

## **Análise temporal de municípios produtores de cana-de-açúcar no estado de São Paulo por meio de agrupamento do NDVI (AVHRR/NOAA) e dados de produtividade e área**

Renata Ribeiro do Valle Gonçalves<sup>1</sup>

Luciana Alvim Santos Romani<sup>2,3</sup>

Jurandir Zullo Jr.<sup>1</sup>

Ana Maria Heuminski de Ávila<sup>1</sup>

Agma Juci Machado Traina<sup>2</sup>

<sup>1</sup>Cepagri / Feagri – Unicamp  
Cidade Universitária – 13083-970 - Campinas - SP, Brasil  
{renata, jurandir, avila}@cpa.unicamp.br

<sup>2</sup>Universidade de São Paulo – USP/ICMC  
Av.Trabalhador são-carlense, 400 - 13566-590 - São Carlos - SP, Brasil  
agma@icmc.usp.br

<sup>3</sup>Embrapa Informática Agropecuária  
Caixa Postal 6041 – 13083-886 - Campinas - SP, Brasil  
luciana@cnptia.embrapa.br

**Abstract.** Sugar cane has become strategic to the Brazilian economy, especially due to its importance in the replacement of fossil fuels. In recent years, the development of bi-fuel car engines led to a fast increase in the demand for ethanol, one of the most important derivatives of sugar cane. Therefore, it is very important to monitor the crop seasons in order to estimate the yield of sugar cane as well as to evaluate the expansion of this crop in the country. There are several models to estimate the sugar cane production involving physiological, meteorological and pedological variables. However, the use of satellite images to aid in the monitoring of sugar cane production has increased greatly in recent years. In this context, this paper presents an assessment of the distribution of sugar cane regions in the last eight crop seasons, according to three different clusters. The test area was the most important region of sugar cane production in the São Paulo state, Brazil. In the experiments, we have used annual variables of NDVI (from AVHRR/NOAA), planted area and productivity values. These variables were used to define the three clusters by a clustering technique. Results showed that there was an increasing in the number of regions in the cluster where the planted area is greater over the years, confirming the sugar cane expansion to other areas.

**Palavras-chave:** remote sensing, K-means, biofuel, sensoriamento remoto, K-médias, biocombustível.

### **1. Introdução**

A cultura da cana-de-açúcar possui função cada vez mais estratégica na economia do país devido à necessidade da substituição dos combustíveis fósseis por fontes de energia renovável, como a produção de etanol, com o propósito de diminuir as emissões de gases de efeito estufa (GEEs). O aumento da concentração desses gases, resultante da intensificação das atividades antrópicas, tem causado alterações significativas no clima (IPCC, 2007). A temperatura da superfície do planeta aumentou, aproximadamente, 0,2°C por década nos últimos trinta anos, e 0,8°C no século passado. Este fato ilustra a importância da cultura da cana-de-açúcar para o país e evidencia a necessidade da existência de sistemas precisos, antecipados e objetivos de monitoramento e previsão de safras que auxiliem o planejamento envolvido na produção e comercialização de produtos estratégicos para os mercados interno e externo, como são o açúcar e o álcool.

Atualmente, o Brasil cultiva 7,2 milhões de hectares (ha) de cana-de-açúcar e a produção está fortemente concentrada no estado de São Paulo, com 59% do total nacional. Segundo Manzatto et al. (2009), as estimativas obtidas demonstram que o país dispõe de cerca de 64,7 milhões de ha de áreas aptas à expansão do cultivo com cana-de-açúcar, sendo essas áreas

cultivadas com pastagens, em 2002, representando cerca de 37,2 milhões de ha. Essas estimativas demonstram que o país não necessita incorporar áreas novas e com cobertura nativa ao processo produtivo, podendo expandir ainda a área de cultivo com cana-de-açúcar sem afetar diretamente as terras utilizadas para a produção de alimentos.

Entretanto, séries temporais de imagens de satélites meteorológicos, como os da série AVHRR/NOAA (*Advanced Very High Resolution Radiometer / National Oceanic and Atmospheric Administration*), estão sendo muito utilizadas no monitoramento da vegetação em escalas regional e global. Vários estudos vêm sendo realizados visando o aumento da utilização efetiva de imagens de satélite no monitoramento agrícola. Exemplos de estudos são a obtenção de modelos que sejam aplicáveis, simultaneamente, a vários municípios produtores de cana-de-açúcar e a avaliação das relações entre o NDVI (*Normalized Difference Vegetation Index*) e o ISNA (Índice de Satisfação das Necessidades de Água) (Gonçalves et al. 2009); a produtividade (Nascimento et al. 2009); e a detecção da similaridade entre os municípios por meio de funções de distância (Romani et al. 2010).

Devido à sua alta resolução temporal, garantia de cobertura global e gratuidade das imagens, é elevada a probabilidade de obtenção de imagens em boas condições ao longo do ciclo de desenvolvimento da cultura comercial. Os satélites da série NOAA têm grande potencial de utilização em métodos operacionais e objetivos de previsão de safras, tanto na estimativa da área plantada quanto na da produtividade. Uma das etapas na estimativa de produtividade é o conhecimento sobre as diferentes áreas produtoras de cana-de-açúcar numa determinada região. Dessa forma, o objetivo deste trabalho foi detectar municípios produtores de cana-de-açúcar no estado de São Paulo, similares, pelo método de agrupamento, analisando variáveis espectrais (NDVI), área plantada e produtividade no período de 2001 a 2009. O resultado dessa análise exploratória dos dados mostrou que a técnica é apropriada para a determinação de grupos de municípios com características semelhantes o que permite classificar as regiões automaticamente.

## 2. Metodologia do trabalho

A área de estudo foi o estado de São Paulo, que está situado entre as coordenadas geográficas 54° e 43°30' de longitude oeste e 25°30' e 19° 30' de latitude sul. Foram selecionados os 50 municípios maiores produtores de cana-de-açúcar do estado, como ilustra a Figura 1. Esses municípios foram selecionados pelos dados de produção do IBGE no período de 2001 a 2009.

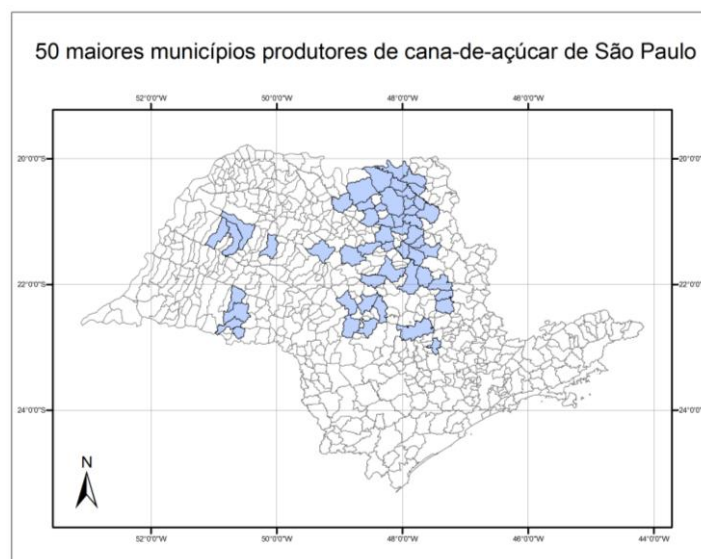


Figura 1. Cinquenta maiores municípios produtores de cana-de-açúcar do estado de São Paulo.

Após a seleção dos municípios, as imagens foram processadas, os valores de NDVI extraídos e o agrupamento dos dados foi gerado utilizando o próprio NDVI, a área plantada e a produtividade, conforme o Fluxograma ilustrado na Figura 2.

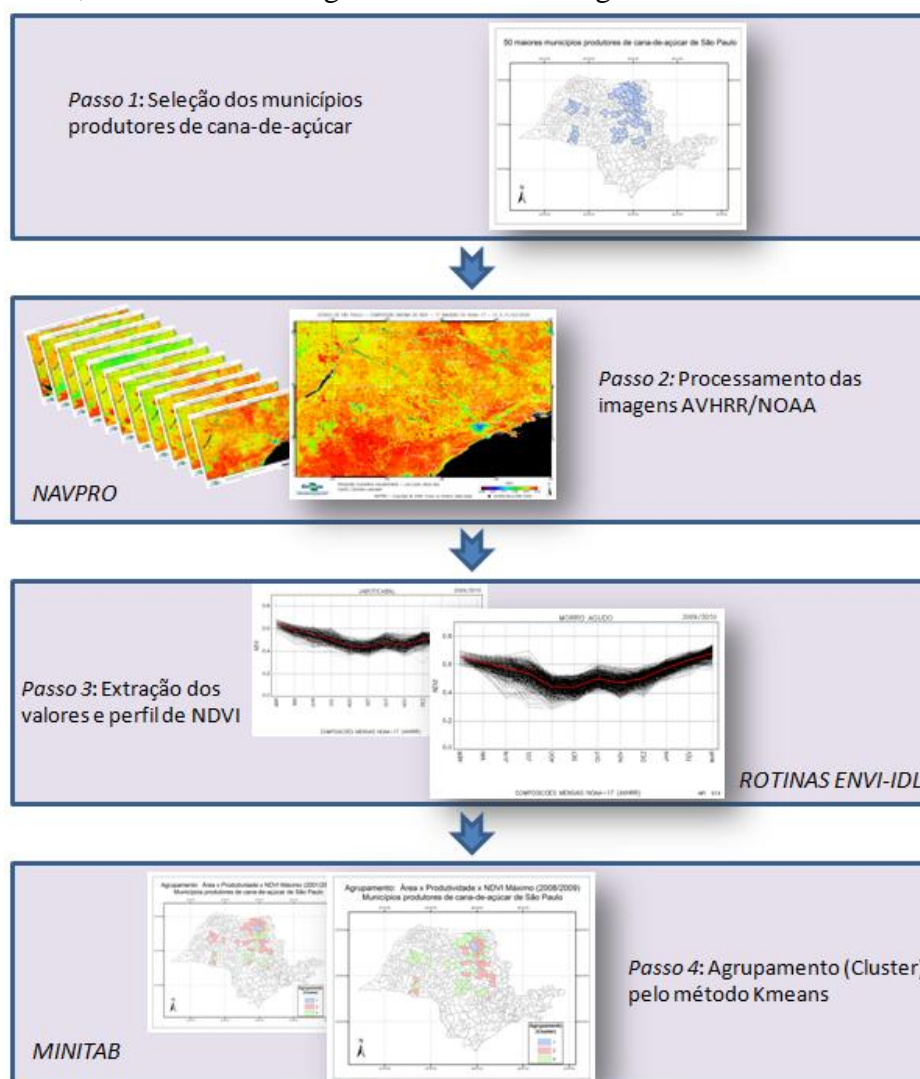


Figura 2. Fluxograma indicando os módulos do sistema desenvolvido neste trabalho.

Para a execução deste trabalho, utilizou-se o banco de imagens do satélite AVHRR/NOAA do Centro de Pesquisas Meteorológicas e Climáticas Aplicadas à Agricultura, da Universidade Estadual de Campinas (Cepagri/Unicamp). Foram utilizadas imagens dos satélites NOAA-16 e NOAA-17 de abril de 2001 a março de 2009. O processamento das imagens do AVHRR/NOAA foi realizado pelo sistema automático de navegação NAVPRO. O NAVPRO é um sistema para processamento e geração automática de produtos das imagens AVHRR/NOAA, desenvolvido por pesquisadores do *Colorado Center for Astrodynamics Research (CCAR), Aerospace Engineering Sciences*, da Universidade do Colorado (Boulder/EUA) (Esquerdo et al. 2006). O sistema realiza todo o processamento das imagens de forma automática, fazendo a conversão de formato, calibração radiométrica, correção geométrica (georreferenciamento preciso) e geração de produtos, como o NDVI.

O cálculo do NDVI foi feito a partir de imagens diárias do AVHRR/NOAA já pré-processadas, isto é, imagens com a calibração radiométrica, georreferenciamento preciso e recorte de São Paulo. O NDVI foi obtido pelas bandas 1 e 2, que correspondem ao vermelho e infravermelho próximo, respectivamente. Nesse processo, foram excluídos os pixels com

ângulo zenital solar maior que  $70^\circ$  e ângulos de varredura maiores que  $42^\circ$ . Esses pixels receberam o valor -1,1. Para amenizar o efeito da atmosfera nas imagens, geraram-se Composições de Valor Máximo (MVC) de NDVI.

Para gerar os perfis temporais com valores de NDVI das imagens MVC utilizou-se uma rotina em IDL desenvolvida por Esquerdo (2007). Para executar essa rotina foi necessário obter um conjunto de coordenadas (Latitude/Longitude) das áreas de cana-de-açúcar de cada município selecionado, que foram fornecidas pelo projeto Canasat/Inpe (<http://www.canasat.inpe.br>). Foram obtidos também os vetores e uma série temporal de imagens AVHRR/NOAA (2001 a 2009) desses municípios. Foram feitas oito séries temporais de imagens, uma para cada ano-safra (de abril a março). Como resultado, essa rotina gerou uma planilha e um gráfico com todos os valores de NDVI dos pixels analisados para cada município em cada ano-safra.

A última etapa do trabalho foi o agrupamento dos municípios. Nessa etapa foram utilizados dados de produção, produtividade e área plantada (que é igual a área colhida) de cana-de-açúcar obtidos pelo IBGE (<http://www.ibge.gov.br>). Para isso foi avaliado um dendrograma com as variáveis anuais: área plantada, produção e produtividade, e os valores de NDVI (máximo, médio, soma e mínimo) para eliminar as variáveis correlacionadas reduzindo o tamanho do conjunto a ser analisado. Em seguida, foi feito o agrupamento pelo método *K-means* (Han & Kamber, 2001), utilizando o software Minitab 14 (<http://www.minitab.com>), esse método classifica as observações em grupos, quando os grupos são desconhecidos inicialmente. O método começa com um agrupamento de observações em um número pré-definido de classes (clusters), neste estudo foram três classes definidas empiricamente. Cada observação é avaliada e movida para o grupo, cujo centróide está mais próximo, utilizando a distância euclidiana. O resultado do agrupamento foi inserido no ARCGIS (<http://www.esri.com/software/arcgis/index.html>) para a visualização da distribuição espacial dos municípios de cada classe em todos os anos-safra. Os resultados são apresentados na próxima seção.

### 3. Resultados e Discussão

O resultado do dendrograma, ilustrado na Figura 3, mostrou que as variáveis área plantada e produção de cana-de-açúcar são similares (98%), o valor médio e a soma de NDVI também são similares (100%), e a produtividade e o NDVI máximo apresentaram 58% de similaridade. O dendrograma também mostrou que as variáveis área plantada, produtividade e NDVI máximo não são similares e, por isso, foram as variáveis selecionadas para a execução dos experimentos.

Utilizando-se o agrupamento com três classes, o resultado foi equivalente para as variáveis área plantada, produtividade e NDVI máximo e para as variáveis área plantada, produtividade e soma de NDVI. Dessa forma, são apresentados apenas os resultados para o agrupamento das variáveis: área plantada, produtividade e NDVI máximo.

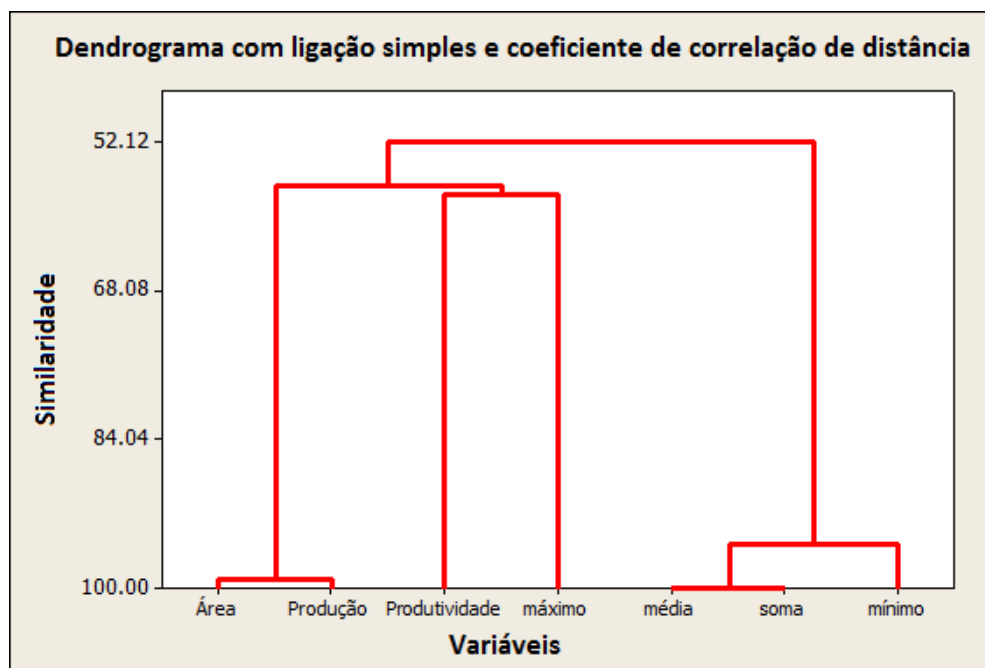


Figura 3. Dendrograma com ligação simples e coeficiente de correlação de distância das variáveis.

A Figura 4 ilustra o agrupamento das variáveis, área plantada, produtividade e NDVI máximo para os 50 municípios maiores produtores de cana-de-açúcar nas safras 2001/2002 a 2008/2009. Esses mapas (A a H) mostram que foi possível selecionar municípios similares ao longo de uma série temporal de imagens do AVHRR/NOAA, com baixa resolução espacial, e com dados de área plantada e produtividade de cana-de-açúcar do IBGE.

O município de Morro Agudo apresenta um padrão diferente de todos os outros, pois sozinho constitui um único agrupamento em toda a série temporal. O município apresenta área plantada diferente dos demais, isto é, apresenta a maior área plantada do estado de São Paulo. Isso fica evidente no agrupamento em todos os anos-safra. Os demais municípios variam a classe do agrupamento de ano para ano, em função da área plantada, produtividade e NDVI de cada safra.

A Tabela 1 apresenta os valores dos centróides dos agrupamentos (classes) 1, 2, 3 e do centróide da análise das variáveis área plantada, produtividade e NDVI máximo. O agrupamento 1 (classe 1) apresentou maior área plantada, representado somente pelo município de Morro Agudo. O agrupamento 2 (classe 2) apresentou os municípios com as menores áreas e a classe 3, os demais municípios com áreas maiores do que as da classe 2 e menores do que Morro Agudo. Pode-se observar na Tabela 1 que a área plantada tem aumentado ao longo dos últimos oito anos em todos os agrupamentos, embora a produtividade tenha se mantido sem alterações significativas.

O aumento da área plantada também pode ser acompanhado observando os mapas de agrupamento (de A a H) por ano na Figura 4. No mapa A, a maioria dos municípios estão na classe 2 (vermelho) que tem área plantada menor. Já, o mapa H possui mais municípios na classe 3 (verde) com área plantada maior. O ano safra 2005/2006 possui um padrão bastante diferente dos demais anos, pois é o único em que a classe 2 possui menos municípios do que a classe 3 (Tabela 2). Possivelmente, as condições climáticas tenham influência nessa alteração de padrão.

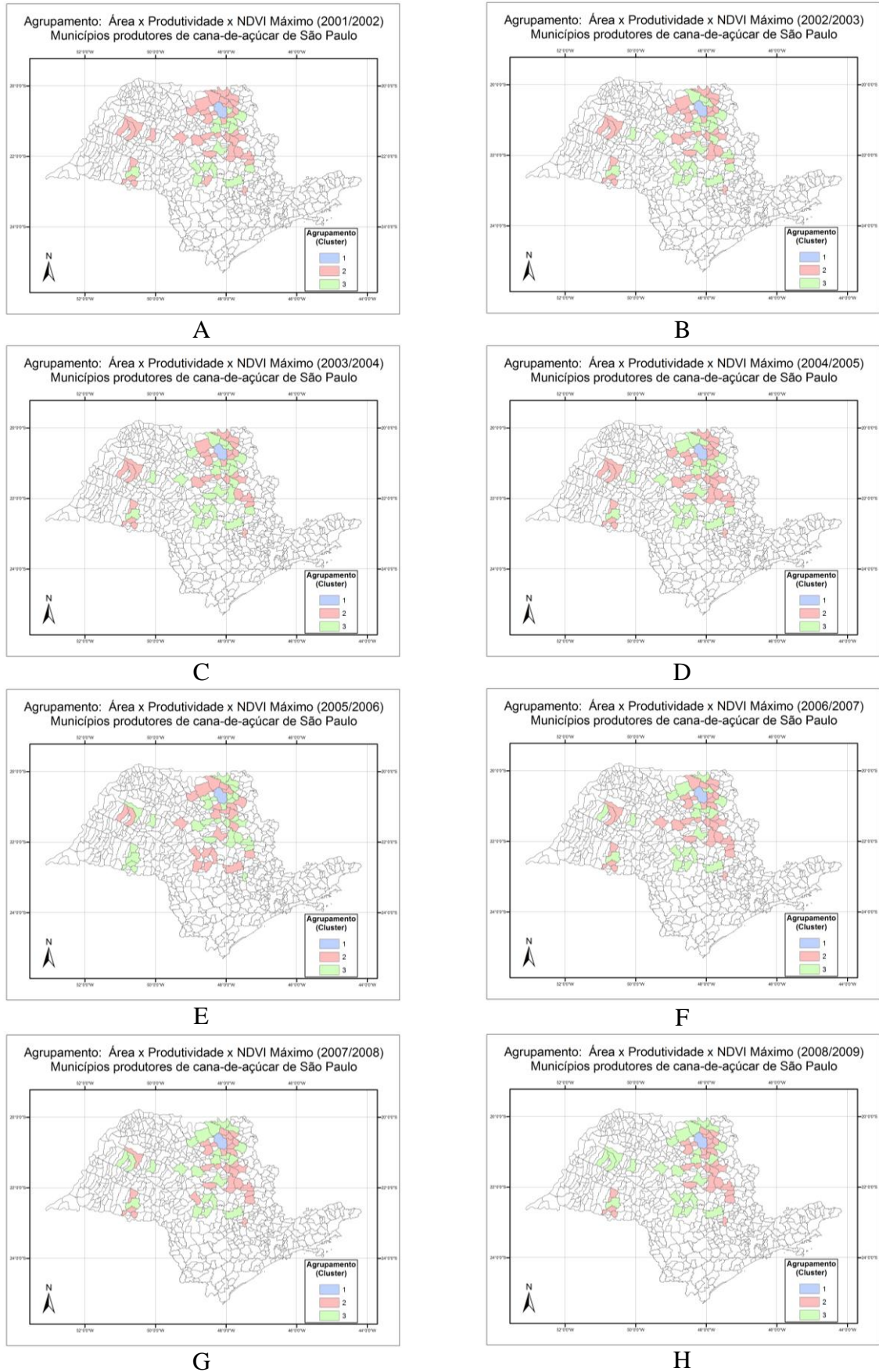


Figura 4. Agrupamento das variáveis, área plantada, produtividade e NDVI máximo para os 50 municípios produtores de cana-de-açúcar nas safras 2001/2002 a 2008/2009.

Tabela 1. Valores dos centróides dos agrupamentos (classes) 1, 2, 3 e do centróide da análise das variáveis área plantada (ha), produtividade (%) e NDVI máximo.

Safra	Variável	Classe 1	Classe 2	Classe 3	Centróide
2001/2002	Área plantada	78.000	18.432	33.629	23.669
	Produtividade	80,00	81,10	75,30	79,52
	NDVI Máximo	0,574	0,564	0,573	0,567
2002/2003	Área plantada	60.000	17.265	30.554	23.709
	Produtividade	100,00	83,23	79,19	81,84
	NDVI Máximo	0,662	0,634	0,630	0,633
2003/2004	Área plantada	93.000	17.875	29.713	24.783
	Produtividade	90,00	81,43	80,64	81,23
	NDVI Máximo	0,614	0,613	0,610	0,612
2004/2005	Área plantada	90.000	19.577	31.874	25.661
	Produtividade	88,00	81,13	81,96	81,58
	NDVI Máximo	0,642	0,623	0,616	0,621
2005/2006	Área plantada	86.602	18.257	30.957	25.189
	Produtividade	90,47	85,06	84,09	84,74
	NDVI Máximo	0,616	0,611	0,610	0,611
2006/2007	Área plantada	93.000	21.191	36.201	28.056
	Produtividade	82,00	84,40	81,53	83,30
	NDVI Máximo	0,634	0,620	0,612	0,617
2007/2008	Área plantada	93.000	22.152	39.621	30.905
	Produtividade	82,00	84,47	83,56	84,04
	NDVI Máximo	0,627	0,628	0,625	0,626
2008/2009	Área plantada	114.000	23.190	42.523	33.488
	Produtividade	90,00	85,62	85,50	85,66
	NDVI Máximo	0,670	0,633	0,620	0,628

De acordo com os valores da Tabela 2, o número de municípios na classe 2 tem uma tendência de queda, embora em alguns anos este número tenha aumentado, como é o caso dos anos de 2004/2005 e 2006/2007. A classe 3, ao contrário, mostrou um aumento no número de municípios a cada ano, o que pode estar relacionado com a expansão das áreas produtoras de cana-de-açúcar em cada município.

Tabela 2. Número de municípios produtores em cada classe do agrupamento.

Classe	2001/2002	2002/2003	2003/2004	2004/2005	2005/2006	2006/2007	2007/2008	2008/2009
1	1	1	1	1	1	1	1	1
2	36	29	27	30	23	31	29	27
3	14	21	23	20	27	19	21	23

Os municípios analisados estão localizados em áreas de alta e média aptidão agrícola, de acordo com o tipo de solo, atualmente utilizadas para agricultura de acordo com o zoneamento agroecológico da cana-de-açúcar, como pode ser visto em [http://www.cnps.embrapa.br/zoneamento\\_cana\\_de\\_acucar/SP\\_ZAE\\_Cana.pdf](http://www.cnps.embrapa.br/zoneamento_cana_de_acucar/SP_ZAE_Cana.pdf). Segundo Manzatto et al. (2009), o zoneamento agrícola fornece o potencial de terras aptas ao plantio da cultura, que se conjugado com as áreas atualmente sob plantio, fornece o potencial de expansão da cultura por regiões de interesse. Isso explica a expansão da área plantada de cana-de-açúcar em determinados municípios e os diferentes municípios em cada classe na série temporal. Vale ressaltar que mesmo usando um satélite de baixa resolução como o

AVHRR/NOAA, foi possível perceber uma expansão da cultura em todo o estado de São Paulo.

#### 4. Conclusão

Neste trabalho foi apresentada uma análise de agrupamento das variáveis área plantada, produtividade e NDVI máximo dos municípios produtores de cana-de-açúcar do estado de São Paulo. Foram definidas três classes de observações como entrada para o método K-means que foi executado para os oito anos-safra utilizados.

Os resultados mostraram que há uma variação na quantidade de municípios por classe para cada ano-safra. O ano-safra de 2001/2002 apresentou 36 municípios na classe 2 (vermelho) e 14 na classe 3 (verde). O último ano-safra analisado, de 2008/2009, tem um aumento de municípios na classe 3 (23) e uma diminuição na classe 2 (27) mostrando claramente um aumento do número de municípios com áreas plantadas maiores levando a um aumento na produção de cana-de-açúcar no estado de São Paulo.

O uso do NDVI gerado a partir do satélite AVHRR/NOAA conjuntamente com as variáveis área plantada e produtividade permitiu acompanhar a distribuição espacial dos municípios produtores de cana-de-açúcar ao longo dos últimos oito anos-safra. Embora o AVHRR/NOAA seja um satélite de baixa resolução espacial (1km x 1km), seus produtos podem auxiliar no estudo de culturas agrícolas produzidas em grandes extensões e em campos contíguos como é o caso da cana-de-açúcar em São Paulo.

Como trabalho futuro, pretende-se avaliar as razões pelas quais o ano de 2005/2006 teve um comportamento atípico se comparado aos demais anos-safra. Além disso, novos algoritmos de agrupamento que não necessitam de entrada inicial do número de classes deverão ser testados.

#### Agradecimentos

Às agências CNPq, Fapesp, Capes, Microsoft Research e Embrapa pelo apoio financeiro, ao Cepagri/Unicamp pelas imagens de satélite AVHRR/NOAA e ao projeto Canasat do INPE pelos vetores de cana-de-açúcar.

#### Referências Bibliográficas

- Esquerdo, J. C. D. M.; Antunes, J. F. G.; Baldwin, D. G.; Emery, W. J.; Zullo Jr, J. An automatic system for AVHRR land surface product generation. **International Journal of Remote Sensing**, v.27, p.3925-3942, 2006.
- Esquerdo, J. C. D. M. Utilização de perfis multi-temporais do NDVI/AVHRR no acompanhamento da safra de soja no oeste do Paraná. 2007. 168p. Tese (Doutorado em Engenharia Agrícola) – Faculdade de Engenharia Agrícola, Universidade Estadual de Campinas, Campinas. 2007.
- Gonçalves, R. R. V.; Nascimento, C. R.; Zullo Jr., J.; Romani, L. A. S. Relationship between the spectral response of sugar cane, based on AVHRR/NOAA satellite images, and the climate condition, in the state of sao paulo (Brazil), from 2001 to 2008. In: International Workshop on the Analysis of Multi-temporal Remote Sensing images (MultiTemp), 5, 2009, Groton. **Proceedings...** 2009, p. 315–322.
- Han, J.; Kamber, M. **Data Mining - Concepts and Techniques**. New York: Morgan Kaufmann Publishers, 2001, 550p.
- IPCC. Climate change 2007: the physical science basis: summary for policymakers. Geneva: IPCC, 2007. 18 p. Disponível em: <<http://www.ipcc.ch/SPM2feb07.pdf>>. Acesso em: 3 ago. 2010.
- Manzatto, C. V.; Assad, E. D.; Bacca, J. F. M.; Zaroni, M. J.; Pereira, S. E. M. **Zoneamento agroecológico da cana-de-açúcar**. Rio de Janeiro: Embrapa Solos, 2009. 55 p. (Documentos / Embrapa Solos, ISSN 1517-2627; 110).
- Nascimento, C. R.; Gonçalves, R. R. V.; Zullo Jr., J.; Romani, L. A. S. Estimation of sugar cane productivity using a time series of AVHRR/NOAA-17 images and a phenology-spectral model. In: International Workshop on the Analysis of Multi-temporal Remote Sensing images (MultiTemp), 5, 2009, Groton. **Proceedings...** 2009, p. 365–372.
- Romani, L. A. S.; Gonçalves, R. R. V.; Zullo Jr., J.; Traina Jr., C.; Traina, A. J. M. New DTW-Based method to similarity search in sugar cane regions represented by climate and remote sensing time series. In: International Geoscience and Remote Sensing Symposium (IGARSS), 2010, Honolulu. **Proceedings...** 2010. p. 355-358.