

## **Análise Espacial da Produção de Soja no Estado do Paraná, durante as safras de 2003/2004 a 2007/2008**

Victor Hugo Rohden Prudente<sup>1</sup>  
Marlus Luciano do Prado<sup>1</sup>  
Briane Kreitlov e Silva<sup>1</sup>  
Erivelto Mercante<sup>1</sup>  
Miguel Angel Uribe Opazo<sup>1</sup>

<sup>1</sup> Universidade Estadual do Oeste do Paraná - UNIOESTE

Caixa Postal 701 – 85819-110 - Cascavel - PR, Brasil

victor.rohden@yahoo.com, marlusprado@yahoo.com.br, brianekreitlov@yahoo.com.br,  
erivelto@unioeste.br, miguel.opazo@pq.cnpq.br

**Abstract.** In this study, we carried out the analysis of spatial distribution from the production figures for the municipal crop in Parana state. Data were obtained from SEAB (Secretary of Agriculture and Supply of Paraná), and refer to values of five crops, which were consecutive from 2003/2004 to 2007/2008. Were prepared five maps, one for each crop, with the aim of presenting as switching occurred in this time interval. Moran's index was used to show levels of spatial correlation on a global level between the municipalities and local indicator LISA was used in studying the spatial correlation at the local level. By the methods it was possible to demonstrate how cities relate spatially and soybean production, and can indicate clusters of high production as well as clusters of low production, thus enabling the knowledge on the main soybean producing counties, assisting members of the production chain with information so they can improve their plans for the strengthening of culture in Paraná.

**Palavras-chave:** spatial correlation, Moran's index, *LISA*, correlação espacial, índice de Moran, *LISA*.

### **1. Introdução**

Atualmente o Brasil ocupa uma posição de destaque no cenário mundial em produção de grãos de soja, perdendo apenas para os Estados Unidos da América (EUA). Esta posição se deve principalmente ao investimento do país em novas tecnologias para o meio rural e incentivo financeiro para o produtor, aumentando as perspectivas de crescimento. O estado do Paraná vem figurando a várias safras como o segundo maior produtor de soja do Brasil, atrás do estado do Mato Grosso.

Desta maneira, tornou-se fundamental estudar a distribuição espacial desse grande volume de grãos, e para programar estes estudos, necessita-se aprimorar metodologias clássicas em detrimento à inclusão de novas tecnologias, tais como técnicas de sensoriamento remoto, sistemas de informações geográficas - SIGs e estatística espacial, as quais têm sido utilizadas para solucionar problemas e obter informações acuradas e precisas em aplicações agrícolas e ambientais (Dorigo et al., 2007; Prasad et al., 2006).

A evolução da informática é um dos fatores responsáveis pela expansão da estatística espacial de áreas, tendo em vista a praticidade fornecida pelos SIGs para análise de dados (Dalposso et. al, 2008). Entre os diversos métodos de análises espaciais, o Índice Global, ou de Moran (*I*) e o *LISA* (Índice Local de Associação Espacial), se destacam pela sua grande utilização e pela sua forma de análise, onde por meio de uma matriz de peso (*W*), pode-se determinar se há autocorrelação espacial sobre os dados em questão (Anselin, 1995).

O objetivo do presente estudo, foi desenvolver uma metodologia para caracterizar a intensidade da produção de soja no Paraná entre as safras de 2003/2004 a 2007/2008,

analisando a distribuição espacial da produção de soja nos 399 municípios do estado, evidenciando há relação espacial em termos de produção entre os municípios.

## 2. Metodologia de Trabalho

Tendo como base um arquivo *shapefile* que engloba todos os 399 municípios constituintes do estado do Paraná, adicionou-se ao mesmo os dados referentes à produção municipal de soja para cinco safras, sendo elas, 2003/2004, 2004/2005, 2005/2006, 2006/2007, 2007/2008. Os dados referentes à produção de soja são oficiais e foram obtidos junto a SEAB-PR (SEAB, 2010). O arquivo *shapefile* obtido junto ao IBGE (Instituto Brasileiro de Geografia e Estatística) compreende apenas a área territorial total de cada município.

Com o auxílio do *software* Geoda, pode-se agrupar estes dados em um único *shapefile* e realizar assim as análises de correlação espacial. Em tais análises, utilizou-se o Índice de Moran ( $I$ ), o qual é um indicador global da autocorrelação espacial, mostrando como os valores observados de um atributo numa determinada região (polígono) são dependentes desta mesma variável em outra região (polígono).

Para a aplicação do Índice de Moran, é necessária uma matriz de pesos  $W$ , também denominada como matriz de proximidade espacial que pode considerar três critérios de vizinhança, na qual o critério Queené utilizado neste trabalho, considera o elemento  $w_{ij} = 1$ , se a região  $A_i$  compartilha um lado comum com a região  $A_j$ , caso contrário  $w_{ij} = 0$ ; para  $i \neq j = 1, 2, \dots, n$ .

Nos cálculos, se insere a matriz de proximidade espacial normalizada  $W'$ , isto é, multiplicam-se os elementos das linhas desta matriz por um determinado escalar, a fim de que sua soma seja normalizada para o valor 1. Na determinação do Índice de Moran, utiliza-se também o vetor dos desvios  $Z$ , onde se calcula a média da variável ( $\mu$ ), e posteriormente acha-se a componente  $z_i$  para  $i = 1, 2, \dots, n$ , do vetor  $Z$ , sendo  $z_i = (x_i - \mu)$ .

Outro elemento básico, utilizado na determinação do Índice de Moran ( $I$ ), é o vetor das médias ponderadas ( $Wz$ ), que resulta da multiplicação do vetor transposto do vetor dos desvios  $Z^T$  pela matriz de proximidade normalizada  $W'$ , obtendo-se com isto uma matriz  $Wz$  tendo em cada elemento  $Wz_i$  a média dos desvios dos vizinhos deste objeto  $i$ .

O Índice de Moran ( $I$ ), esta expresso na equação (1), a equação dos vetores de desvios esta expresso na equação (2), a equação referente à média ponderada esta apresentada na equação (3).

$$I = \frac{Z^T \cdot Wz}{Z^T \cdot Z} \quad (1)$$

$$Z = (z_i - \mu)1^T \quad (2)$$

$$Wz = Z^T \cdot W' \quad (3)$$

Os valores do Índice de Moran variam de -1 a 1, sendo que os valores de 0 a +1, indicam correlação espacial positivas ( $I > 0$ ), áreas com valores altos (baixos) de um certo atributo, cercados de áreas com valores altos (baixos) deste mesmo atributo. Já valor de 0 a -1, representam correlação negativa ( $I < 0$ ), áreas com valores altos (baixos) de um certo atributo cercado com áreas com valores baixos (altos) deste mesmo atributo. Valor zero ( $I=0$ ) para o Índice de Moran, representa a inexistência de correlação espacial.

No que se refere a testar a significância do Índice de Moran, foi gerada uma distribuição de referência usando valores aleatórios de produção nas áreas e calculando um novo valor de produção 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 que o nível de significância, não se rejeita a hipótese nula.

Para uma melhor interpretação do Índice de Moran, criaram-se gráficos de espalhamento de Moran (*Moran Scatter Plots*), que são gráficos compostos de quatro quadrantes, representando o relacionamento entre os valores do vetor  $Z$  e os valores das médias ponderadas  $Wz$ . O quadrante denominado alto-alto (AA) representa valores positivos para  $Z$  e  $Wz$ , ou seja, áreas com valores altos para certo atributo, cercado por municípios que possuem valores altos para este mesmo atributo; o quadrante denominado baixo-baixo (BB) representa valores negativos para  $Z$  e  $Wz$ , áreas com baixo valor para determinado atributo, cercado de áreas com baixos valores para o mesmo atributo; o quadrante baixo-alto (BA) representa áreas com valores negativos para  $Z$  e positivos para  $Wz$ , sendo estas áreas com baixo valor para certo atributo cercados por vizinhos com alto valor para este mesmo atributo e; o quadrante alto-baixo (AB), que representa valores positivos para o vetor de desvios  $Z$  e valores negativos para a média ponderada  $Wz$ , observando áreas com alto valor para certo atributo cercados por vizinhos com baixo valor para o mesmo atributo.

Neste gráfico o valor ( $I$ ) do Índice de Moran, representa o coeficiente angular da reta de regressão, ou seja, esse valor indica a inclinação da reta de  $Wz$  para  $Z$ .

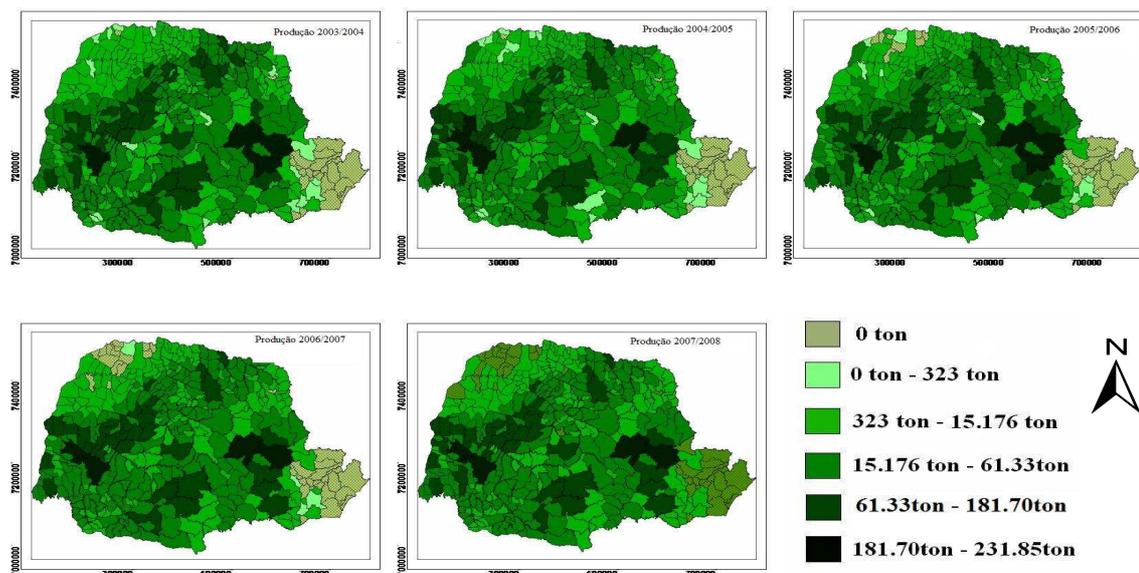
O Índice de Moran fornecem um único valor de associação como medida para todo o conjunto de dados, sendo necessário, desagregar tal índice global, de acordo com seus constituintes locais. O índice local de associação espacial *LISA* constituiu um exemplo da desintegração do Índice de Moran, determinando um valor específico para cada área (polígono), permitindo assim a identificação de regiões com valores atípicos (*outliers*) e de aglomerados com valores semelhantes (*clusters*). O índice *LISA* é representado pela equação (4),

$$I_i = (Z_i \cdot Wz_i) / \sigma^2 \quad (4)$$

em que,  $\sigma^2$  representa a variância dos dados  $x$ .

### 3. Resultados e Discussões

A Figura 1 contém cinco mapas que representam a produção de soja por município nas safras de 2003/2004, 2004/2005, 2005/2006, 2006/2007 e 2007/2008. Observa-se nestes mapas, que não há grande alternância em quantidade de produção de cada município ao longo das cinco safras em questão, sendo que se pode destacar uma predominância dos municípios de Cascavel, Toledo, Assis Chateaubriand, localizados na região Oeste do estado e, Tibagi, Castro, Toledo, Ponta Grossa, localizados na região dos Campos Gerais, com as maiores produções. Enquanto principalmente nos municípios de parte da região norte e de praticamente toda a região litorânea predomina a produção nula ou quase nula.



**Figura 1** - Mapas referentes a produção em toneladas (t) por município nas safras em estudo.

A Tabela 1 apresenta os valores do Índice de Moran ( $I$ ) obtidos em cada safra estudada. Para cada safra foram geradas distribuições empíricas por meio de 999 permutações escolhidas independentes e com igual probabilidade dentre todas as possíveis permutações dos valores de produções entre os municípios estudados.

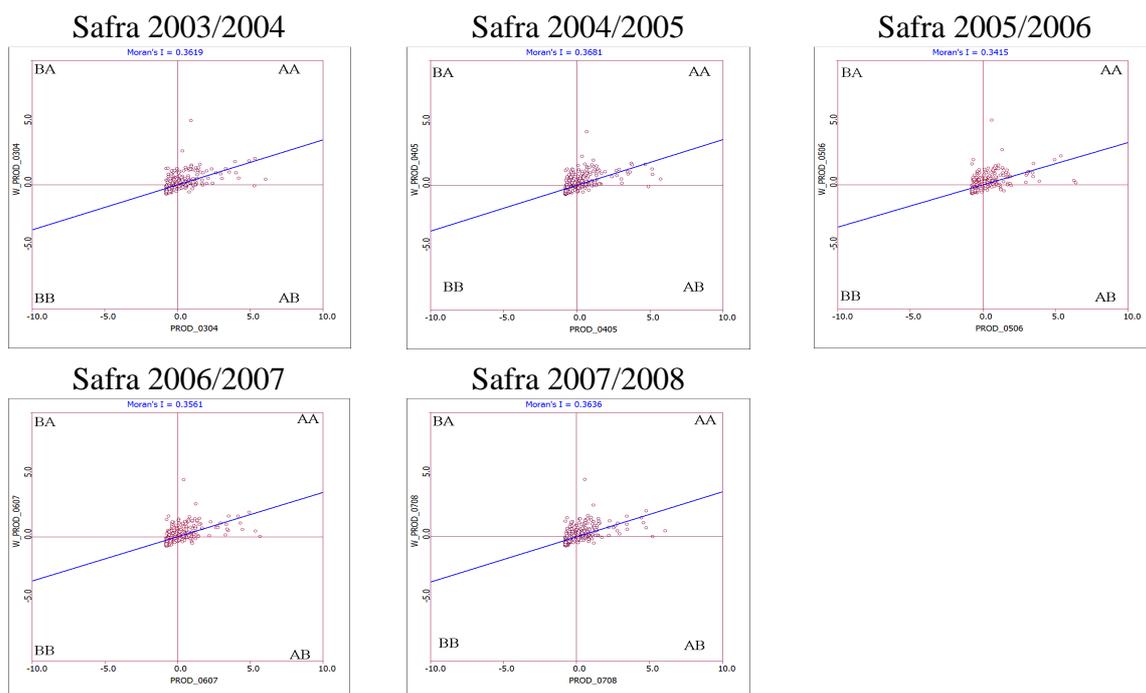
Observa-se na Tabela 1, que todos os valores para o Índice de Moran ( $I$ ) estudado foi maior que zero, sendo assim a hipótese de não existência de correlação espacial descartada.

Tabela 1: Valores do Índice de Moran ( $I$ ) para as safras em estudo.

Safras	( $I$ )
2003/2004	0,3619
2004/2005	0,3681
2005/2006	0,3415
2006/2007	0,3561
2007/2008	0,3636

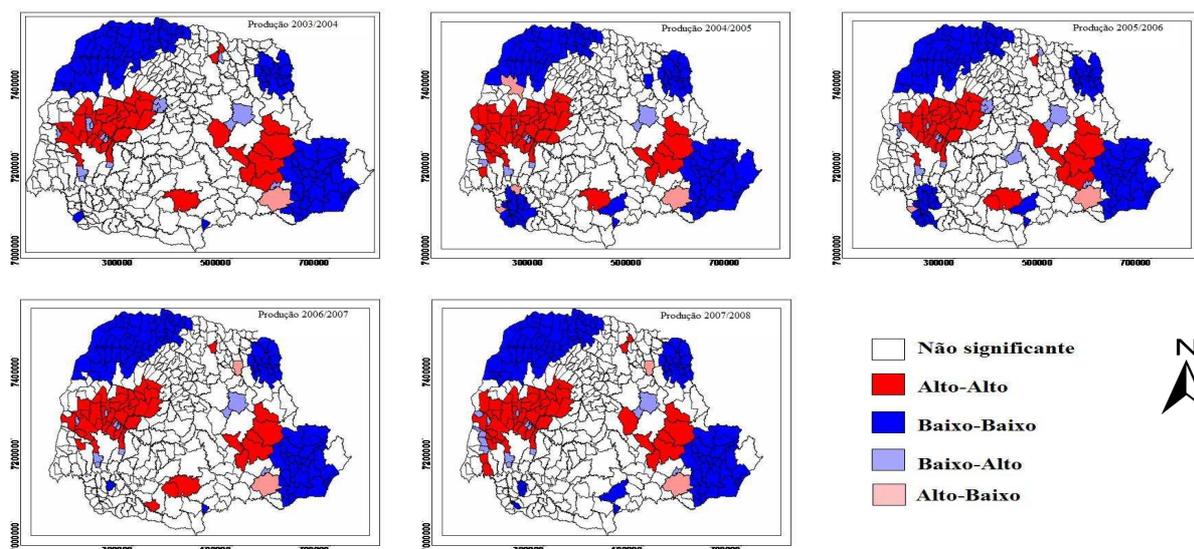
A Figura 2 apresenta os gráficos de espalhamento do Índice de Moran (*Moran Scatter Plots*) para as safras 2003/2004, 2004/2005, 2005/2006, 2006/2007 e 2007/2008. Sendo que o Índice de Moran representa o coeficiente angular da reta de regressão e os índices *LISA* por município os pontos nos gráficos de dispersão, a exemplificação deste em um gráfico de dispersão permite uma fácil interpretação do valor de produção em um determinado município e a média de seus vizinhos. Observa-se nos cinco mapas que há maior concentração dos municípios (*LISA* local) no quadrante do tipo AA (alto-alto) e no quadrante BB (baixo-baixo), o que justifica o valor positivo para o Índice de Moran.

Nota-se na Figura 2 que também há municípios nos quadrantes AB e BA, que retratam uma associação espacial negativa. Há pouca mudança na ordem de apresentação dos dados, sendo que para as cinco safras os municípios praticamente não mudaram as suas ordens de associação espacial, o que nos leva a concluir que municípios altamente produtores continuam sua larga escala produtiva, enquanto municípios nulos ou praticamente não produtores continuam com suas baixas ou nenhuma produção.



**Figura 2** – Gráfico de espalhamento de Moran da produção de soja nas safras em estudo.

Para analisar o valor específico para cada município, o índice *LISA* permite analisar o comportamento da produção com maiores detalhes durante a safra estudada. Desenvolveram-se mapas denominados *LISA CLUSTERS MAP* para cada safra, o qual retrata o gráfico de espalhamento de Moran, determinando os aglomerados com baixas produções e altas produções, aonde apenas os dados com um índice de significativos (0,05 de probabilidade) são apresentados.



**Figura 3** – *LISA CLUSTER MAPS* referentes a distribuição espacial da produção de soja nas safras em estudo.

Na interpretação dos *LISA CLUSTER MAPS*, destaca-se durante as 5 safras, as regiões próximas aos municípios com as maiores produções, formando aglomerados (*clusters*) do tipo AA, ou seja, municípios que possuem alta produção cercados de municípios altamente produtores, destacando assim dois grandes pólos, Oeste do Paraná, municípios localizados próximos as cidades de Cascavel e Toledo e; a região dos Campos Gerais, principalmente os municípios próximos as cidades de Castro e Tibagi. Também houve a formação de dois grandes *clusters* do tipo BB, municípios com pouca ou nenhuma produção, cercados de municípios com nível de produção semelhante, sendo que esses dois *clusters* estão localizados na região litorânea e na região norte do estado.

Através destes dados é possível elaborar planejamentos de infra-estrutura a fim de se otimizar a lucratividade dessa cultura nestas regiões altamente produtoras, principalmente melhorando as condições de escoamento, que se realiza em rodovias diminuindo o lucro para o produtor.

#### **4. Conclusões**

Definiu-se assim um perfil da espacialização da produção para a cultura de soja no estado do Paraná, evidenciando que os municípios formadores da região dos Campos Gerais e da região Oeste, possuem as maiores produções, formando dois grandes aglomerados para a produção da soja. Com estas informações, é possível elaborar um sistema de escoamento e armazenagem dos grãos, para que nestas regiões o produtor consiga escoar ou armazenar sua produção de forma segura e principalmente de forma a otimizar os custos, consequentemente aumentando o lucro para estes produtores.

#### **Agradecimentos**

Os autores agradecem a Secretaria de Agricultura e Abastecimento do Paraná - SEAB-PARANÁ, Fundação Araucária e CNPq.

#### **Referências Bibliográficas**

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

DALPOSSO, G. H.; URIBE-OPAZO, M. A.; MERCANTE E. Analysis of the spatial autocorrelation of bean's productivity for the 2005/2006 harvest in Brazil. In: CIGR International Conference of Agricultural Engineering - XXXVII Congresso Brasileiro de Engenharia Agrícola (CONBEA). Foz do Iguaçu. 2008.

DORIGO, W.A.; ZURITA-MILLA, R.; WIT, A.J.W.; BRAZILE, J. A review on reflective remote sensing and data assimilation techniques for enhanced agroecosystem modeling. *International Journal of Applied Earth Observation and Geoinformation*, v. 9, p. 165–193, 2007.

PRASAD, A.K.; CHAI, L., SINGH, R. P., KAFATOS, M. Crop yield estimation model for Iowa using remote sensing and surface parameters. *International Journal of Applied Earth Observation and Geoinformation*, v. 8, p. 26-33, 2006.

SEAB. Secretaria da Agricultura e do Abastecimento do Paraná. em: <<http://www.seab.pr.gov.br/>> Acesso em Outubro de 2010.