

Identificação da mudança de uso da terra usando técnicas de agrupamento de séries temporais de imagens de satélite

Rachel Scrivani¹
Bruno Ferraz do Amaral²
Renata Ribeiro do Valle Gonçalves³
Elaine Parros Machado de Sousa²
Jurandir Zullo Jr.^{1,3}
Luciana Alvim Santos Romani⁴

¹ Universidade Estadual de Campinas
Faculdade de Engenharia Agrícola – Feagri
Av. Cândido Rondon, 501 - Barão Geraldo
13083-875 - Campinas - SP, Brasil
rachelscrivani@gmail.com

² Universidade Estadual de São Paulo
Avenida Trabalhador São-Carlense, 400 - Centro
CEP: 13566-590 - São Carlos - SP, Brasil
{amaral, parros}@icmc.usp.br

³ Universidade Estadual de Campinas
Centro de Pesquisas Meteorológicas e Climáticas Aplicadas a Agricultura – Cepagri
Cidade Universitária "Zeferino Vaz"
13083-970 - - Campinas – SP, Brasil
{renata, jurandir}@cpa.unicamp.br

⁴ Embrapa Informática Agropecuária
Av. André Toselo, 209 - Caixa Postal 6041
13083-886 - Campinas - SP, Brasil
luciana.romani@embrapa.br

Resumo. A disponibilidade de dados orbitais, aliada à necessidade crescente de monitoramento de grandes extensões e de preservação de regiões ambientalmente sensíveis, gera uma oportunidade para o desenvolvimento/adaptação de métodos computacionais. A fim de gerar informações de avaliação temporal e espacial que possam constituir importante ferramenta de planejamento e de orientação à tomada de decisão para o manejo e conservação destas áreas, o objetivo deste trabalho é propor uma abordagem baseada em métodos de agrupamento de séries de imagens de satélite para auxiliar na análise espaço-temporal da mudança do uso da terra, identificando de forma automática, com alta probabilidade de serem áreas inundáveis, floresta, culturas agrícolas e pastagem em escala regional. Séries temporais de imagens do sensor MODIS com valores dos índices de vegetação NDVI e EVI, do período de 2008/2009 a 2013/2014, foram agrupadas por meio do algoritmo K-means. Para uma avaliação da qualidade dos agrupamentos obtidos foi utilizado o coeficiente de Silhueta. Séries temporais de precipitação do satélite TRMM foram utilizadas a fim de correlacionar com os demais resultados obtidos para avaliar a pré-classificação gerada pelo método de agrupamento em anos secos, chuvosos e com distribuição de chuva diferente da normalidade. A abordagem proposta pode auxiliar no monitoramento de áreas ambientalmente sensíveis, tendo o EVI apresentado melhor desempenho em áreas de densa vegetação e concentração de água.

Palavras-chave: sensoriamento remoto, MODIS, séries temporais, índice de vegetação.

Abstract. The wide availability of orbital data, combined with the growing need for monitoring large areas as well as the preservation of environmentally sensitive areas, generates an opportunity for the development/adaptation of computational methods. In order to generate information of temporal and spatial assessment that may be an important tool for planning and policy for decision making to the management and conservation of these areas, the goal of this work is to propose an approach based on cluster analysis of satellite image time series in order to aid in analysis of spatio-temporal change of land use, automatically identifying wetlands, forest, agricultural crops and pasture on a regional scale. Time series of MODIS images with values of NDVI and EVI vegetation indices, the period from 2008/2009 to 2013/2014 were grouped through the k-means algorithm. Silhouette coefficient was employed to verify the quality of generated clusters. Temporal data from the TRMM satellite rainfall series were used to correlate with the other results obtained from the pre-classification generated by the clustering method in dry and wet years with different rainfall distribution normality. The proposed approach can aid at applying to the monitoring of environmentally sensitive areas, and the EVI series showed to be better than NDVI ones when applied in areas of dense vegetation and water concentration.

Key-words: remote sensing, MODIS, time series, vegetation index.

1. Introdução

O Pantanal Brasileiro, uma planície contínua de inundação, está localizado na região Centro-Oeste do Brasil inserido na Bacia do Alto Paraguai com uma área aproximada de 150 mil km² nos estados do Mato Grosso (40,3%km²) e Mato Grosso do Sul (59,7%km²), apresentando exuberantes e diversificadas fauna e flora, Ibama (2014). É declarado Reserva da Biosfera e Patrimônio Mundial Natural pela Unesco (Organização das Nações Unidas para Educação, a Ciência e a Cultura), porém apenas 4,4% do Pantanal encontram-se protegidos por unidades de conservação, MMA (2014).

Uma fonte de preocupação atual com a conservação ambiental deste importante bioma brasileiro está na expansão da agricultura nas áreas que circundam o Pantanal, pois segundo Santos (2006) ela pode causar a intensificação da erosão laminar e do assoreamento, a contaminação com biocidas da rede hidrográfica e alterações limnológicas, a remoção de matas ciliares, dentre outros fatores, ameaçando em grande escala as características ecológicas e a biodiversidade da região.

No entanto, o monitoramento de uma grande extensão territorial pode beneficiar-se do uso de técnicas de sensoriamento remoto, que permite, por exemplo, a avaliação das características biofísicas da vegetação pelo uso dos índices de vegetação, Yamasoe (2004).

Na região pantaneira há vários estudos mostrando a aplicação do sensoriamento remoto para identificar padrões de vegetação ciliar, Silva et al. (2009), para classificação de pastagens degradadas Abdon et al. (2009), para discriminação de formações vegetais Victoria et al. (2009), para identificação de áreas alagadas, Santos et al. (2009), avaliações utilizando índices de vegetação, Viana e Alvalá (2011) dentre outros.

Nos últimos anos, a tecnologia utilizada no desenvolvimento dos sensores orbitais evoluiu provocando um aumento na quantidade e na qualidade das imagens captadas, o que aumenta significativamente o volume de dados disponíveis para análise. Com isso, cresce a necessidade de desenvolvimento de novos métodos e técnicas computacionais para auxiliar os especialistas na tarefa de análise e extração de conhecimento do conjunto de dados. Novos métodos computacionais para lidar com grandes volumes de dados estão sendo cada vez mais utilizados em sensoriamento remoto. No entanto, ainda há poucos trabalhos desenvolvidos para o bioma Pantanal.

A disponibilidade de dados orbitais, aliada à necessidade crescente de monitoramento de grandes extensões e de preservação da região, gera uma oportunidade para o desenvolvimento/

adaptação de métodos computacionais. Neste cenário, uma questão relevante é: dados de sensoriamento remoto classificados de forma não supervisionada podem ser úteis para identificação de áreas de expansão agrícola em regiões ambientalmente sensíveis? Existem alternativas para classificação de áreas para as quais não se tem imagens que representam a “verdade terrestre”? Como acompanhar de forma automática a dinâmica de mudança no uso da terra? Como forma de solucionar estas questões e gerar informações de avaliação temporal e espacial que possam constituir importante ferramenta de planejamento e de orientação à tomada de decisão para o manejo e conservação de áreas afins, este trabalho propõe o uso de algoritmo de agrupamento para analisar séries temporais de imagens de satélite.

Este artigo está organizado em 5 seções. A Seção 2 apresenta os objetivos do trabalho. A Seção 3 apresenta material e métodos. Seção 4 apresenta os resultados. Finalmente a Seção 5 apresenta as conclusões do trabalho.

2. Objetivo

Este trabalho propõe uma abordagem baseada em métodos de agrupamento de séries de imagens de satélite para auxiliar na análise espaço-temporal da mudança do uso da terra, identificando de forma automática, com alta probabilidade de serem áreas inundáveis, floresta, culturas agrícolas e pastagem em escala regional.

3. Material e Métodos

A área de estudo é o município de Poconé em Mato Grosso. Localizado nas coordenadas 16° 15' 25" S, 56° 37' 22" W inserido no Pantanal Brasileiro. Este município possui regiões de agendamento, florestas, pastagem e áreas agrícolas. As imagens utilizadas são do sensor MODIS a bordo do satélite Terra, produto MOD13Q1, com resolução de 16 dias, composição e resolução espacial de 250m, de 2008/2009 a 2013/2014, disponibilizadas pela Embrapa Informática Agropecuária. Foram utilizadas imagens com valores dos índices de vegetação NDVI e EVI para o estado de Mato Grosso.

O NDVI (do inglês, *Normalized Difference Vegetation Index*) proposto por Rouse et al. (1973) teve por objetivo produzir um Índice de Vegetação espectral que separa vegetação verde em relação às demais áreas e também é capaz de medir a quantidade e a condição da vegetação em uma determinada área, pois quanto maior o valor do índice, que varia de -1 a 1, maior é a quantidade de biomassa verde. É expresso como a diferença entre a banda do infravermelho próximo e vermelho normalizada pela soma das bandas, definido na seguinte equação:

$$NDVI = (\rho_{ivp} - \rho_v) / (\rho_{ivp} + \rho_v)$$

Onde: ρ_{ivp} é a refletância no infravermelho próximo; ρ_v é a refletância no vermelho.

O EVI (do inglês, *Enhanced Vegetation Index*) proposto por Huete et al. (1997) foi desenvolvido para melhorar a qualidade do NDVI. Seu cálculo é semelhante ao do NDVI, porém ele corrige algumas distorções na luz refletida causadas pelas partículas no ar, bem como a cobertura do solo por baixo da vegetação. Os dados de EVI não se tornam saturados com tanta facilidade como o NDVI ao imagear florestas tropicais e outras áreas da Terra com grande quantidade de clorofila. O índice é calculado a partir da seguinte equação:

$$EVI = 2,5[(\rho_{ivp} - \rho_v) / (\rho_{ivp} + 6,0\rho_v - 7,5\rho_a + 1,0)]$$

Onde: ρ_{ivp} é a refletância no infravermelho próximo; ρ_v é a refletância no vermelho; ρ_a é a refletância no azul.

As imagens de NDVI e EVI adquiridas tinham os valores dos índices em números inteiros multiplicados por 10.000, então a primeira etapa do trabalho foi a conversão desses valores para -1 a 1. A segunda etapa foi o uso do algoritmo K-means e distância Euclidiana para realizar a análise de agrupamento das séries temporais de NDVI e EVI no software SatImagExplorer¹, Chino et al. (2010). O K-means é um algoritmo de agrupamento baseado em particionamento, em que cada objeto do conjunto de dados a ser agrupado pode pertencer a apenas um agrupamento. Inicialmente são escolhidos aleatoriamente os elementos representantes (centróides) dos agrupamentos, e cada elemento do conjunto de dados é associado ao agrupamento cujo representante é o mais próximo. Utilizando o valor médio dos elementos de cada grupo, o algoritmo recalcula iterativamente os centróides, e redistribui todos os objetos, até que não haja mais alterações nos agrupamentos, Hans e Kamber (2006).

Para avaliar a qualidade do processo de agrupamento, foi utilizado o coeficiente de Silhueta (do inglês, *Silhouette coefficient*) de interpretação e validação de dados de agrupamento. É uma medida que combina coesão e separação. A coesão mede o quanto os objetos dentro de um grupo se aglomeram perto do centro do grupo. Enquanto, a separação mede o quanto os centros dos agrupamentos estão bem separados entre si, Rousseeuw (1987).

Para a realização dos experimentos foram processadas 23 imagens por ano para cada índice de vegetação, perfazendo um total de 138 imagens de cada índice de vegetação. Considerando os seis anos de estudo, o total geral foi de 276 imagens processadas. Foram extraídas 290.000 séries temporais para cada índice por ano, levando a um total de 580.000 séries. Cada série temporal corresponde a um pixel da imagem com resolução de 250 m cada um. Ao final foram analisadas 580.000 séries por ano, totalizando 3.480.000 séries submetidas como dado de entrada para o algoritmo de agrupamento, disponível na ferramenta SatImagExplorer. Valores de $k=5$, $k=10$, $k=15$ e $k=20$ foram utilizados, sendo k o número de grupos a formar.

A fim de correlacionar os resultados obtidos com a série temporal de precipitação pluviométrica foram utilizados dados obtidos do satélite TRMM (*Tropical Rainfall Measuring Mission*), produto 3B42, disponibilizados para este estudo pela Embrapa Informática Agropecuária. Os dados de precipitação foram organizados e são apresentados na forma de gráficos.

4. Resultados e Discussão

Dentre os valores de k testados, definiu-se por meio da avaliação do coeficiente de Silhueta, da avaliação realizada pelos especialistas e ainda tomando por base os tipos de vegetação presentes na região, que o número $k=10$ era adequado para identificar os principais tipos de cobertura vegetal para a área de estudo. A vegetação de Poconé se divide em quatro sistemas: Naturais, de Transição, Secundários e Artificializados. Nos sistemas Naturais estão as florestas, cerrado e savanas. No sistema de Transição está a floresta sempre inundável. Nos sistemas Secundários e Artificializados estão as áreas antropizadas, Almeida et.al (2010).

O clima em Poconé é sazonal, com temperatura média anual em torno de 25°C. A precipitação anual em média é de 1.200 mm e pelo menos 80% das chuvas estão concentradas no período de dezembro a março. Poconé tem grande parte de sua área inundável num ciclo anual de inundações que varia de intensidade em decorrência das chuvas do período.

Nas **Figuras 1 a 6**, pode ser visto que os grupos de cor lilás sempre tem o menor valor de

¹ SatImagExplorer é um software livre para mineração de dados desenvolvido pelo ICMC/USP, Embrapa Informática Agropecuária e Cepagri/Unicamp.

índice de vegetação. Para o NDVI, os valores deste grupo variam entre 0,10 e 0,15 e para o EVI os valores variam entre 0,07 e 0,09, indicando que esta região é de rios ou de alagamentos. Por outro lado, os grupos de vermelho apresentam altos valores de índice de vegetação, para NDVI na média de 0,79 e para o EVI com média de 0,50 onde se concentram áreas de vegetação mais densa, florestas. Nas **Figuras 4, 5 e 6**, com resultados para o EVI, os grupos cor de laranja e amarelo esverdeado correspondem à região de cerrado/savana, com valores de EVI em torno de 0,45. O grupo dos tons azuis, com valores médios de EVI em 0,35 são áreas de pastagem ou vegetação rasteira. Finalmente, os tons verdes, com EVI médio de 0,39, identificam áreas de culturas agrícolas. Para o NDVI não foi possível identificar precisamente áreas de cerrado/savana, pastagem, vegetação rasteira e áreas de culturas agrícolas, pois os resultados dos agrupamentos não apresentaram padrões esperados como indicou a análise dos especialistas. O método de agrupamento utilizado apresentou grupos com mais coesos e separados utilizando as séries de EVI, independente do regime de chuvas dos anos estudados. Desta forma, o resultado foi coerente tanto para o ano de 2008/2009 que foi muito chuvoso, como para 2010/2011 com chuvas na média, porém com irregularidade na distribuição, como pode ser observado nas **Figuras 4 e 5**.

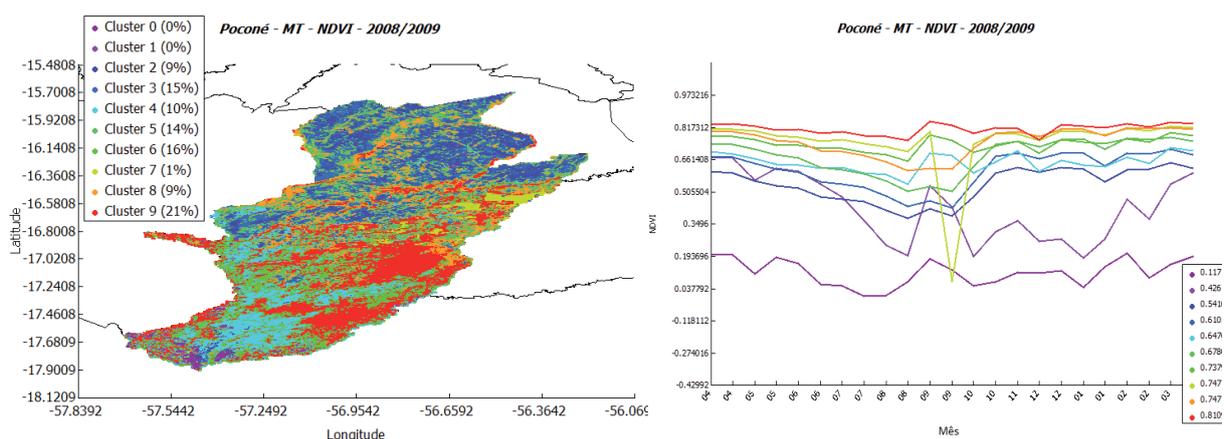


Figura 1. Mapa e gráfico dos agrupamentos resultantes do algoritmo K-means, com distância Euclidiana, para os índices de vegetação NDVI, 2008/2009, Poconé – MT.

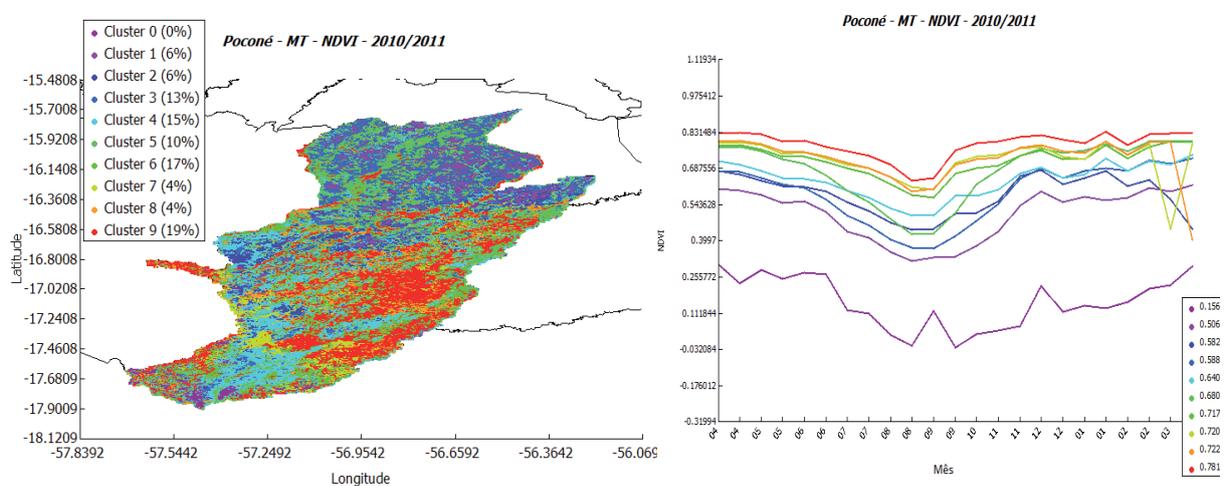


Figura 2. Mapa e gráfico dos agrupamentos resultantes do algoritmo K-means, distância Euclidiana, para os índices de vegetação NDVI, 2010/2011, Poconé – MT.

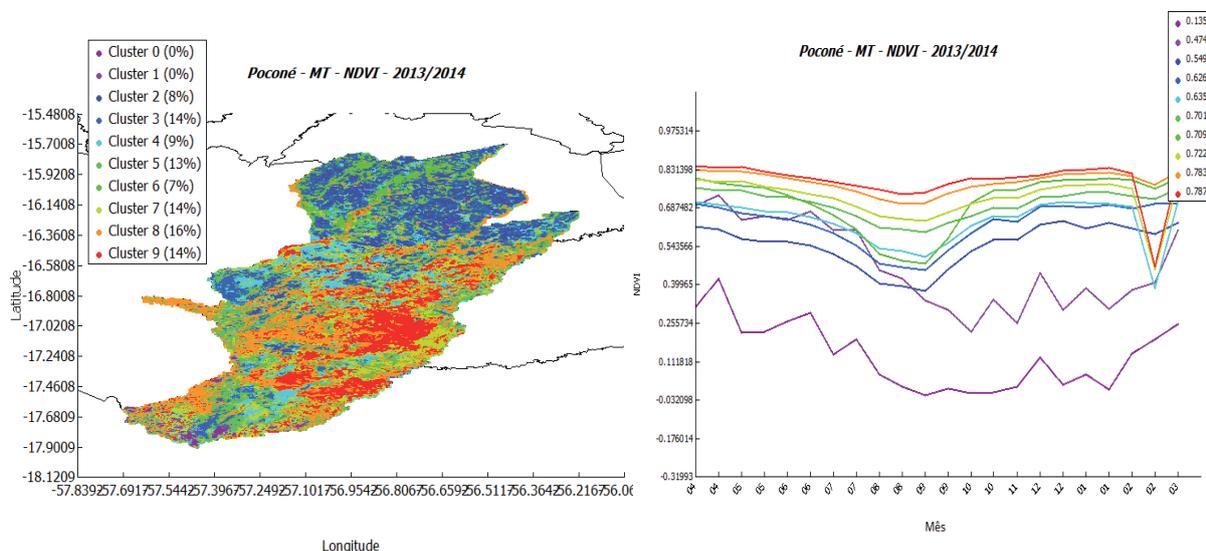


Figura 3. Mapa e gráfico dos agrupamentos resultantes do algoritmo K-means, distância Euclidiana, para os índices de vegetação NDVI, 2013/2014, Poconé – MT.

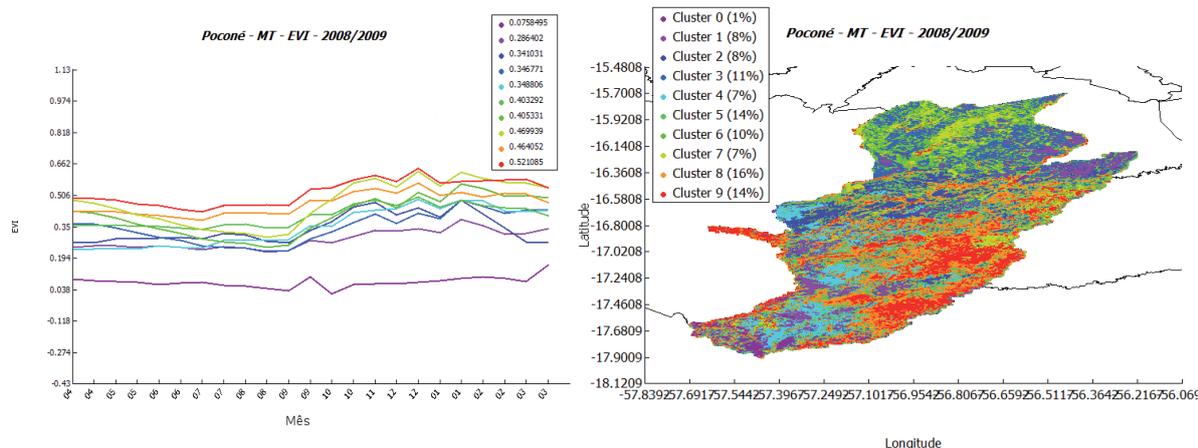


Figura 4. Mapa e gráfico dos agrupamentos resultantes do algoritmo K-means, distância Euclidiana, para os índices de vegetação EVI, 2008/2009, Poconé – MT.

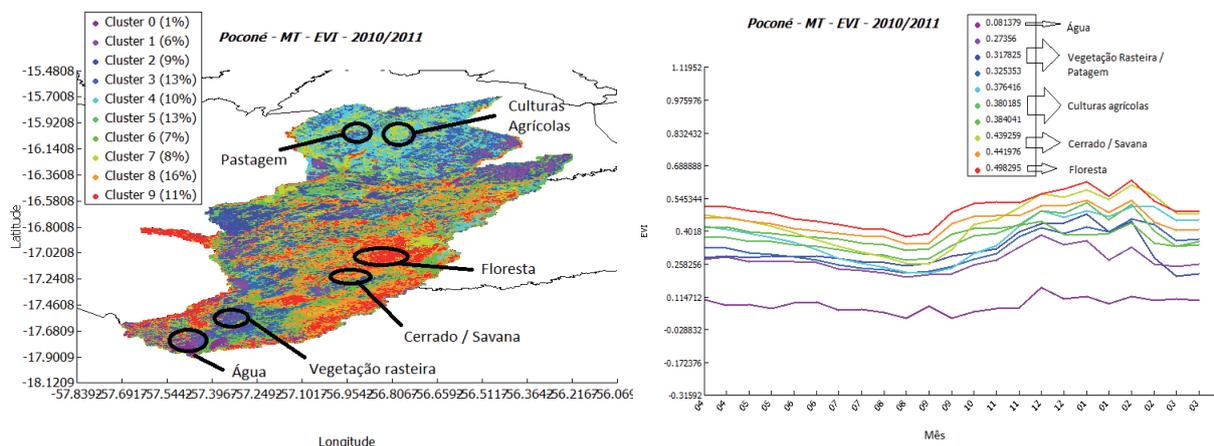


Figura 5. Mapa e gráfico dos agrupamentos resultantes do algoritmo K-means, distância Euclidiana, para os índices de vegetação EVI, 2010/2011, Poconé – MT.

clidiana, para os índices de vegetação EVI, 2010/2011, Poconé – MT; com legenda informativa dos tipos de vegetação de acordo com a região e valor de EVI.

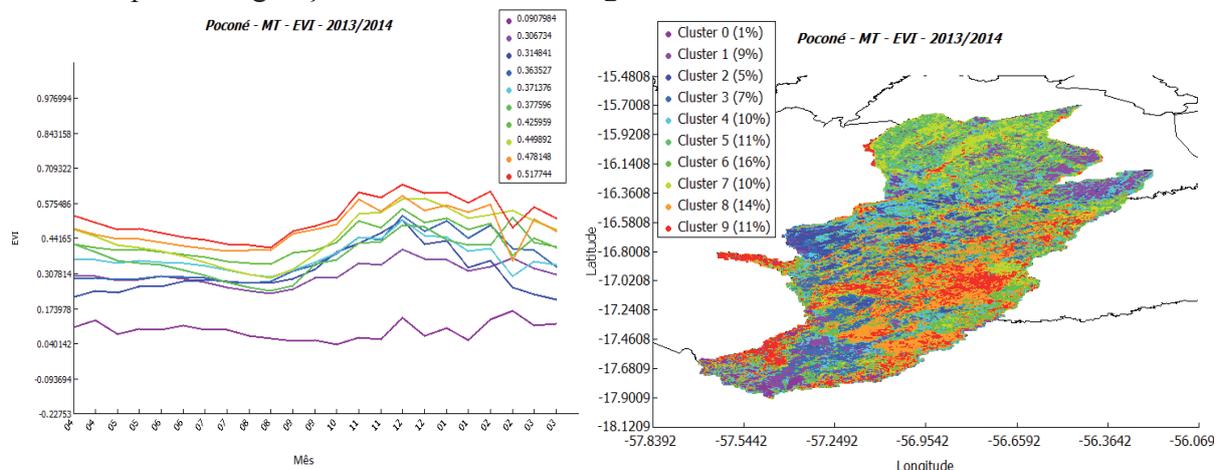


Figura 6. Mapas e gráficos dos agrupamentos resultantes do algoritmo k-means, distância euclidiana, para os índices de vegetação EVI, 2013/2014, Poconé – MT.

Analisando-se os gráficos gerados pelo método *Silhouette*, apresentado na **Figura 7**, pode-se observar que para os grupos encontrados utilizando-se as séries de EVI, praticamente todos os valores são positivos, ou seja, estão à direita do eixo principal. Isto indica que os grupos encontrados possuem alta coesão e são separados uns dos outros. No entanto, avaliando-se o resultado para os grupos obtidos após a aplicação do método nos agrupamentos gerados com as séries de NDVI, observa-se que parte dos valores em praticamente todos os grupos é negativo, ou seja, estão do lado esquerdo do eixo principal do gráfico de *Silhouette*, ou estão muito próximos ao valor zero, como pode ser visto na **Tabela 1**. Neste caso, os grupos gerados não estão coesos e bem separados. Os valores variam entre -1 e 1.

Os melhores resultados obtidos pela análise das séries de EVI se dão pelas suas características aperfeiçoadas em relação ao NDVI no que se refere à identificação de água e vegetações densas, compatível com a região analisada.

Pode-se constatar a partir dos dados de precipitação do TRMM que a região apresentou padrão aumentado de precipitação nos anos analisados **Figura 8 (a)**. No entanto, a distribuição de chuva oscilou nos anos analisados, sendo 2008/2009 um ano com alto volume de chuvas, mantendo um padrão típico da região com concentração de chuvas nos meses de verão e escassez de chuva no período de inverno. Os anos de 2010/2011 apresentaram um volume anual de chuvas dentro da média dos anos analisados, porém com uma distribuição irregular de chuvas, com valores muito baixos durante o inverno e grandes volumes de água no verão, **Figura 8 (b)**. Devido à escassez de dados de estações meteorológicas de superfície na região de Poconé, os dados estimados pelo TRMM puderam auxiliar na avaliação da mudança dos agrupamentos identificados ao longo dos anos.

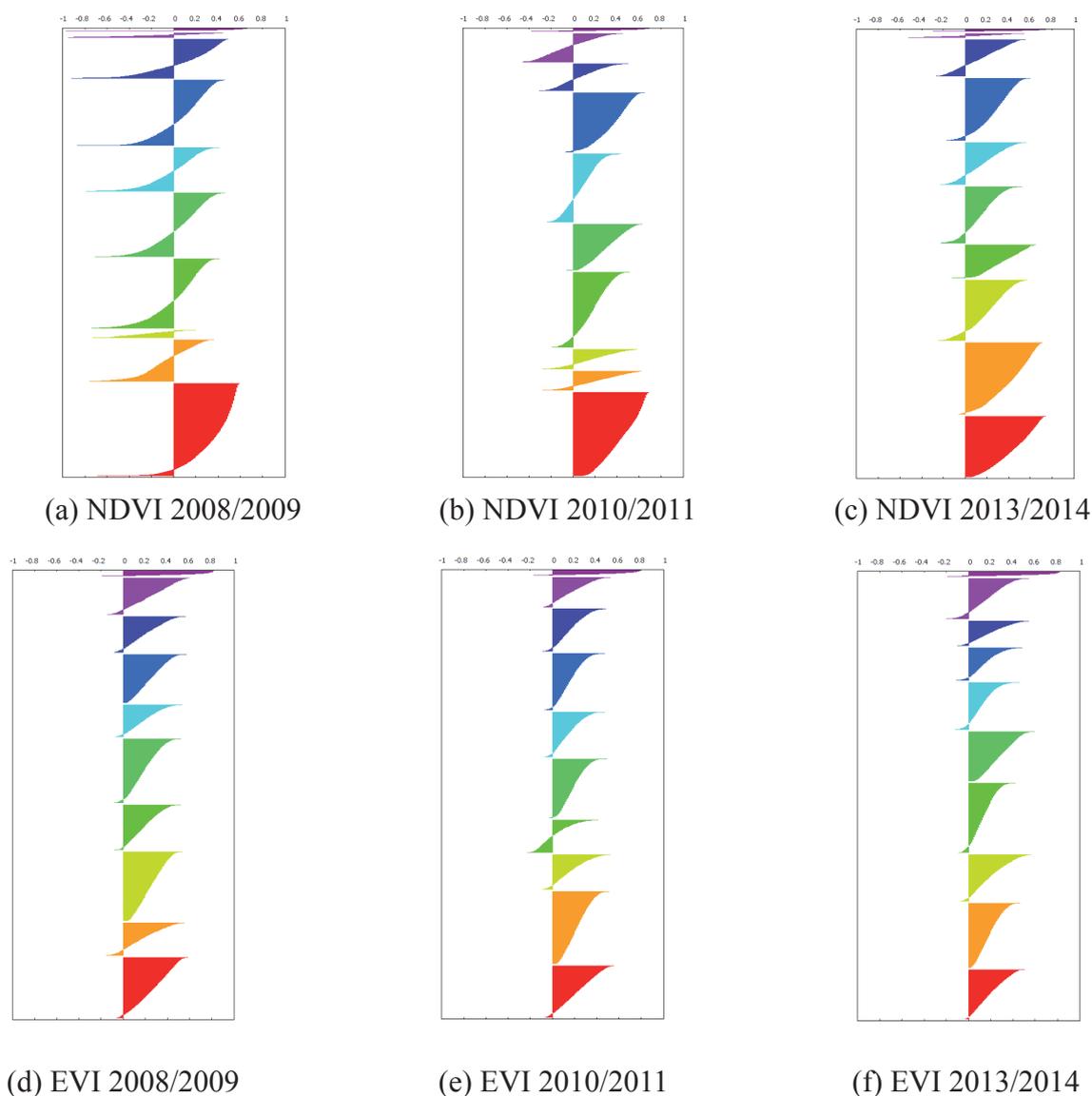


Figura 7. Gráficos gerados pelo método *Silhouette*, para agrupamentos com valores de NDVI e EVI, Poconé-MT.

Tabela 1. Valores gerados pelo método *Silhouette* para agrupamentos de NDVI e EVI.

2008/2009 Grupos	Silhouette		2010/2011 Grupos	Silhouette	
	NDVI	EVI		NDVI	EVI
0	0,36	0,61	0	0,41	0,58
1	-0,13	0,24	1	-0,05	0,18
2	0,1	0,2	2	0,06	0,16
3	0,08	0,25	3	0,35	0,16
4	0	0,19	4	0,07	0,15
5	0,04	0,19	5	0,31	0,19
6	0,02	0,19	6	0,17	0,02
7	-0,24	0,25	7	0,16	0,16
8	-0,07	0,17	8	0,2	0,22
9	0,36	0,27	9	0,45	0,22

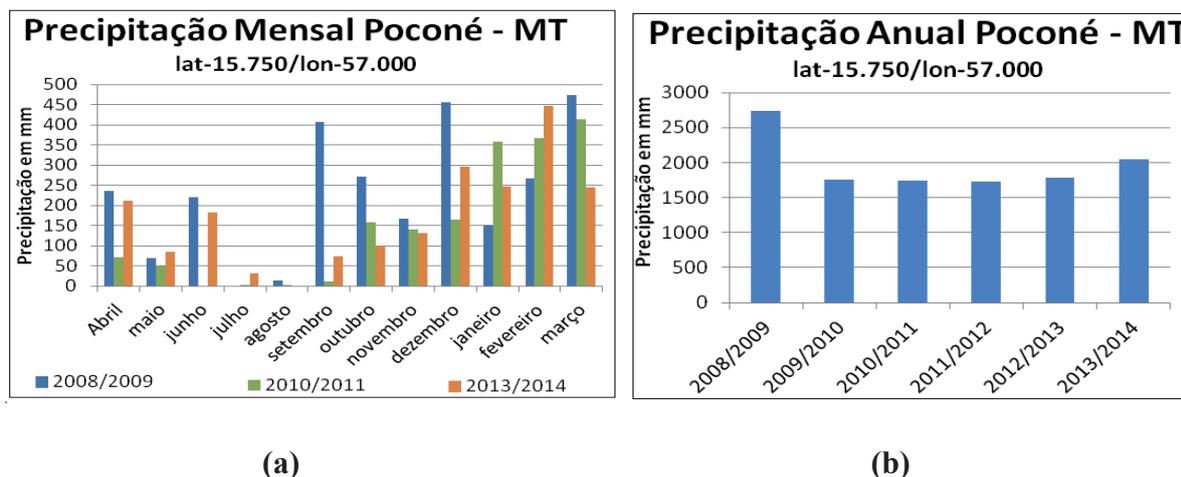


Figura 8. (a) Precipitação acumulada anual para o período entre 2008/2009 e 2013/2014 em Poconé-MT. (b) Distribuição mensal de chuva em Poconé-MT para os anos 2008/2009, 2010/2011 e 2013/2014.

5. Conclusões

Neste trabalho foi apresentada uma metodologia para detectar de forma automática a mudança de uso da terra ao longo dos últimos anos utilizando mineração de séries temporais de imagens de índices de vegetação. A principal contribuição deste trabalho é apresentar um método capaz de realizar uma pré-classificação regionalizada, automatizada, usando grande volume de dados (mais de três milhões de séries temporais) e em áreas nas quais nem sempre se tem uma verdade terrestre. Outra contribuição é mostrar a superioridade da qualidade dos resultados do EVI em relação ao NDVI em áreas de grande concentração de água e/ou vegetação densa como a região estudada. Destaca-se ainda que o EVI possibilitou a identificação de água e mais 4 grupos de vegetação, sendo vegetação rasteira ou pastagem, áreas de culturas agrícolas e região de cerrado/savana. Em trabalhos futuros podem-se realizar novos experimentos de agrupamento com o índice NDVI alterando-se o valor de k na busca de melhores resultados e fazer aplicação do mesmo método usando tanto outros sensores quanto outros municípios. Também podem ser realizadas novas análises separadas para períodos secos e chuvosos.

6. Referências

- Abdon, M. M.; Luciano, A. C. S.; Silva, J. S. V.; Silva, M. S. Classificação de pastagens degradadas nos municípios de Corguinho e Rio Negro, MS, utilizando fusão de imagens CBERS. *Geografia*, v.1,n.1, 2009. P. 709-720.
- Almeida, N. N.; Silveira, E. A.; Barros, L. T. L. P. Mapa de vegetação e uso do solo da da Região de Poconé, MT: I Descrição das unidades. **III Simpósio sobre Recursos Naturais e Socioeconômicos do Pantanal – Os Desafios do Novo Milênio**. Corumbá, MS, novembro 2010.
- Chino, D. Y. T.; Romani, L. A. S.; Traina, A. J. M. Construindo séries temporais de imagens de satélite para sumarização de dados climáticos e monitoramento de safras agrícolas. In: *REIC*, v. 10, p. 1-16. 2010.
- Han, J and Kamber, M. *Data Mining: Concepts and Techniques*, 2nd edition, Morgan Kaufmann. 2006.
- Huete, A; Liu, H. Q.; Batchily, K.; Leween, W. A comparison of vegetation indices over a global set of TM images for EOS-MODIS. *Remote Sensing of Environment*, n. 59, p. 440–451, 1997.

Ibama. **Projeto de Monitoramento do Desmatamento dos Biomas Brasileiros por Satélite – PMDBBS – Pantanal**. Disponível em: <<http://siscom.ibama.gov.br/monitorabiomas/pantanal/pantanal.htm>>. Acesso em: 06 ago. 2014.

MMA. **Biomas – Pantanal**. Disponível em: <<http://www.mma.gov.br/biomas/pantanal>>. Acesso em: 06 ago. 2014.

Rouse, J.W., Haas, R. H., Schell, J. A., and Deering, D.W. Monitoring vegetation systems in the Great Plains with ERTS. In: **Proceedings of the Third ERTS Symposium**, v. 1, 1973. p. 309–317.

Rousseeuw, P. J. **Silhouettes: a Graphical Aid to the Interpretation and Validation of Cluster Analysis**. *Computational and Applied Mathematics*, v. 20, 1987. p. 53–65.

Santos, J. R. Avanços das pesquisas e aplicações de sensoriamento remoto no monitoramento da paisagem: contribuições aos estudos do Pantanal. **Anais 1º Simpósio de Geotecnologias no Pantanal**, Campo Grande, Brasil, 11-15 novembro 2006, Embrapa Informática Agropecuária/INPE, p.675-683.

Santos, J. S.; Pereira, G.; Shimabukuro, Y.E.; Rudorff, B. F. T. Identificação de áreas alagadas no Bioma Pantanal – Brasil – utilizando dados multitemporais TERRA/MODIS. *Geografia*, v.1, n.1, 2009. p. 745-755.

Viana, D. R.; Alvalá, R. C. S. Vegetation index performance for the pantanal region during both dry and rainy seasons. *Geografia*, v. 1, n.1, 2011. P. 143-158.

Victoria, D. C.; Andrade, R. G.; Paz, A. R. Série temporal de imagens EVI/MODIS para discriminação de formações vegetais no Pantanal. *Geografia*, v.1, n.1, 2009. p.721-729.

Yamasoe, M. **Apostila da disciplina Meteorologia por Satélite**. Disponível em < http://www.dca.iag.usp.br/www/material/satelite/apostila_completa.pdf>. Acesso em: 06 ago. 2014.