

Comparação de classificadores para o mapeamento de culturas agrícolas anuais em Campo Novo do Parecis – MT, utilizando NDVI/MODIS.

Victor Danilo Manabe¹

Jansle Vieira Rocha¹

¹Univerdidade Estadual de Campinas – UNICAMP/FEAGRI
Av. Candido Rondon, 501 – 13083-874 – Campinas – SP, Brasil
{victor.manabe, jansle.rocha}@feagri.unicamp.br

Abstract. This study aims to evaluate the performance of four classifiers (MaxVer, Decision Tree, SVM and Neural Networks) in mapping areas of annual crops in Campo Novo do Parecis - Mato Grosso, using NDVI/MODIS time series data for 2010/11 season. Parameters used were selected from the time series of 23 images of the MODIS/NDVI. The classifiers MaxVer, Decision Tree, SVM and Neural Networks showed very similar values of accuracy, 0.927, 0.971, 0.930 and 0.927, and Kappa, 0.824, 0.805, 0.829 and 0.825, respectively. The SVM and MaxVer classifiers presented the characteristics most balanced among the classes of annual crops and other uses. Decision Tree is a classifier more aggressive in areas of annual crops and SVM more conservative for these areas.

Palavras-chave: Vegetation Index, Classification, Data Mining, Índice de Vegetação, Classificação, Mineração de Dados.

1. Introdução

O agronegócio corresponde a aproximadamente a 25% do PIB nacional e apresentou em 2011 um crescimento de 5,73%, tendo o estado do Mato Grosso papel fundamental no crescimento, sendo o cultivo de grãos o destaque no agronegócio brasileiro (CEPEA, 2012). Desta maneira, o monitoramento e a estimativa das áreas plantadas de agricultura anual (soja, milho, algodão e trigo) são de extrema importância.

A utilização de séries temporais de índices de vegetação vem sendo bastante utilizado para a identificação e mapeamento de culturas agrícolas (Chang et al., 2007; Wardlow et al., 2007; Victoria et al., 2009). As imagens de NDVI (Índice de Vegetação da Diferença Normalizada) do sensor MODIS apresentam larga utilização, tendo o produto MOD13Q1 imagens a partir de composições de 16 dias com resolução de 250 m (MODIS, 2010). A obtenção de diversas imagens para um ano agrícola, possibilita a identificação de períodos onde é possível a realização da diferenciação de uso do solo na serie temporal (Wardlow et al., 2007), assim como a criação de parâmetros a partir da série temporal que irão ajudar na classificação (Arvor et al., 2012).

Existem diversos métodos de classificação de imagens estabelecidos para sensoriamento remoto, como o classificador MaxVer (aplicado por Arvor et al., 2011). Porém técnicas de mineração de dados estão sendo cada vez mais aplicadas para esta finalidade. Segundo Fayyad et al. (1996) mineração de dados é o processo não trivial de identificação de novos padrões em um conjunto de dados. Desta forma, diversos trabalhos foram feitos com aplicações de técnicas de mineração de dados em classificação de imagens (Foody e Mathur, 2006, Peña-Barragán et al., 2011, Panda et al., 2010) , dentre aquelas mais utilizadas e que apresentam melhores resultados estão Arvore de Decisão, SVM (Suport Vector Machine) e Redes Neurais.

Desta forma, este trabalho tem como objetivo avaliação do desempenho de quatro classificadores (MaxVer, Arvore de Decisão, SVM e Redes Neurais) no mapeamento de áreas de cultura agrícola anual no município de Campo Novo do Parecis – Mato Grosso, utilizando dados da série temporal NDVI/MODIS para a safra 2010/11.

2. Metodologia de Trabalho

A área de estudo é o municio de Campo Novo do Parecis, Figura 1, um dos principais produtores de grãos do estado do Mato Grosso, tendo também áreas de plantio de cana de açúcar e pastagem, assim como vegetação natural.

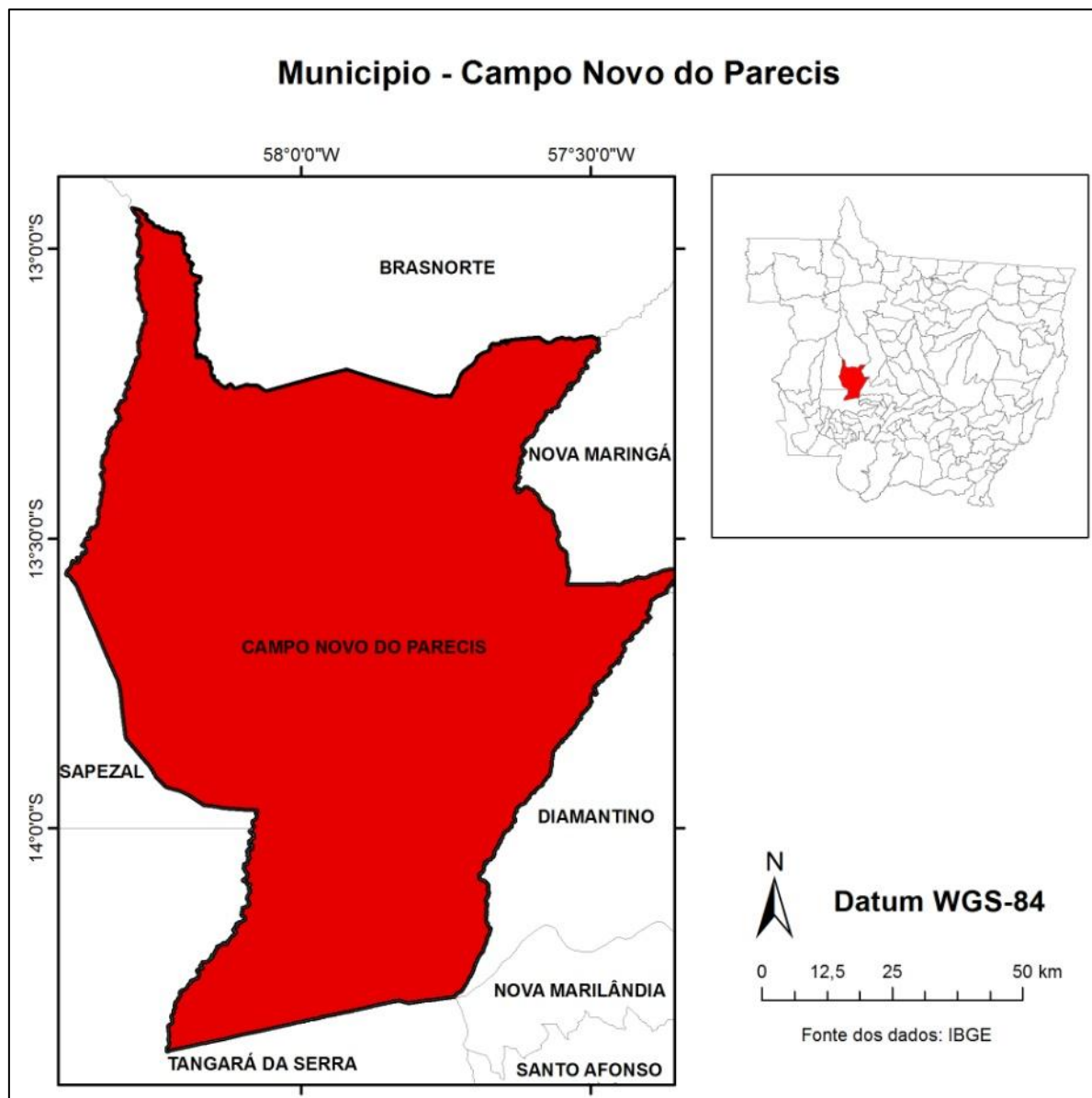


Figura 1: Localização do município de Campo Novo de Parecis - MT.

Foram utilizadas 23 imagens, no período de 12 de julho de 2010 até 26 de Junho de 2011, de NDVI do produto MOD13Q1 do sensor MODIS para o estado do Mato Grosso, que é disponibilizado na Base Estudual Brasileira pela Embrapa (<http://www.modis.cnptia.embrapa.br>) (Esquerdo, 2011). A partir desta serie de imagens foram extraídas informações em relação a serie temporal de NDVI, estes dados foram trabalhados em conjunto com as imagens de NDVI, Quadro 1.

Quadro 1: Dados extraídos para seleção de atributos.

NDVI mínimo da serie	NDVI máximo da serie
Diferença entre NDVI mínimo e máximo	Média de NDVI da série
Inclinação máxima da serie	Data do valor de NDVI maximo
Data do valor de NDVI mínimo	NDVI dia 12/07/10
NDVI dia 28/07/10	NDVI dia 13/08/10
NDVI dia 29/08/10	NDVI dia 14/09/10
NDVI dia 30/09/10	NDVI dia 16/10/10
NDVI dia 01/11/10	NDVI dia 17/11/10
NDVI dia 03/12/10	NDVI dia 19/12/10
NDVI dia 01/01/11	NDVI dia 17/01/11
NDVI dia 02/02/11	NDVI dia 18/02/11
NDVI dia 06/03/11	NDVI dia 22/03/11
NDVI dia 07/04/11	NDVI dia 23/04/11
NDVI dia 09/05/11	NDVI dia 25/05/11
NDVI dia 10/06/11	NDVI dia 26/06/11

Foi realizada uma técnica de seleção de atributos, para utilização dos dados que mais possuem correlação, o que permite maior rapidez, assim como melhoria na acurácia do classificador. Foi utilizada a técnica CFS (Seleção Baseada em Correlação), onde os atributos são avaliados na sua correlação com a classe e na sua não correlação entre si. Assim os atributos que foram utilizados para a classificação se apresentam no Quadro 2.

Quadro 2: Dados utilizados para as classificações de cultura anual.

NDVI minimo da serie	NDVI máximo da serie
Diferença entre NDVI mínimo e máximo	NDVI dia 12/07/10
NDVI dia 01/01/11	NDVI dia 13/08/10
NDVI dia 14/09/10	NDVI dia 16/10/10
NDVI dia 22/03/11	

Foram feitas duas amostras de pixels para treinamento uma contendo áreas de cultura anual e outra com todas outras classes presentes na área de estudo. Através do software de mineração de dados WEKA 3.6 foram realizados testes para a seleção dos parâmetros utilizados nos classificadores SVM e Redes Neurais e a geração da Arvore de Decisão, utilizando o algoritmo J48. No software ENVI 4.7 foram geradas as mascaras de agricultura para os quatro classificadores (MaxVer, Arvore de Decisão, SVM e Redes Neurais) utilizando as amostra de treinamento e os parâmetros definidos.

Para a análise das classificações foram gerados 313 pontos aleatórios em todo o município de Campo Novo do Parecis. Os pontos foram analisados com auxílios de imagens Landsat 5, e apresentam a seguinte referencias de distribuição 98 pontos de agricultura anual, 182 pontos de vegetação, 12 pontos de cana de açúcar e 21 pontos de pasto.

3. Resultados e Discussão

Realizou-se a classificação do município de Campo Novo do Parecis para os quatro classificadores testados, MaxVer, Arvore de Decisão, SVM e Redes Neurais, as mascaras de agricultura anual se encontram na Figura 1.

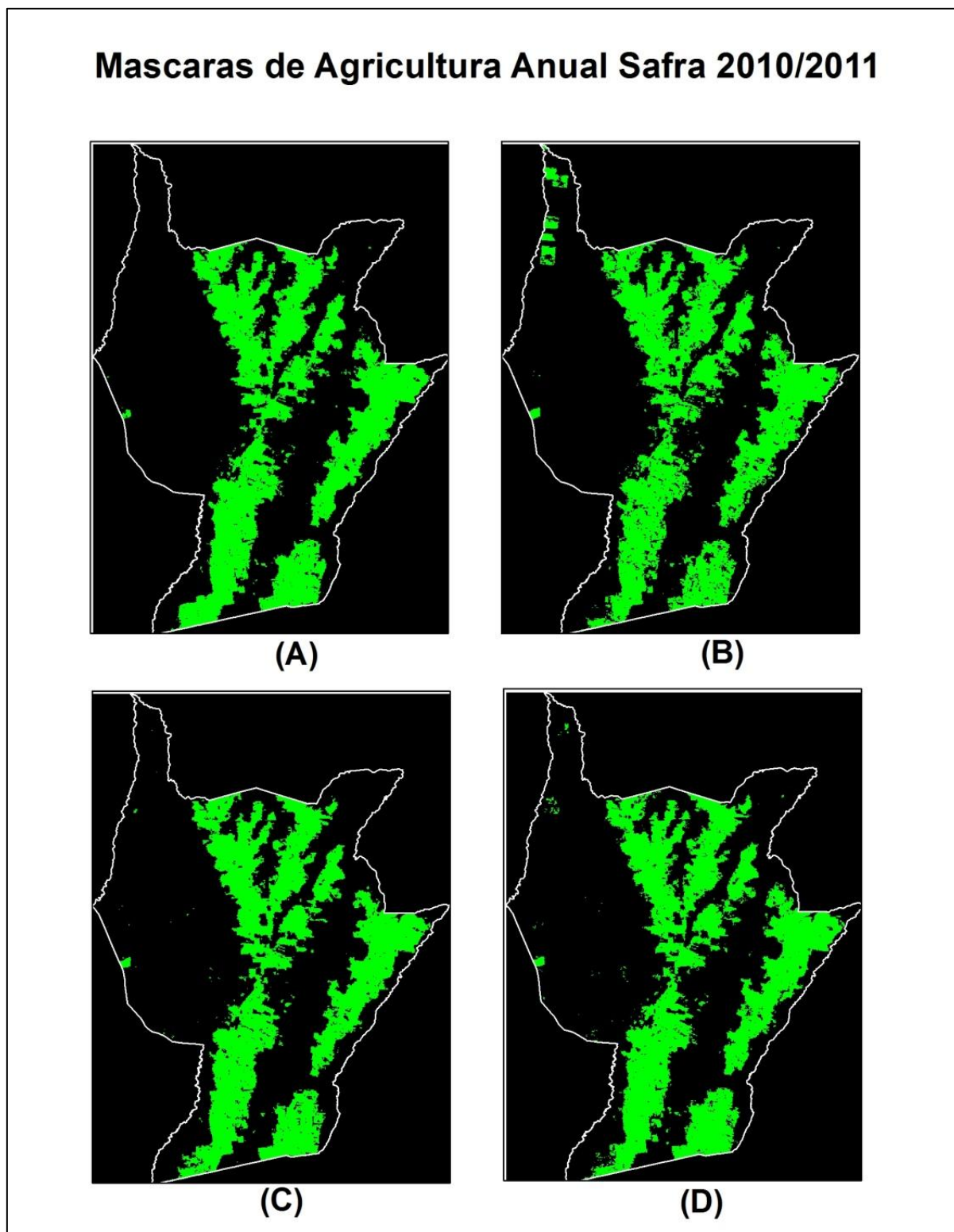


Figura 1: Mascaras de cultura anual para Campo Novo do Parecis. (A) MaxVer; (B) Arvore de Decisão; (C) SVM; (D) Redes Neurais.

Ao analisar os pontos aleatórios gerados pode-se notar que os classificadores de uma maneira geral apresentaram resultados muito próximos, Tabela 1. MaxVer, SVM e Redes Neurais apresentaram valores bastante semelhantes de acurácia, 0,927, 0,930 e 0,927, e de índice Kappa, 0,824, 0,829 e 0,825, respectivamente. Arvore de Decisão apresentou acurácia

de 0,917 e índice Kappa de 0,805, valores relativamente inferiores ao dos outros classificadores.

Tabela 1: Matriz de confusão dos classificadores.

	MaxVer			Arv.Decisão			SVM			Redes Neurais		
	Agric	Outros		Agric	Outros		Agric	Outros		Agric	Outros	
Agricultura	81	17	98	83	15	98	79	19	98	82	16	98
Outros	6	209	215	11	204	215	3	212	215	7	208	215
	87	226	313	94	219	313	82	231	313	89	224	313
Acuracia Geral	0,927			0,917			0,930			0,927		
Indice Kappa	0,824			0,805			0,829			0,825		

Por outro lado a classificação da Arvore de Decisão apresentou uma especificidade maior para as áreas de cultura anual, 0,847, ou seja, possui um erro de omissão de 0,153, menor em comparação aos outros classificadores, Tabela 2. Porém seu erro de inclusão, 0,117, é consideravelmente maior, o que leva a compreensão de que seu maior acerto das áreas de agricultura anual é em decorrência de uma maior generalização desta classificação.

Tabela 2: Analise resultado dos classificadores

	MaxVer	Arv.Decisão	SVM	Redes Neurais
Erro Omissão Agric	0,173	0,153	0,194	0,163
Erro Inclusão Agric	0,069	0,117	0,037	0,079
Especificidade Agric	0,827	0,847	0,806	0,837

O classificador SVM, neste caso, apresentou a melhor acurácia geral, entretanto, é aquele com a pior especificidade para cultura anual, 0,806, e com o menor erro de inclusão 0,037, Tabela 2. Isto ocorre, porque ele é um classificador conservador, ou seja, apresenta um erro de omissão da classe agricultura anual muito maior do que o de inclusão.

MaxVer e Redes Neurais se comportaram de maneira muito similares em relação aos acertos da classe agricultura anual, obtiveram respectivamente, erro de omissão de 0,173 e 0,163 e erro de inclusão de 0,069 e 0,079, Tabela 2. Apesar de um erro de omissão consideravelmente maior que aquele de inclusão, na comparação com as classificações da Arvore de Decisão e de SVM, apresentam características de classificadores intermediários, pois equilibram melhor o acerto geral da classificação, com os erros de inclusões e omissões das classes, sendo os dois classificadores mais balanceados dentre os trabalhados.

As áreas de cana de açúcar foram as que causaram maior dificuldade para diferenciação de agricultura anual, sendo aquela que percentualmente obteve maior quantidade de erro por inclusão na maioria das classificações, Tabela 3.

Tabela 3: Erros de inclusão.

	Agricultura			
	Max_ver	Arvore	Redes	SVM
Vegetação	0,011	0,074	0,037	0,011
Cana	0,046	0,043	0,037	0,022
Pasto	0,011	0,000	0,012	0,000

4. Conclusões

A classificação de áreas de agricultura anual utilizando informações provenientes de séries temporais de NDVI utilizando as técnicas MaxVer, Arvore de Decisão, SVM e Redes Neurais, apresentaram bons valores de acurácia e índice Kappa, assim como, valores relativamente próximos.

Os classificadores MaxVer e SVM foram os que apresentaram características mais equilibradas entre os acertos das classes de cultura anual e outros usos. Sendo Arvore de Decisão um classificador mais agressivo para áreas de agricultura anual e o SVM mais conservador para estas áreas.

Agradecimentos

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela concessão de bolsa de mestrado.

Referências Bibliográficas

Arvor, D., Jonathan, M., Meinelles, M. S. P., Dubreuil, V., & Durieux, L. Classification of MODIS EVI time series for crop mapping in the state of Mato Grosso. Brazil: **International Journal of Remote Sensing**. vol. 32, issue 22, pp. 7847-7871, 2011

CEPEA. **PIB do agronegócio 2011** <http://www.cepea.esalq.usp.br/comunicacao/Cepea_PIB_BR_dez_2011.pdf>, acesso em 08/11/2012

Chang, J.; Hansen, M.C.; Pittman, K.; Carrol, M.; Dimiceli, C. Corn and Soybean Mapping in the United States Using MODIS Time-Series Data Sets. **Agron. J.** 99, p. 1654–1664, 2007

Esquerdo, J. C. D. M. ; Antunes, J. F. G. ; Andrade, J. C. . Desenvolvimento do Banco de Produtos MODIS na Base Estadual Brasileira. In: XV Simpósio Brasileiro de Sensoriamento Remoto, 2011, Curitiba-PR. **Anais...** São José dos Campos: INPE, 2011. p. 7596-7602.

Fayyad, Usama; Piatetski-Shapiro, Gregory; Smyth, Padhraic (1996) The KDD Process for Extracting Useful Knowledge from Volumes of Data. In: **Communications of the ACM**, pp.27-34, Nov.1996

Foody, G. M., & Mathur, A. The use of small training sets containing mixed pixels for accurate hard image classification: Training on mixed spectral responses for classification by a SVM. **Remote Sensing of Environment**, 103, 179–189, 2006.

MODIS Vegetation Indices (MOD13) C5. **User's Guide**. Disponível em: <<http://tbrs.arizona.edu/project/MODIS/UsersGuide.pdf>>. Acesso em: 08/11/2012.

Panda, Sudhanshu Sekhar; Ames, Daniel P.; Panigrahi, Suranjan Application of Vegetation Indices for Agricultural Crop Yield Prediction Using Neural Network Techniques. **Remote Sensing**, vol. 2, issue 3, pp. 673-696 2010

José M. Peña-Barragán, Moffatt K. Ngugi, Richard E. Plant, Johan Six Object-based crop identification using multiple vegetation indices, textural features and crop phenology. **Remote Sensing of Environment**, Volume 115, Issue 6, 15 June 2011, Pages 1301–1316

Victoria, D.C.; Oliveira, A.F.; Greco, C.R. Análise harmônica de séries temporais de imagens NDVI/MODIS para discriminação de coberturas vegetais. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 14. Natal. **Anais**. São José dos Campos: INPE, p. 1589-1596. 2009.

Wardlow, B.D.; Egbert, S.L.; Kastens, J.H. Analysis of time-series MODIS 250 m vegetation index data for crop classification in the U.S. Central Great Plains. **Remote Sensing of Environment** 108, p. 290-310, 2007.