

MAPEAMENTO DE CULTURAS AGRÍCOLAS UTILIZANDO DADOS MULTITEMPORAL SAR POLARIMÉTRICOS DO SENTINEL-1A

Lucas Volochen Oldoni¹, Victor Hugo Rohden Prudente¹, Denis Corte Vieira¹ e Ieda Del'Arco Sanches¹

¹ Instituto Nacional de Pesquisas Espaciais – INPE, Caixa Postal 515 - 12227-010 - São José dos Campos - SP, Brasil
lucas.oldoni@inpe.br, victor.prudente@inpe.br, denis.vieira@inpe.br, ieda.sanches@inpe.br

RESUMO

A utilização de dados de sensores de microondas é pouco explorada no Brasil para o monitoramento agrícola. Diante disso, o objetivo do presente trabalho é avaliar a utilização de dados polarimétricos do Sentinel-1A para a classificação de culturas agrícolas. Foram utilizados como atributos o retroespalhamento na polarização VV e VH, a razão VV/VH, e a Entropia e Ângulo α obtidos da decomposição polarimétrica H-Alpha. Foi utilizada a abordagem multitemporal, com imagens de 10 datas para a primeira safra e 9 para a segunda safra. A exatidão global para as duas safras foi superior a 0,83 e as acurácias de produtor e usuário foram superiores a 0,83 para as principais culturas da área de estudo (soja, milho e algodão). Conclui-se que os dados polarimétricos do Sentinel-1A são adequados para classificar culturas agrícolas.

Palavras-chave — Random Forest, Soja, Milho, Algodão.

ABSTRACT

The use of data from microwave sensors is little explored in Brazil for agricultural monitoring. Therefore, the objective of this work is to evaluate the use of Sentinel-1A polarimetric data for the classification of agricultural crops. The VV / VH polarization, the VV / VH ratio, and the Entropy and Angle α obtained from the H-Alpha polarimetric decomposition were used as attributes. A multitemporal approach was used, with images of 10 dates for the first harvest and 9 for the second harvest. The overall accuracy for the two harvests was over 0.83 and the producer and user accuracy were greater than 0.83 for the major crops in the study area (soybean, maize and cotton). Thus, the Sentinel-1A's polarimetric data seems to be useful to classify agricultural crops.

Key words — Random Forest, Soybean, Maize, Cotton.

1. INTRODUÇÃO

Com o contínuo crescimento populacional e a estagnação da quantidade de áreas agrícolas disponíveis mundialmente, a segurança alimentar se torna uma preocupação global [1]. Assim, os esforços atuais de agricultores, técnicos e interessados diretos, devem garantir que a produção de alimentos seja maximizada e ao mesmo tempo sustentável, exigindo uma atualização regular das informações espaciais e temporais sobre as atividades agrícolas [2].

A discriminação e mapeamento de culturas agrícolas é um passo crítico para várias aplicações, como monitoramento agrícola e planejamento econômico [3]. Dados de

sensoriamento remoto óptico, principalmente com abordagens multitemporais, são adequados para monitorar e discriminar espécies agrícolas [4]. Entretanto, mesmo em abordagens temporais, a cobertura de nuvens pode afetar a disponibilidade de dados de sensores ópticos e consequentemente a confiabilidade do monitoramento da agricultura [5].

Neste sentido, sensores de microondas, como os *Synthetic Aperture Radar* (SAR), coletam dados independentemente das condições de iluminação e são pouco afetados pelas condições atmosféricas e por nuvens [6]. Possibilitando assim a obtenção de densas séries temporais de dados, mesmo em áreas com grande coberturas de nuvens, tais como regiões tropicais.

As interações entre a radiação eletromagnética e a vegetação na faixa de microondas depende de fatores ligados a geometria (estrutura) e constante dielétrica do dossel, além de fatores ligados ao sistema de manejo, como densidade de plantio e da orientação da linha, e de fatores ligados ao sensor, como ângulo de incidência e polarização da onda incidente [6]. Por estes motivos, o retroespalhamento mostra variações distintas nos estágios fenológicos das culturas [7], sendo uma mistura de espalhamento de superfície, espalhamento duplo e dispersão de volume. Para qualquer cultura, a contribuição de cada uma dessas fontes de espalhamento para o total de retroespalhamento é principalmente uma função do estágio de crescimento no ciclo da cultura [6].

Muitos SARs obtêm dados em mais de uma polarização, fornecendo informações sobre o espalhamento polarimétrico, que podem melhorar a capacidade do mapeamento da cobertura da terra, a detecção de alvos e o monitoramento de mudanças nas características do solo dentro das áreas de interesse [6].

A decomposição polarimétrica de alvos é um método que representa o espalhamento dos mesmos por vários mecanismos básicos de dispersão [8]. Neste sentido, com o emprego da decomposição de Cloude-Pottier, é possível obter os atributos polarimétricos de entropia (H) e ângulo alfa (α) [9]. O Sentinel-1A transporta um SAR de banda C que coleta dados em duas polarizações (HH+HV ou VV+VH) [10]. Porém, os SARs com polarização HH-HV e HV-VV (e.g. Sentinel-1A) não extraem adequadamente mecanismos de dispersão no plano H- α , como ocorre com dados de polarização completa, indicando que a co-polarização é vital para a extração dos mecanismos de dispersão [8].

Nesse contexto, o objetivo do presente trabalho é avaliar a utilização de dados polarimétricos Sentinel-1A para classificação de culturas agrícolas.

2. MATERIAIS E MÉTODOS

2.1. Dados de campo

A área de estudo compreende o município de Campo Verde, no estado do Mato Grosso, Brasil (Figura 1). Este município está localizado no bioma Cerrado, com extensão territorial de 4794,56 km², possui temperatura média anual de 22,3°, precipitação média de 1726 mm e clima Tropical Aw de acordo com a classificação de Köppen-Geiger [11]. Na safra de 2015 as principais culturas semeadas em Campo Verde foram soja, milho e algodão, com área de 210.100 ha, 88.760 ha e 81.996 ha, respectivamente [12].

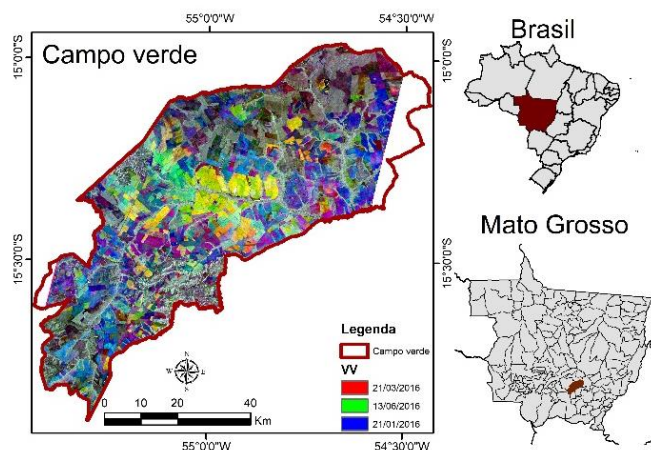


Figura 1. Município de Campo Verde, estado de Mato Grosso, Brasil.

Para esta localidade foram utilizadas dados de culturas agrícolas provenientes da *Campo Verde Database* [13], de outubro de 2015 a julho de 2016, sendo que as campanhas de campo foram realizadas nos meses de dezembro e maio, período da primeira e segunda safra, respectivamente.

As 14 classes existentes no *Campo Verde Database* foram agrupadas em quatro classes para a primeira safra (setembro de 2015 a fevereiro de 2016): Soja, Pastagem, Outros (Eucalyptus e Cerrado) e OAA (Outra Áreas Agrícolas – Beans, NCC brachiaria, NCC grasses, NCC millet, Sorghum, Turfgrass, Soil); e cinco para a segunda safra (fevereiro a agosto de 2016): Algodão, Milho, Pastagem, Outros e OAA (Beans, NCC brachiaria, NCC grasses, NCC millet, Sorghum, Turfgrass, Soil, NCC crotalaria). Na primeira safra foram utilizadas a informação de 353, 49, 31 e 80 polígonos correspondente ao mês de dezembro, para as classes Soja, Pastagem, Outros e OAA, respectivamente. Na segunda safra utilizou-se informação de 199, 29, 205, 36 e 55 polígonos correspondente a maio, para as classes Algodão, Outros, Milho, OAA e Pastagem, respectivamente.

2.2. Dados SAR/Sentinel-1A

Foram utilizados dados SAR do sensor Sentinel-1A, no modo de imageamento IW (*Interferometric Wide Swath*) em Level-1 SLC (*Single Look Complex*), em duas polarizações (VV e VH),

com faixa de imageamento de 250 km e resolução espacial de 5x20m. As imagens neste nível de processamento (SLC) compreendem dados complexos com amplitude e fase.

Para a primeira safra foram utilizadas imagens de 10 datas (30 de agosto de 2015 a 14 de fevereiro de 2016). Para a segunda safra utilizou-se imagens de 9 datas (14 de fevereiro a 24 de agosto de 2016). Para cobrir a área de estudo foram necessárias duas cenas da mesma data (exceto para a data de 21/01/2016). O ângulo de incidência variou de 41,4° para *near range* à 53,4° para *far range*, na área de estudo. As imagens foram obtidas através do Sentinel Scientific Data Hub, e pré-processadas utilizando os softwares *Sentinel Application Platform-SNAP* e o *Sentinel-1 Toolbox* [14].

2.3. Processamento dos dados SAR/Sentinel-1A

O processo de *TOPSAR Split* foi realizado inicialmente para selecionar a faixa de imagem, e apenas os *bursts* que cobriam a área de estudo. Na sequência realizou-se o *TOPSAR Deburst*. Após, dividiu-se o processamento, para obter os atributos polarimétricos de entropia e ângulo α em uma das partes e outra para obter o coeficiente de retroespalhamento para as polarizações VV e VH.

Na primeira parte, dos coeficientes de retroespalhamento, foram realizados os processos de correção de órbita (*Apply Orbit File*) e a remoção do ruído termal (*Thermal Noise Removed*). Posteriormente, usando o módulo “*Calibration*” do SNAP, se deu a etapa de calibração radiométrica para se obter os coeficientes de retroespalhamento, sigma zero (σ^0) nas polarizações VV e VH. Com estes coeficientes gerou-se a razão σ_{VV}/σ_{VH} . Em seguida aplicou-se o processamento denominado de *Multilook* com um tamanho de janela de 4x1 pixels, sendo o espaçamento entre os pixels convertido de 2,33 e 14,05 m para 14,05 m nas direções de *range* e azimute, respectivamente. Este procedimento auxilia na redução do ruído *Speckle*.

Também com o objetivo de reduzir o ruído *Speckle*, foi então aplicado um filtro de *Lee* com janela de tamanho de 5x5 pixels. Posteriormente, os dados SAR foram submetidos ao processo de correção do terreno (*Range Doppler Terrain Correction*) utilizando um modelo digital de elevação (DEM) do *Shuttle Radar Topography Mission* (SRTM) para geocoficar as imagens. Este processo gerou imagens com resolução espacial de 14,05 m. Como última etapa, as polarizações VV e VH foram convertidas da escala linear para a escala logarítmica (dB). Após todas as correções, foi realizado o mosaico das imagens.

Na segunda parte, para a obtenção dos atributos polarimétricos, após o *TOPSAR Deburst*, as imagens também foram submetidas ao processo de reamostragem através do *Multilook* de 4x1 pixels e filtradas com o filtro *Refined Lee* com janela de tamanho 5x5 pixels para reduzir o ruído *Speckle*. Na sequência, foi gerado a matriz de Covariância (C2). Então, foi realizada a decomposição polarimétrica H-Alpha, para se obter os atributos de entropia e ângulo α . Analogamente a primeira parte, também foi aplicada a correção do terreno utilizando o

DEM do SRTM e realizado o processo de mosaico para as imagens que necessitavam.

2.4. Treinamento, classificação e validação

Para treinamento dividiu-se aleatoriamente 70% dos polígonos de cada classe. Foram extraídos todos os pixels que se encontravam dentro dos polígonos e selecionado aleatoriamente 60000 pixels para cada classe.

Para a classificação utilizou-se o algoritmo *Random Forest* (RF) [15] da biblioteca Scikit-Learn [16] de Machine Learn para Python. O RF é adequado para classificar culturas agrícolas usando grande volume de dados [17].

Para avaliação do desempenho do modelo de classificação foi utilizada validação cruzada com duas partições sendo 90% para treinamento e 10% para validação e 10 repetições. Para avaliação da classificação foram utilizados todos os pixels que se encontravam dentro dos 30% de polígonos que não foram utilizados para treinamento, para se ter amostras independentes daquelas utilizadas para treinamento. Gerou-se então a matriz de erros e calculou-se a exatidão global, acurácia de produtor (AP) e usuário (AU) [18].

3. RESULTADOS E DISCUSSÃO

A classificação obtida para a primeira (Figura 2) e segunda safra (Figura 3) apresentam distribuição espacial semelhante para as classes Pastagem e Outros. Na classificação para primeira safra há a presença de mais pixels isolados de OAA classificados entre áreas de Soja. O mesmo ocorre entre as classes Pastagem e Outros. Já na segunda safra este efeito ocorre principalmente entre as classes Pastagem e Outros. Isso acontece devido ao efeito do ruído *Speckle* que mesmo após aplicação de filtro ainda causa influência.

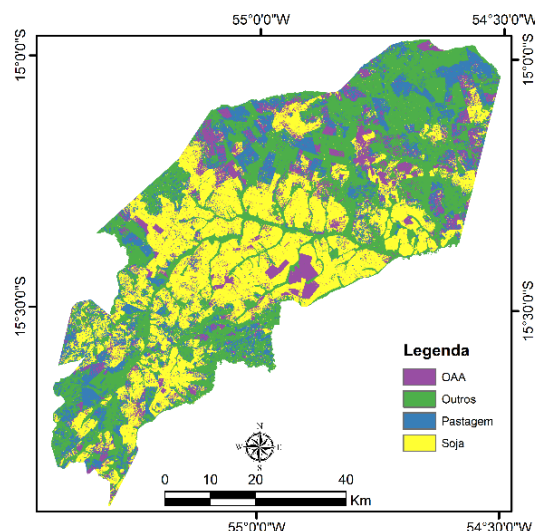


Figura 2. Classificação para primeira safra.

Considerando a avaliação do modelo de classificação a EG foi de 93,00% e 99,97%. Levando em consideração a avaliação da classificação com os polígonos independentes daqueles

utilizados para treinamento, a EG para a primeira (Tabela 1) e segunda safra (Tabela 2) foi de 83,29% e 89,08%, respectivamente. A classe de soja obteve maior acurácia de produtor (0,83) e usuário (0,99) na primeira safra. Apenas 1% das áreas classificadas como soja não são soja. Enquanto que 17% das áreas de soja não foram classificadas como soja, sendo classificadas principalmente como OAA. Isso ocorre provavelmente devido à algumas das culturas que integram a classe OAA possuírem estrutura e/ou ciclo semelhante à soja.

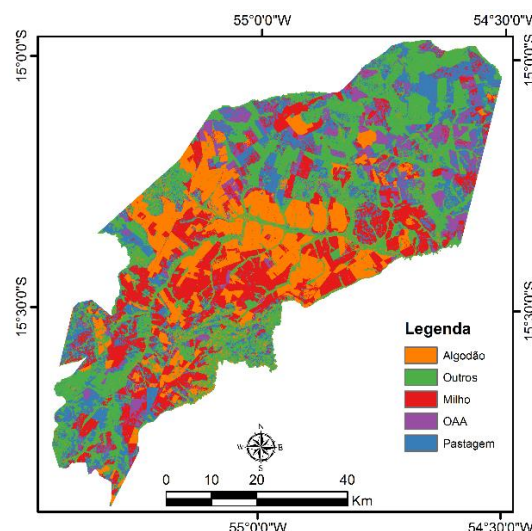


Figura 3. Classificação para segunda safra.

A classe Pastagem obteve a menor acurácia de produtor (0,67), pois muitas áreas de pastagem foram classificadas como Outros. Dentro da classe de Outros existe áreas de cerrado que possui campos limpos [19], em que a vegetação é semelhante à pastagem. Isso explica a confusão ocorrida entre as classes. A classe Pastagem também obteve a menor acurácia de usuário (0,4), pois uma grande parcela das áreas classificadas como Pastagem são Soja. O mesmo ocorre com a classe OAA, que também teve baixa acurácia de usuário (0,54), pois muitas das áreas classificadas nesta classe são Soja na referência.

Tabela 1. Matriz de erros e acurácias para primeira safra.

		Referência				Acurácia do usuário
		OAA	Outros	Pastagem	Soja	
Classificação	OAA	91051	183	1246	75082	0.54
	Outros	5203	74798	13072	16143	0.68
	Pastagem	9494	3476	29258	30379	0.40
	Soja	6198	124	323	606966	0.99
Acurácia do produtor		0.81	0.95	0.67	0.83	
Exatidão global						0.83

Para a segunda safra (Tabela 2), as maiores acurácias de usuário foram obtidas para as classes Algodão (1,00) e Milho (0,89). Sendo que a classe OAA obteve a menor acurácia de usuário, 0,47, principalmente por parte das áreas classificadas nesta classe pertencerem à classe milho. Uma das culturas que

compõe a classe OAA é o sorgo, que possui estrutura semelhante ao milho, o que pode ser um dos motivos da confusão entre estas classes. As acurácias de produtor foram superiores a 0,90 para todas as classes, com exceção da OAA (0,54). Isto porque parte das áreas que pertenciam a esta classe OAA nos dados de referência foram classificadas principalmente como Milho, de maneira errônea. Confusão que ocorre por motivos que já foram explicados anteriormente.

Tabela 2. Matriz de erros e acurácias para segunda safra.

		Referência					Acurácia do usuário
		Algodão	Outros	Milho	OAA	Pastagem	
Classificação	Algodão	379252	55	1178	31	17	1,00
	Outros	2939	57136	6001	1441	12	0,85
	Milho	18140	1817	292936	15509	310	0,89
	OAA	16521	292	18294	33016	2581	0,47
	Pastagem	2709	267	5989	11126	95811	0,83
Acurácia do produtor		0,90	0,96	0,90	0,54	0,97	
Exatidão global							0.89

4. CONCLUSÕES

A utilização dos dados polarimétricos do Sentinel-1A mostrou bons resultados de acurácia com exatidão global acima de 0,83 e acurácias de usuário e produtor para as classes soja, milho e algodão também superiores a 0,83. Porém, por causa do efeito do ruído *Speckle* houve muitos pixels classificados de forma isolada.

Como trabalhos futuros, sugere-se a utilização de segmentação de imagens e classificação por regiões. Outra possibilidade é avaliar a utilização conjunta de dados do Sentinel-1 com dados de sensores ópticos, tais como Sentinel-2/MSI e Landsat-8/OLI.

5. AGRADECIMENTO

O presente trabalho foi realizado com apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ) e da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001, pela concessão de bolsa.

6. REFERÊNCIAS

[1] Thenkabail, P. S. "Global croplands and their importance for water and food security in the twenty-first century: Towards an evergreen revolution that combines a second green revolution with a blue revolution," *Remote Sens.*, vol. 2, no. 9, pp. 2305–2312, 2010.
[2] Kenduywo, B.K.; Bargiel, D. e Soergel, U. "Crop type mapping from a sequence of TerraSAR-X images with dynamic conditional random fields," *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. III-7, no. July, pp. 59–66, Jun. 2016.
[3] Tamiminia, H.; Homayouni, S.; McNairn, H. e Safari, A. "A particle swarm optimized kernel-based clustering method for crop mapping from multi-temporal polarimetric L-band SAR observations," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 58, pp. 201–212, 2017.

[4] Skriver, H.; Mattia, F.; Satalino, G.; Balenzano, A.; Pauwels, V.R.N.; Verhoest, N.E.C.; Davidson, M. "Crop Classification Using Short-Revisit Multitemporal SAR Data," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 4, no. 2, pp. 423–431, 2011.
[5] McNairn, H.; Shang, J.; Champagne, C. e Jiao, X. "TerraSAR-X and RADARSAT-2 for crop classification and acreage estimation," *Int. Geosci. Remote Sens. Symp.*, vol. 2, pp. 898–901, 2009.
[6] Liu, C.; Shang, J.; Vachon, P.W. e McNairn, H. "Multiyear Crop Monitoring Using Polarimetric RADARSAT-2 Data," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 4, pp. 2227–2240, Apr. 2013.
[7] Hariharan, S.; Mandal, D.; Tirodkar, S.; Kumar, V.; Bhattacharya, A. e Lopez-Sanchez, J.M. "A Novel Phenology Based Feature Subset Selection Technique Using Random Forest for Multitemporal PolSAR Crop Classification," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, no. September, pp. 1–15, 2018.
[8] Ji K. e Wu, Y. "Scattering mechanism extraction by a modified Cloude-Pottier decomposition for dual polarization SAR," *Remote Sens.*, vol. 7, no. 6, pp. 7447–7470, 2015.
[9] Cloude S.R. e Pettier, E. "A review of target decomposition theorems in radar polarimetry," *IEEE Trans. Geosci. Remote Sens.*, vol. 34, no. 2, pp. 498–518, 1996.
[10] ESA Report, "Sentinel-1 Product Definition" 2016. Disponível em: <https://sentinel.esa.int/documents/247904/1877131/Sentinel-1-Product-Definition>. Acesso em: 19 de julho de 2018.
[11] Peel, M.C.; Finlayson, B.L. e McMahon, T.A. "Updated world map of the Köppen-Geiger climate classification," *Hydrol. Earth Syst. Sci.*, vol. 11, no. 5, pp. 1633–1644, 2007.
[12] IBGE - Instituto Brasileiro de Geografia e Estatística, "Produção agrícola municipal," 2015. Disponível em: <https://sidra.ibge.gov.br/pesquisa/pam/tabelas>. Acesso em: 08 de outubro de 2018.
[13] Sanches, I.D.; Feitosa, R.Q.; Diaz, P.M.A.; Soares, M.D.; Luiz, A.J.B.; Schultz, B. e Maurano, L.E.P. "Campo Verde Database: Seeking to Improve Agricultural Remote Sensing of Tropical Areas," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 3, pp. 369–373, Mar. 2018.
[14] ESA, 2018. The Sentinel Application Platform (SNAP), a Common Architecture for all Sentinel Toolboxes Being Jointly Developed by Brockmann Consult, Array Systems Computing and C-S. Download em: <http://step.esa.int/main/download/> European Space Agency (ESA).
[15] Breiman, L. "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.
[16] Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M.; Duchesnay, É. "Scikit-learn: Machine Learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2011.
[17] Hao, P.; Zhan, Y.; Wang, L.; Niu, Z. e Shakir, M. "Feature selection of time series MODIS data for early crop classification using random forest: A case study in Kansas, USA," *Remote Sens.*, vol. 7, no. 5, pp. 5347–5369, 2015.
[18] Congalton, R.G. "A review of assessing the accuracy of classifications of remotely sensed data," *Remote Sens. Environ.*, vol. 37, no. 1, pp. 35–46, 1991.
[19] Ribeiro, J. F.; Walter, B. M. T. As principais fitofisionomias do bioma Cerrado. In: Sano, S. M.; Almeida, S. P.; Ribeiro, J. F. (Ed.). Cerrado: ecologia e flora. Brasília: EMBRAPA Informações Tecnológicas, 2008. p. 152-212.