

## Autocorrelação espacial do índice NDVI de imagens TM para cultura da soja no Oeste Paranaense, no ano agrícola 2004/2005

Gustavo Henrique Dalposso<sup>1</sup>  
Miguel Angel Uribe Opazo<sup>1</sup>  
Erivelto Mercante<sup>1</sup>  
Rubens Augusto Camargo Lamparelli<sup>2</sup>  
Jansle Vieira Rocha<sup>3</sup>

<sup>1</sup> Universidade Estadual do Oeste do Paraná - UNIOESTE  
Caixa Postal 701 – 85819-110 - Cascavel - PR, Brasil  
gustavodalposso@hotmail.com, mopazo@unioeste.br, erivelto@unioeste.br

<sup>2</sup> Centro de Pesquisas Meteorológicas e Climáticas Aplicadas a Agricultura – CEPAGRI  
13083-970 - Campinas - SP, Brasil  
rubens@cpa.unicamp.br

<sup>3</sup> Faculdade de Engenharia Agrícola – FEAGRI / UNICAMP  
13083-875 - Campinas - SP, Brasil  
jansle.rocha@agr.unicamp.br

**Abstract.** In this work we present a spatial statistic study using the soybean vegetation index NDVI obtained from images in west of Paraná State. The images were collected by TM (Thematic Mapper) sensor of Landsat 5 during the 2004/2005 harvest. The data were grouped and divided in five classes of equal magnitude, allowing building a map of soybean crop temporal evolution. The Moran Index  $I$  was used to study the spatial correlation in a global level and the LISA index to study the correlation in a local level. Through the indexes is possible identify the culture profile in municipalities of the region, allowing producers, cooperatives, consumers and other members of soybean chain become more informed and can better trace their plans for strengthening of culture.

**Palavras-chave:** spatial statistics, índice  $I$  de Moran, LISA, Landsat, estatística espacial, índice  $I$  de Moran.

### 1. Introdução

O Brasil sempre ocupou lugar de destaque entre os maiores produtores de soja. Já faz alguns anos que ele ocupa a segunda colocação mundial sendo que na safra atual contou com uma produção de aproximadamente 50,19 milhões de toneladas, figurando atrás somente dos Estados Unidos da América (EUA), que produziram 82,82 milhões de toneladas. A soja, como lavoura comercial, chegou ao Estado do Paraná em meados dos anos 50 e até então, a cultura era uma quase curiosidade. Sua produção era pequena e as reduzidas lavouras de soja existentes na região destinavam-se ao consumo doméstico sendo que o total da produção não passava de 60 toneladas (EMBRAPA, 2008). Na safra atual estima-se uma produção de 11.896.000 toneladas para o Estado, que na safra estudada (2004/2005) obteve uma produção de 9.707.300 toneladas, ficando atrás do Estado do Mato Grosso que produziu 17.937.100 toneladas (CONAB, 2008).

Uma das alternativas para realizar estimativas da produção agrícola em nível regional e nacional é a utilização de tecnologias da área de sensoriamento remoto, pois elas permitem a obtenção de informações precisas, em tempo hábil e relativamente com baixo custo, sobre a extensão, as condições de desenvolvimento e o potencial de produção das culturas (MOTTA et al., 2001). Entre os diversos produtos provenientes de imagens de satélite, pode-se citar os índices de vegetação, que são combinações de dados espectrais de duas ou mais bandas, selecionadas com o objetivo de sintetizar e melhorar a relação desses dados com os parâmetros biofísicos da vegetação (MOREIRA, 2000). EPIPHANIO et al. (1996) relatam

que a relação entre sensoriamento remoto e os parâmetros biofísicos pode ser feita por meio da aplicação de índices de vegetação.

Entre os diversos índices de vegetação existentes, o índice de vegetação por diferença normalizada (NDVI - *Normalized Difference Vegetation Index*) (ROUSE et al., 1973), obtido a partir da razão entre a subtração e soma das refletâncias das bandas da região do infravermelho próximo e do vermelho do espectro eletromagnético, é o que vem sendo mais largamente utilizado (KLERING et. al, 2007).

Em estudos de monitoramento, freqüentemente é buscado o estabelecimento de relações entre o padrão de evolução do crescimento das plantas ao longo do tempo com o padrão de evolução temporal do NDVI. A utilidade do índice NDVI foi demonstrada em diversos campos, tais como, agricultura de precisão, estudo da fenologia da vegetação, em classificações espectrais e avaliações primárias de produtividades (SUGAWARA, 2001). A evolução da informática é um dos fatores responsáveis pela expansão da estatística espacial de áreas, visto que a praticidade fornecida pelos sistemas de informação geográfica, para análise de dados, está atraindo um número cada vez maior de pesquisadores, o que resulta em trabalhos nos diversos campos do conhecimento.

Neste contexto, o objetivo deste trabalho é estudar o comportamento espectral da biomassa da soja em 36 municípios da região Oeste do Paraná a partir da análise da autocorrelação espacial do índice NDVI de dados referentes a refletância de superfície das imagens do sensor TM, facilitando a elaboração de estratégias aos participantes da cadeia produtiva da soja.

## 2. Material e Métodos

Construiu-se um banco de dados espaciais com 36 municípios da região Oeste do Paraná alocados dentro da cena 223/77 do satélite Landsat 5, utilizando um mapa da rede municipal e um conjunto de índices NDVI de refletância de superfície da safra 2004/2005 obtidos por MERCANTE (2007) nas datas de 23/11/2004, 09/12/2004, 25/12/2004, 26/01/2005 e 11/02/2005. A Figura 1 apresenta os 36 municípios da Região Oeste.

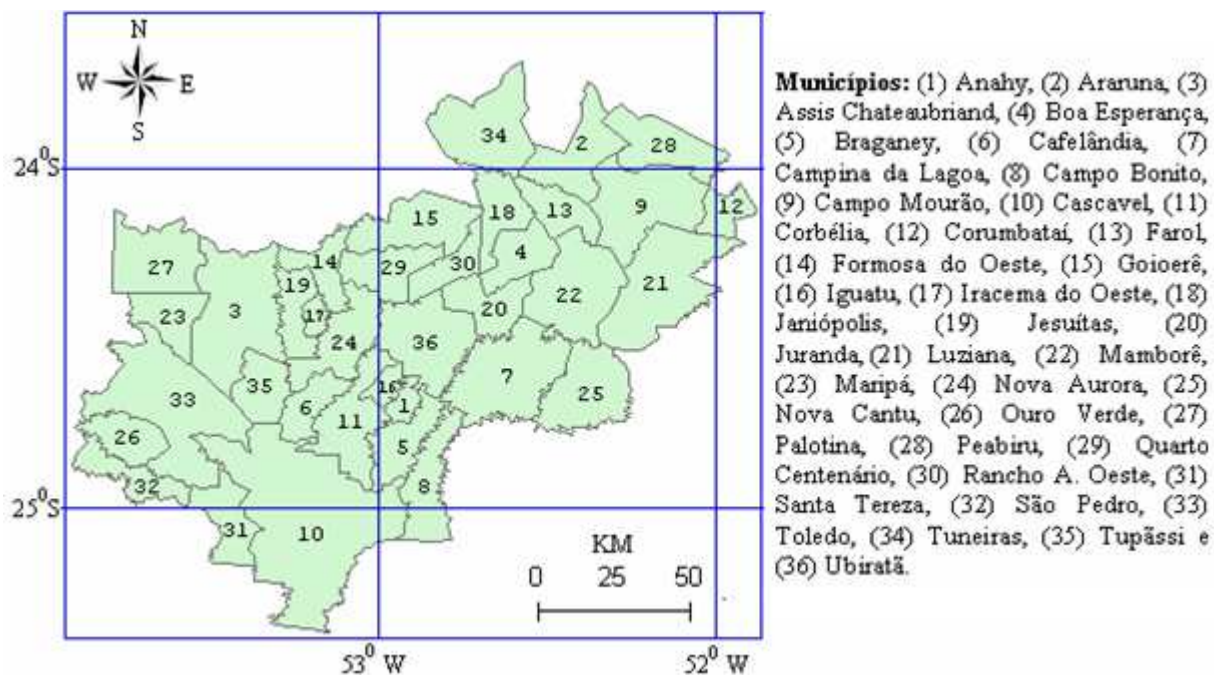


Figura 1 – Localização dos 36 municípios estudados na Região Oeste do Paraná.

Os dados de refletância de superfície das imagens TM utilizados para gerar o índice NDVI, foram obtidos com emprego do sistema de correção de imagens de satélite denominado SCORADIS, desenvolvido por ZULLO JR. (1994), a partir do qual se estimou o sinal recebido pelo satélite sem a interferência de fatores atmosféricos como gases e aerossóis, tendo como o objetivo melhorar a qualidade das imagens. Este procedimento foi descrito de forma mais completa por MERCANTE (2007).

Visando obter uma melhor compreensão da dinâmica espacial do NDVI na região estudada, utilizaram-se alguns recursos da análise exploratória de dados espaciais. A ESDA, *exploratory spatial data analysis*, é definida na literatura como uma coleção de técnicas para descrever e visualizar distribuições espaciais, identificar situações atípicas, descobrir padrões de associação espacial, agrupamento de valores semelhantes (*clusters*) e sugerir regimes espaciais ou outras formas de heterogeneidade espacial (ANSELIN e BAO, 1997). A verificação da dependência espacial foi realizada por meio do Índice de Moran (I), apresentado na Equação (1) na sua forma matricial, o qual é um indicador global da autocorrelação espacial e mostra como os valores estão correlacionados no espaço.

Com auxílio do *software* Geoda foram criados a matriz de proximidade espacial (W), utilizando o critério de vizinhança *Queen*<sup>1</sup> para definição das fronteiras, o vetor dos desvios apresentado na Equação (2), e o vetor de médias ponderadas apresentado na Equação (3), que é o produto entre o vetor Z transposto e a matriz W normalizada<sup>2</sup>.

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

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

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

O índice I de Moran presta-se a um teste de ausência de correlação espacial ( $\rho = 0$ ). Os valores positivos do índice I de Moran (entre 0 e +1) indicam correlação positiva ( $\rho > 0$ ), isto é, áreas que apresentam valores altos (baixos) de um atributo estão cercadas de áreas com valores altos (baixos) e os valores negativos do índice I de Moran, (entre -1 e 0) representam correlação negativa ( $\rho < 0$ ), ou seja, áreas com valores altos (baixos) de um atributo estão cercadas por áreas com valores baixos (altos). Para testar a significância do índice I de Moran, foi gerada uma distribuição de referência aleatorizando os valores dos índices de vegetação nas áreas e calculando um novo valor do índice para cada permutação feita, permitindo a realização do seguinte teste de hipóteses:  $H_0: \rho = 0$ , que representa a hipótese nula de ausência de correlação espacial versus a hipótese alternativa  $H_1: \rho \neq 0$ . 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. Ao contrário, se o nível descritivo for maior que o nível de significância, não se rejeita a hipótese nula.

Para complementar os estudos da autocorrelação espacial, foram criados gráficos de espalhamento de Moran (*Moran scatter plots*), que são gráficos que auxiliam a interpretação do índice I de Moran e também se realizou estudos no nível local utilizando a estatística LISA (ANSELIN, 1995). O *Moran scatter plot* procura visualizar espacialmente o relacionamento entre os valores observados Z e os valores das médias locais  $W_z$ . Este gráfico é dividido em quatro quadrantes: AA (Z positivo e  $W_z$  positivo) e BB (Z negativo e  $W_z$  negativo), que indicam pontos de associação espacial positiva, no sentido que uma localização possui

<sup>1</sup> O modelo *Queen* considera fronteiras em comum com extensão não nula e nós comuns.

<sup>2</sup>  $W'$  representa a matriz W normalizada, ou seja, cada linha é multiplicada por uma constante de tal forma que a soma dos valores da linha seja 1.

vizinhos com valores parecidos e AB ( $Z$  positivo e  $W_z$  negativo) e BA ( $Z$  negativo e  $W_z$  positivo), que indicam pontos de associação espacial negativa.

Quando se trabalha com um grande número de áreas é muito provável que ocorram diferentes regimes de associação espacial, o que torna conveniente o uso de técnicas que desagreguem as estatísticas globais segundo seus constituintes locais, como por exemplo, o índice local de associação espacial (*LISA*). Esta estatística produz um valor específico para cada área, permitindo a identificação de aglomerados com valores semelhantes e valores atípicos. Este índice é fornecido por  $I_i = (Z_i \cdot W_{zi}) / \sigma^2$ , em que  $\sigma^2$  representa a variância da distribuição dos valores dos desvios.

### 3. Resultados e Discussão

A Figura 2 apresenta a evolução temporal do índice de vegetação. Os intervalos foram definidos agrupando os valores de NDVI de todas as datas para obter o máximo (0,8750) o mínimo (0,2609) e amplitude geral (0,61403) dos índices. Posteriormente, dividiu-se a amplitude geral pelo número de classes desejado (5) para obter o intervalo de classe. Observe-se que a evolução temporal dos índices de vegetação esteve relacionada com as diferentes fases do ciclo fenológico da cultura, pois os valores mais baixos dos índices são observados em 23/11/04, onde a cultura já está em desenvolvimento, porém, em período inicial e os valores mais altos se encontram entre 26/01/05 e 11/02/05, período esse de maior desenvolvimento das plantas. É possível observar uma diferença nas datas de plantio, visto que alguns municípios têm seus ciclos mais tardios que os outros.

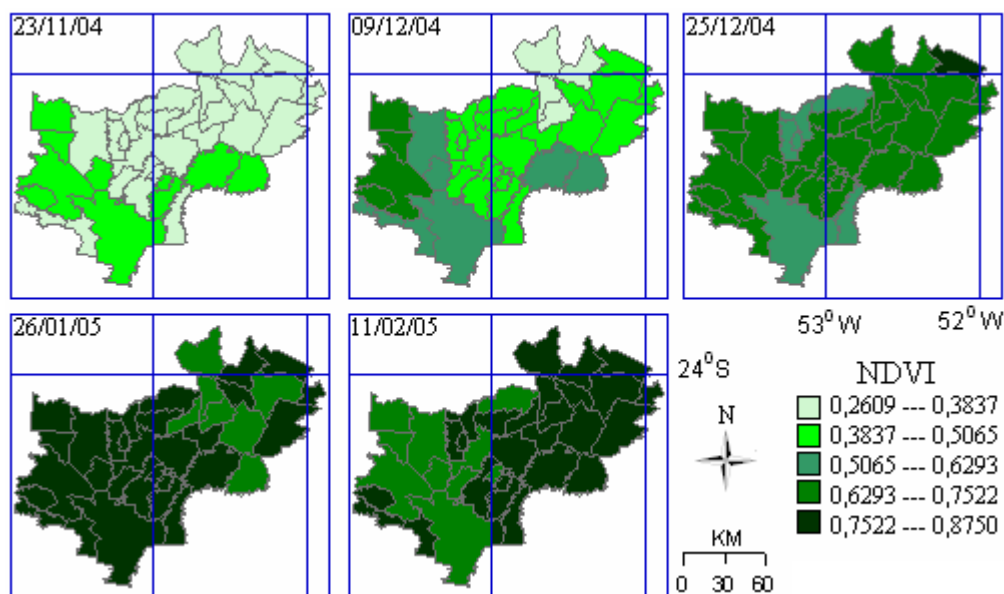


Figura 2 – Evolução temporal do índice NDVI utilizando intervalos de mesma amplitude.

A Tabela 1 apresenta o índice  $I$  de Moran e o seu respectivo nível descritivo (p-valor) referente às hipóteses de correlação espacial. Para cada data foram geradas distribuições empíricas por meio de 999 permutações escolhidas independentes e com igual probabilidade dentre todas as possíveis permutações dos índices NDVI entre os municípios estudados.

Segundo a Tabela 1 observa-se que, para todas as datas em estudo, o índice  $I$  de Moran obtido foi maior que zero, assim, existe uma correlação positiva indicando que a região estudada apresenta municípios com altos (ou baixos) índices de vegetação cercados por municípios com a mesma situação.

TABELA 1 – Índice  $I$  de Moran e nível descritivo (p-valor) do teste de falta de correlação espacial

Data	$I$	p-valor
23/11/04	0,6568	0,0010*
09/12/04	0,7332	0,0010*
25/12/04	0,4193	0,0010*
26/01/05	0,5749	0,0010*
11/02/05	0,7153	0,0010*

\*Rejeita-se a hipótese de falta de correlação espacial ao nível de 5% de probabilidades.

Os níveis descritivos (p-valores) referentes a todas as datas em estudo são inferiores aos níveis de significâncias estipulados (0,05 de probabilidade), mostrando que todos os índices  $I$ 's de Moran calculados são significativos o que implica a existência de correlação espacial dos NDVI's nas datas monitoradas, sugerindo que existem outras causas que contribuem para os padrões espaciais analisados.

A Figura 3 apresenta o gráfico de espalhamento de Moran em 23/11/04 (primeira data analisada) e em 11/02/05 (última data analisada). O índice  $I$  de Moran representa o coeficiente angular da reta de regressão da variável  $W_z$  em função de  $Z$ , o que permite estudar a relação entre o índice de vegetação e a média de seus vizinhos através de um gráfico de dispersão. A maior quantidade de municípios alocados nos quadrantes AA e BB enfatizam a autocorrelação positiva nas duas datas, porém, em vinte e três de novembro observou-se que apenas um município apresenta um NDVI alto cercado de vizinhos com NDVI baixo, sendo alocado no quadrante AB, enquanto no dia onze de fevereiro esse número atinge sete municípios. Como se verificou que o plantio na região não é homogêneo, no sentido de que um expressivo número de municípios, principalmente localizados nas regiões norte e nordeste, apresenta um ciclo vegetativo da soja mais tardio, é natural que na última data aconteça um acréscimo no número de municípios com NDVI alto cercado por municípios com média do NDVI baixa, pois enquanto certos municípios estão atingindo o pico vegetativo, seus vizinhos já passaram do período de maior biomassa, apresentando um NDVI baixo.

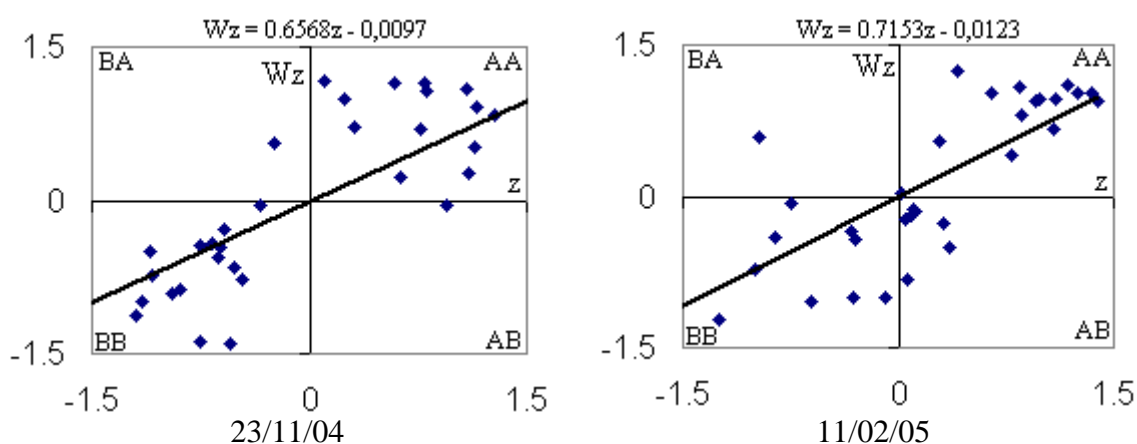


Figura 3 – Moran Scatter do índice NDVI da soja em 23/11/04 e 11/02/05.

A porcentagem de municípios alocados em cada quadrante do gráfico de espalhamento de Moran é apresentada na Tabela 2. Enfatiza-se o resultado mencionado acima, pois na primeira data (23/11/04), 2,78 % dos municípios se encontravam em AB, enquanto que na segunda data (11/02/05) esse valor passou para 19,44 %.

TABELA 2 – Porcentagens de municípios em cada quadrante do gráfico de espalhamento de Moran

DATA	AA	AB	BB	BA
23/11/04	44,44 %	2,78 %	50,00 %	2,78 %
09/12/04	30,56 %	5,56 %	55,56 %	08,33 %
25/12/04	44,44 %	11,11 %	30,56 %	13,89 %
26/01/05	38,89 %	5,56 %	38,89 %	16,67 %
11/02/05	41,67 %	19,44 %	36,11 %	2,78 %

Conforme observado, pontos que se enquadram nos quadrantes AB e BA do gráfico de espalhamento de Moran representam municípios com associação espacial negativa, ou seja, podem-se entender como municípios com alto (baixo) NDVI cercado por municípios com baixo (alto) índice de vegetação. Observando a Tabela 2 é possível constatar que a data que apresenta a maior porcentagem de municípios alocados nos quadrantes AB e BA é 25/12/04 com 25 % dos municípios. Pode-se entender essa situação como um momento de transição, visto que nas regiões aonde o NDVI é alto, alguns municípios começam a apresentar índices menores, em virtude de estarem passando do pico vegetativo. Em contrapartida, municípios que iniciaram o plantio posteriormente, começam a apresentar valores elevados, em virtude de estarem atingindo o estágio de maior biomassa.

Por fornecer um valor específico para cada município, o índice *LISA* permite analisar com maiores detalhes o comportamento dos índices de vegetação da soja durante a safra estudada. A significância do índice local *LISA* é calculada de maneira similar ao caso do índice global, utilizando uma pseudo-distribuição. Para melhor interpretar os índices locais de associação espacial é conveniente utilizar o *LISA MAP* (Figura 4) onde apenas as regiões que apresentam correlação local significativamente diferente do resto dos dados são indicadas. Pela Figura 4 é possível observar o comportamento das correlações espaciais no período estudado. De maneira geral, destacam-se os conglomerados de municípios nas regiões sudoeste e nordeste do mapa. A menor quantidade de índices significativos ocorre em 25/12, visto que apenas seis municípios são classificados.

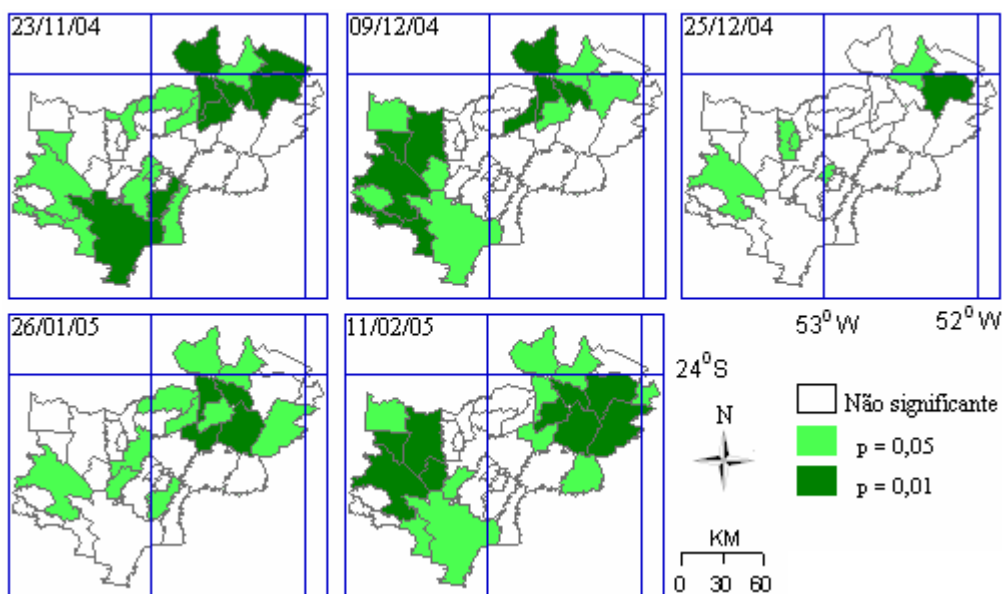


Figura 4 – *LISA MAP* do NDVI da soja na safra 2004/2005. Apenas os valores com significância maior que 95% são mostrados.



A Figura 5 apresenta o *LISA CLUSTER MAP*, que associa ao *LISA MAP* a classificação do diagrama de espalhamento de Moran. Observa-se a existência de dois aglomerados com características distintas. Um deles formado por municípios do sul e oeste e o outro formado por municípios fixados no nordeste.

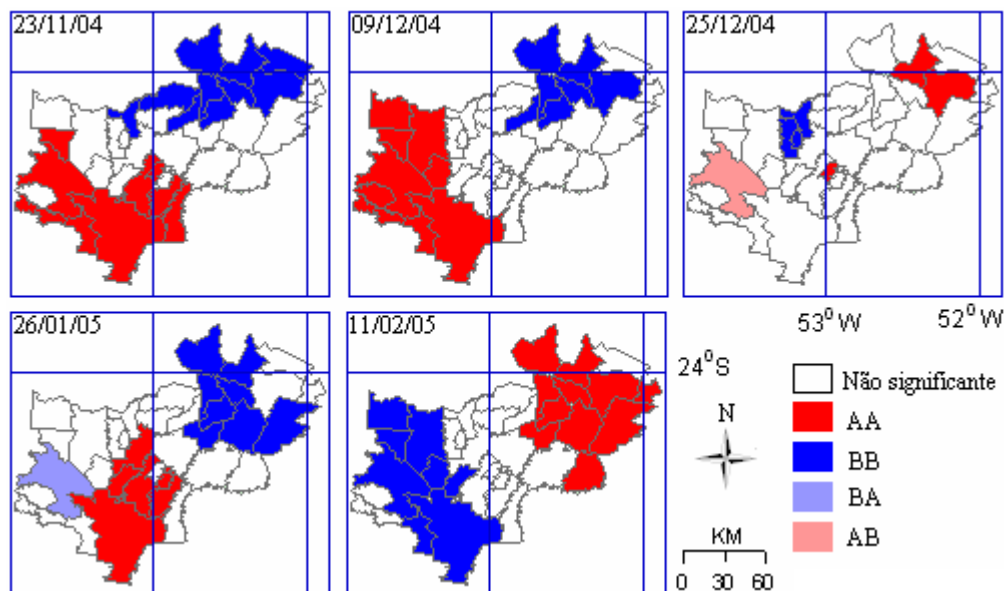


Figura 5 – *LISA CLUSTER MAP* do NDVI da soja na safra 2004/2005.

Essas áreas refletem as diferenças nas datas de plantio no Oeste do Paraná, pois em 23/11/04 os municípios de Cascavel e Toledo apresentavam índices de vegetação altos cercados por municípios com a mesma característica, indicando que a cultura já estava em desenvolvimento, e a região norte apresentava um aglomerado de índices baixos, evidenciando o início do plantio da cultura. A data de 25/12/04 pode ser entendida como um período de transição, pois a região do *cluster* AA começa a obter valores mais baixos e a região do *cluster* BB começa a receber valores baixos. Em 11/02/05 observa-se o comportamento inverso ao ocorrido em 23/11/04, pois o *cluster* que no início apresentou valores altos e agora é composto de valores baixos, indicando o término do ciclo vegetativo e o *cluster* que inicialmente era baixo apresenta agora valores altos, mostrando que a cultura ainda está em desenvolvimento. Essas identificações fornecem valiosas informações tanto para as indústrias que utilizam a soja como matéria prima quanto para os agricultores que são os fornecedores, pois as diferenças nos períodos de plantio indicam que é necessário um planejamento mais cuidadoso da logística de abastecimento, no intuito de que não falte matéria prima para a indústria e também seja uma tarefa rentável para o produtor rural.

#### 4. Conclusões

Os resultados mostraram que a estatística espacial de área associada aos recursos do sensoriamento remoto fornece informações de valiosa importância sobre o atributo observado. No exemplo da soja, foi possível compreender o perfil da cultura nos municípios da Região Oeste do Paraná estudados, identificando conglomerados de municípios com características de plantio semelhantes, no sentido de existirem regiões que realizam um plantio mais tardio. Com essas informações, o produtor pode planejar junto com a indústria qual o momento mais apropriado para fazer o plantio, evitando a sobrecarga de abastecimento e com isso conseguindo um melhor preço pelo fornecimento.

## 5. Agradecimentos

Os autores agradecem a CAPES e Fundação Araucária pelo apoio financeiro.

## 6. Referências Bibliográficas

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

ANSELIN, L.; BAO S. Exploratory Spatial Data Analysis Linking SpaceStat and ArcView. In: Fischer, M. M. & Getis, A. Recent developments in spatial analysis. New York: Springer, p. 35-59. 1997.

CONAB. Soja – Brasil - Série histórica de produtividade (1976 – 2007). Disponível em: <<http://www.conab.gov.br/conabweb/download/safra/SojaSerieHist.xls>> Acesso em: 12. out. 2008.

EMBRAPA. Tecnologias de produção de soja — Paraná – 2004 – A soja no Brasil. Disponível em: <<http://www.cnpso.embrapa.br/producaosojapr/SojanoBrasil.htm>> Acesso em: 12. out. 2008.

EPIPHANIO, J.C.N.; ALMEIDA JÚNIOR, A.C.; FORMAGGIO, A.R. Desenvolvimento do trigo avaliado com dois índices de vegetação. [CD-ROM] In: Simpósio Brasileiro de Sensoriamento Remoto (SBSR), 8, Salvador, 1996, Anais... São Paulo: Imagem Multimídia.

KLERING, E. V.; CUSTÓDIO, M. S.; FONTANA D. C.; BERLATO, M. A. Relação entre os perfis temporais de NDVI/MODIS da cultura do arroz irrigado, a insolação e o rendimento na região da Campanha do Rio Grande do Sul. In: Simpósio Brasileiro de Sensoriamento Remoto (SBSR), 8, Florianópolis, 2007.

MERCANTE, E. Dinâmica Espectral da Cultura da Soja ao Longo do Ciclo Vegetativo e sua Relação com a Produtividade na Região Oeste do Paraná. Tese de Doutorado. FEAGRI – UNICAMP. Campinas. 2007.

MOREIRA, R. C. Influência do posicionamento e da largura de bandas de sensores remotos e dos efeitos atmosféricos na determinação de índices de vegetação. 114p. (Dissertação de Mestrado). Instituto de Pesquisas Espaciais, São José dos Campos, 2000.

MOTTA, J.L.G.; FONTANA, D.C.; WEBER, E. Verificação da acurácia da estimativa de área cultivada com soja através de classificação digital em imagens Landsat. In: Simpósio Brasileiro de Sensoriamento Remoto (SBSR), 10, 2001, Foz do Iguaçu, Anais... São José dos Campos: INPE, 2001.

ROUSE, J. W.; HAAS, R. H.; SCHELL, J. A.; DEERING, D. W. Monitoring vegetation systems in the great plains with ERTS. In: Earth Resources Technology Satellite – 1 Symposium, 3., Washington, D. C., 1973. Proceedings. Washington, D. C.: NASA. Goddard Space Flight Center, v. 1, p. 309-317. (NASA SP-351).1973.

SUGAWARA, L. M. Avaliação de modelo agrometeorológico e imagens NOAA/AVHRR no acompanhamento e estimativa de produtividade da soja no estado do Paraná. 181 p. Dissertação (mestrado em Sensoriamento Remoto) INPE, São José dos Campos, 2001.

ZULLO JR., J. Correção atmosférica de imagens de satélite e aplicações. 189 p. Tese (Doutorado em Engenharia Elétrica) – Faculdade de Engenharia Elétrica, UNICAMP. Campinas. 1994.