovação nos municípios de São Paulo. *Associação Brasileira de Estudos Regionais E Urbanos*, 2(2), 7–31.
- Novaes, M. R. de, Rudorff, B. F. T., Almeida, C. M. de, & Aguiar, D. A. de. (2011). Análise espacial da redução da queima na colheita da cana-de-açúcar: perspectivas futuras ao cumprimento do protocolo agroambiental. *Engenharia Agrícola*, 31(3), 572–583. <https://doi.org/10.1590/S0100-69162011000300017>
- Perobelli, F. S., Almeida, E. S. De, Alvim, M. I. D. S. a., & Ferreira, P. G. C. (2007). Produtividade do setor agrícola brasileiro (1991-2003): uma análise espacial. *Nova Economia*, 17(1), 65–91. <https://doi.org/10.1590/S0103-63512007000100003>
- Predebon, K. M., Mathias, T. A. D. F., Aidar, T., & Rodrigues, A. L. (2010). Desigualdade sócio-espacial expressa por indicadores do Sistema de Informações sobre Nascidos Vivos (SINASC). *Cadernos de Saúde Pública*, 26(8), 1583–1594. <https://doi.org/10.1590/S0102-311X2010000800012>
- Selemani, M., Mrema, S., Shamte, A., Shabani, J., Mahande, M. J., Yeates, K., ... Lutambi, A. M. (2015). Spatial and space–time clustering of mortality due to malaria in rural Tanzania: evidence from Ifakara and Rufiji Health and Demographic Surveillance System sites. *Malaria Journal*, 14(1), 369. <https://doi.org/10.1186/s12936-015-0905-y>
- Sugumaran, R., Larson, S. R., & DeGroot, J. P. (2009). Spatio-temporal cluster analysis of county-based human West Nile virus incidence in the continental United States. *International Journal of Health Geographics*, 8(1), 43. <https://doi.org/10.1186/1476-072X-8-43>
- Wrublack, S. C., Prudente, V. H. R., Sílvia, E. M., & Coelho, R. M. (2013). Spatial distribution of Canola culture in the State of Paraná (Brazil) between the agricultural years of 2005 and 2009. *Ciencia E Investigacion Agraria*, 40(3). <https://doi.org/10.4067/S0718-16202013000300005>
- Wu, W., Guo, J., Guan, P., Sun, Y., & Zhou, B. (2011). Clusters of spatial, temporal, and space-time distribution of hemorrhagic fever with renal syndrome in Liaoning Province, Northeastern China. *BMC Infectious Diseases*, 11(1), 229. <https://doi.org/10.1186/1471-2334-11-229>
- Zhao, F., Cheng, S., He, G., Huang, F., Zhang, H., Xu, B., ... Wang, L. (2013). Space-Time Clustering Characteristics of Tuberculosis in China, 2005-2011. *PLoS ONE*, 8(12), e83605. <https://doi.org/10.1371/journal.pone.0083605>