

MAPEAMENTO DE CULTURAS DE VERÃO E PASTAGEM UTILIZANDO PROCESSO DE SEGMENTAÇÃO E MODELO DIGITAL DE TERRENO

Alex Paludo¹, Willyan Ronaldo Becker¹, Laíza Cavalcante de Albuquerque Silva¹, Jonathan Richetti¹, Jerry Adriani Johann², Marcelo Henrique de Oliveira Santos³, Eduarda Aparecida Rodrigues³, João Felipe Cesar Silveira³

¹Universidade Estadual do Oeste do Paraná, Programa de Pós-Graduação em Engenharia Agrícola, Cascavel, Paraná, Brasil, {paludo.alex@hotmail.com; willyanbecker@hotmail.com; laiza.cavalcante@hotmail.com; j_richetti@hotmail.com};

²Universidade Estadual do Oeste do Paraná, Docente no Programa de Pós-Graduação em Engenharia Agrícola, Cascavel, Paraná, Brasil, {jerry.johann@unioeste.br}, ³Universidade Estadual do Oeste do Paraná, Graduando em Engenharia Agrícola, Cascavel, Paraná, Brasil, {marcelohenriquetricolor@gmail.com; dudahrodri58@gmail.com; joaofelipecs17@gmail.com};

RESUMO

O estado do Paraná se destaca no cenário nacional pela elevada produção de soja (2º) e milho (3º). Sendo que a mesorregião Oeste é responsável por 20% da área semeada de soja do estado. O objetivo central desta pesquisa, consistiu em fazer o mapeamento dessas culturas, no ano-safra 2016/2017, por meio da plataforma online Google Earth Engine, utilizando as imagens do sensor OLI/Landsat-8 (bandas 6, 5 e 4). Utilizou-se os algoritmos SNIC e SLOPE, para a obtenção, respectivamente, da imagem segmentada e do modelo de elevação de terreno (SRTM Plus). Posteriormente, foi feita uma classificação supervisionada, através do classificador *Continuous Naive Bayes*. A verificação da acurácia do mapeamento foi feita através do índice Kappa ($K=0,81$). Desta forma, foram mapeados 871.899 ha de soja, 114.547 ha de milho primeira safra. Mapeou-se também outras áreas antropizadas, destacando as áreas de pastagem com 555.580 ha.

Palavras-chave Landsat-8, Google Earth Engine, SNIC, Continuo Naive Bayes.

ABSTRACT

The state of Paraná stands out in the national scenario due to the high production of soybeans (2nd) and corn (3rd). The West mesoregion accounts for 20 percent of the soybean area of the state. The objective of this research was to map these crops in the 2016/2017 crop year through the online Google Earth Engine platform using OLI / Landsat-8 sensor images (bands 6, 5 and 4). The SNIC and SLOPE algorithms were used to obtain the segmented image and the terrain elevation model (SRTM Plus), respectively. Subsequently, a supervised classification was made through the Continuous Naive Bayes classifier. The mapping accuracy was verified using the Kappa index ($K = 0.81$). In this way, 871,899 ha of soybean were mapped, 114,547 ha of first harvest corn. Other anthropic areas were also mapped, highlighting the pasture areas with 555,580 ha.

Key words — Landsat-8, Google Earth Engine, SNIC, Continuous Naive Bayes.

1. INTRODUÇÃO

A agricultura no estado do Paraná tem uma grande variabilidade devido às condições favoráveis de crescimento de diversas plantas. Culturas comuns ao Paraná incluem soja, milho, feijão, café, algodão, trigo, mandioca, cana-de-açúcar [1].

O estado do Paraná destaca-se como o segundo maior produtor nacional de soja com quase 20 milhões de toneladas, com área semeada, na safra 2016/2017, de aproximadamente 5.271.620 ha [2, 3]. Somente a mesorregião Oeste do estado foi responsável por 1.030.470 ha da área semeada. Já para o milho, segundo dados oficiais a área semeada na mesorregião Oeste do estado do Paraná foi de aproximadamente 75.446 ha [3].

Devido as condições meteorológicas, o estado do Paraná proporciona o cultivo de duas safras por ano agrícola, caracterizando respectivamente, primeira e segunda safra (verão e inverno). Desta forma, tem sido verificado que com o passar dos anos, as áreas de cultura de verão (1ª Safra) semeadas com milho, estão sendo ocupadas com soja. Portanto, a soja, tem sido a principal cultura de verão no estado do Paraná e, o milho, a mais importante para o inverno, principalmente, devido ao vazio sanitário imposto sobre a soja em todo o território nacional.

Com o advento de novas tecnologias, como a plataforma online Google Earth Engine (GEE), a qual proporciona uma análise de dados globais, é possível ter acesso a vários bancos de dados e suas séries históricas de imagens de satélite, sem ter a necessidade de armazenar e processar estas informações para depois poder as utilizar [5]. Outra importante informação que deve ser utilizada em estudos de mapeamento de culturas agrícolas, são as séries temporais de imagens orbitais. Destacando as imagens de satélite Landsat-8, que são utilizadas com muita intensidade e apresentam boas condições temporal, espacial, espectral e radiométrica, proporcionando confiabilidade nos dados a serem analisados [6].

Outra fonte de dados que pode ter grande importância para a realização de mapeamento são os dados de elevação digital de terreno (SRTM) [7], que é um esforço de pesquisa

internacional para obter modelos de elevação digital em uma escala quase global. O produto SRTM V3 (SRTM Plus) é fornecido pela NASA com uma resolução 30m.

Neste processo, além do uso direto destes dados, vem-se utilizando muito o pré-processamento dos dados. A segmentação tornou-se rapidamente uma ferramenta de pré-processamento muito eficaz que simplifica uma imagem [8], tornando-a uma composição de pequenos clusters de pixels conectados, chamado superpixels [9]. O algoritmo SNIC (*Simple Non-Iterative Clustering*) tem sido utilizado pois melhora a eficiência computacional, o consumo de memória e a qualidade da segmentação quando comparado com seu antecessor, o algoritmo SLIC [8].

O processo crucial de um mapeamento com imagens de satélite é a escolha de um método de classificação. Uma abordagem bem utilizada é a classificação bayesiana, uma técnica que se tornou cada vez mais popular nos últimos anos, em parte devido a desenvolvimento recente no aprendizado com redes de crenças bayesianas [10]. O classificador bayesiano mais simples é o *Continuo Naive Bayes*, o qual é um classificador probabilístico simples que calcula um conjunto de probabilidades referentes a frequência e as combinações de valores do conjunto de dados. O algoritmo assume que todos os atributos são independentes, dado o valor da variável de classe. Essa suposição de independência condicional raramente é verdadeira em aplicações no mundo real, daí a caracterização como *Naive* (ingênuo), mas o algoritmo tende a funcionar bem e a aprender rapidamente em vários problemas de classificação supervisionada (PATIL, 2013).

O objetivo deste trabalho foi realizar o mapeamento das culturas da soja, milho primeira safra e pastagens na mesorregião Oeste do estado do Paraná, para o ano safra 2016/2017, desenvolvendo um algoritmo na plataforma online GEE.

2. MATERIAIS E MÉTODOS

A área de estudo compreende a mesorregião Oeste do estado do Paraná, na região Sul do Brasil (Figura 1). É constituída por 3 microrregiões e 50 municípios [11]. De forma geral, o estado possui uma temperatura média anual do ar variando entre 15 e 24 °C. A precipitação anual do estado do Paraná [1] variam de 1.100 a 1.920 mm ano⁻¹. A fim de estudar regiões mais homogêneas, tanto de clima quanto de características regionais mais específicas, o estudo foi conduzido em nível de microrregião (Cascavel, Toledo e Foz do Iguaçu) com a finalidade de se obter um resultado mais preciso e confiável no mapeamento.

Utilizando a plataforma GEE foi desenvolvido o script da metodologia utilizando a linguagem *javascript*, para a realização do mapeamento dos alvos de interesse (soja, milho e pastagem) para o ano-safra 2016/2017. O primeiro passo realizado foi a importação da coleção de imagens Landsat-8.

As imagens selecionadas do sensor Landsat-8/OLI, foram no período de 01/12/2016 a 31/01/2017, que tivessem no

máximo 10% de nuvens. O período escolhido contempla o máximo desenvolvimento vegetativo (fase de frutificação) das culturas de soja e milho de interesse. A partir das imagens filtradas, se realizou a composição RGB-564 aplicando a operação de mediana no processo. Optou-se por utilizar durante a análise apenas o conjunto de bandas (6,5,4) por estas, compreenderem as regiões do visível e do infravermelho próximo e médio. Regiões estas que proporciona uma análise bem detalhada da diferenciação espectral de todo o ciclo produtivo das culturas de interesse.

A partir do mosaico falsa-cor RGB-564, selecionaram-se amostras de treinamento, de forma aleatória, direcionada estratificada, ou seja, pixels puros que sejam representativos de suas classes. Os alvos de interesse foram: milho 1ª safra, soja e pastagem, coletando-se amostras em cada uma das microrregiões. As amostras selecionadas tinham formato de polígonos amostrais pequenos distribuídos pela região de estudo e cada amostra continha aproximadamente 30 pixels.

Com o intuito de auxiliar a diferenciação entre os alvos, aplicou-se o processo de segmentação no mosaico RGB-564. O algoritmo de segmentação utilizado foi o SNIC. Os parâmetros utilizados na aplicação do algoritmo da segmentação foram: seeds (5), size (5), compactness (1), connectivity (8) e neighborhoodSize (10).

Outro fator relevante é a elevação do terreno no local sob estudo, que influencia no uso e ocupação do espaço agrícola. Neste sentido, foi utilizado o modelo digital de elevação (SRTM Plus) para obtenção da elevação aplicando-se a função *slop* que cria uma imagem mostrando a inclinação em graus de um terreno. Este plano de informação foi necessário para diminuir a confusão entre pastagem e as culturas de soja e milho.

Para a classificação, usou-se o classificador *Continuous Naive Bayes*. Os dados de entrada para o classificador foram: as amostras de treinamento; bandas B4 (0,64 – 0,67 µm), B5 (0,85 – 0,88 µm) e B6 (1,57 – 1,65 µm) do Landsat-8; a imagem de segmentação; a imagem *slop*, que indica o grau de inclinação do terreno da imagem SRTM Plus.

Posteriormente, foi retirado do mapeamento os alvos de floresta e reflorestamento, alvos urbanos e corpos hídricos, os quais foram mapeados em outro trabalho.

O cálculo da acurácia foi efetuado pelo índice Kappa [12], utilizando 300 pontos sorteados, sendo 100 pontos por classe. Juntamente a esse índice também foram calculados a exatidão global e os erros de omissão e inclusão. Efetuou-se também a comparação entre as áreas das culturas com os dados oficiais do IBGE.

3. RESULTADOS E DISCUSSÃO

Na realização do mapeamento para se obter um bom resultado foram coletadas 8 amostras para a microrregião de Cascavel, 5 amostras para a microrregião de Toledo e 4 amostras para Foz do Iguaçu. A microrregião de Cascavel necessitou de mais amostras por obter uma maior variação de alvos e de relevo.

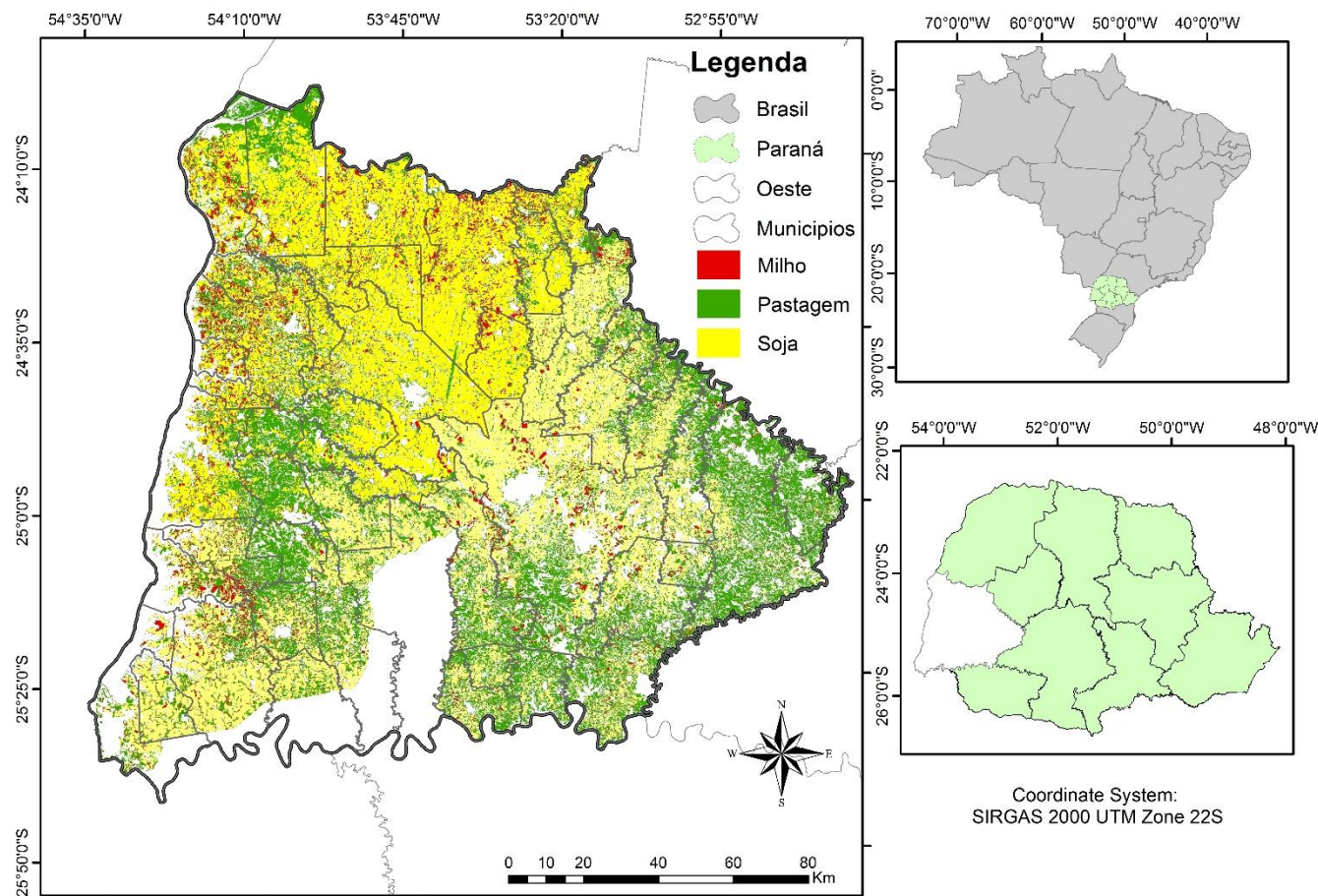


Figura 1. Mapeamento de milho soja e pastagem para a mesorregião oeste do estado do Paraná.

Neste mapeamento (Figura 1) se encontrou uma área total de milho primeira safra de 114.547 ha, contra uma área oficial de 75.446 ha, superestimando a área de milho primeira safra em 51%. Já a área de soja foi de 871.899 ha, no entanto, segundo os dados oficiais, a região possui uma área de 1.030.470 ha, subestimando a área de soja em 15%. Porém, ao se considerar as áreas mapeadas de soja e milho (986.446 ha), constata-se uma subestimação de área de 10,8% em relação aos dados oficiais (1.105.916 ha). A área de pastagem encontrada na mesorregião Oeste foi de 555.580 ha, porém, não foram encontrados dados oficiais para se ter uma comparação. Portanto, foram mapeados 1.542.026 ha (soja, milho e pastagem), o que significa que 67% da área da mesorregião (2.285.100 ha).

A acurácia do mapeamento, medida pelo índice Kappa (0,81) e exatidão global (85,5%), indicaram uma excelente qualidade da classificação. Os erros de inclusão, na classe soja, foram de 4% ocorridos basicamente com pastagem, e os erros de omissão foram de 20%. Já para a classe milho, os erros de inclusão foram de 19% ocorridos em sua maioria com a cultura da soja devido a sua proximidade espectral em algumas regiões, os seus erros de omissão foram de 10%. Na classe pastagem o erro de inclusão foi de 17% tendo ocorrido

principalmente em áreas de bordas dos talhões das classes de soja e milho, e foram encontrados 15% de erros de omissão.

Como demonstrado foi possível realizar o mapeamento da mesorregião Oeste do estado do Paraná utilizando imagens Landsat-8 através da metodologia proposta, assim como outros autores também vem realizando mapeamentos através de imagens *Landsat-8* como [13, 14, 15]. Destaca-se que houve dificuldade no mapeamento por conta da presença de nuvens nas imagens utilizadas, tendo em vista, que não foram encontradas imagens totalmente livres de nuvens para o período em estudo. Outros trabalhos também fazem menção a esta dificuldade com o uso de imagens *Landsat-8* [16, 17, 18].

As regiões que apresentaram nuvens consideradas finas, a classificador mapeou de forma correta. Porém, a medida que as nuvens se tornavam mais espessas, estas regiões não foram adequadamente classificadas, o que gerou erros de omissão e confusão no mapeamento com outros alvos.

Nas regiões em que não houve problemas com nuvens, a metodologia utilizada apresentou ótimos resultados, realizando a separação dos alvos estudados de forma muito precisa. Credita-se parte de sucesso a utilização do processo de segmentação que criou agrupamentos de pixels

homogêneos, deixando-se de se analisar pixels de individualmente, mas sim regiões semelhantes, diminuindo assim a possibilidade de confusão no processo de mapeamento.

5. CONCLUSÕES

Esta metodologia de mapeamento mostrou-se muito eficaz, pois utiliza um baixo número de amostras por classe de solo, o que a torna muito mais rápida e de prática para tarefas de mapeamentos de alvos agrícolas. O mapeamento do Oeste Paranaense, apresentou uma excelente acurácia espacial com exatidão global de 85,5% e índice kappa de 0,81 para o ano-safra 2016/2017.

Foram mapeadas 1.542.026 ha de área de milho (114.547 ha), soja (871.899 ha) e pastagem (555.580 ha) na região Oeste do Paraná, o que corresponde a 67% de ocupação do território com estas culturas.

6. AGRADECIMENTOS

A UNIOESTE - Campus Cascavel, a Pós-Graduação em Engenharia Agrícola (PGEAGRI) e ao Núcleo de Pesquisa GeoScience pela infraestrutura e apoio técnico-científico. A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES), ao CNPq e a Fundação Araucária (FA) pelo apoio financeiro.

7. REFERÊNCIAS

- [1] Aparecido, L. E. de O.; Rolim, G. de S., Richetti, J., Souza, P. S. e Johann, J. A., “Köppen, Thornthwaite and Camargo climate classifications for climatic zoning in the State of Paraná, Brazil”. *Ciência e Agrotecnologia*, volume (40), número (4), páginas (405–417), 2016.
- [2] SEAB, Secretaria da Agricultura e Abastecimento, “Soja - Análise da Conjuntura”. páginas (1–6), 2018. Disponível em: <http://www.agricultura.pr.gov.br/arquivos/File/deral/Prognosticos/2018/Soja_2017_18.pdf>.
- [3] IBGE, Instituto Brasileiro de Geografia e Estatística., “Levantamento Sistemático da Produção Agrícola”. Disponível em: <<https://sidra.ibge.gov.br/tabela/1618#resultado>>. Acesso em: 2018.
- [4] CONAB, Companhia Nacional de Abastecimento. “5º Levantamento - Safra 2017/18 – Grãos”, número (5), páginas (1–142), 2018.
- [5] GOOGLE INC. “Google Earth Engine API”, Disponível em: <<https://developers.google.com/earth-engine/>>. Acesso em: 2018.
- [6] Vanhellemont, Q. e Ruddick, K. “Turbid wakes associated with offshore wind turbines observed with Landsat 8”. *Remote Sensing of Environment*, volume (145.), páginas (105–115.), 2014.

- [7] Farr, T. G.; Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D. e Alsdorf, D. “The Shuttle Radar Topography Mission”, *American Geophysical Union*, número (2005), página (1–33), 2007
- [8] Achanta, R. e Susstrunk, S., “Superpixels and Polygons Using Simple Non-iterative Clustering”, [S.l.]: IEEE, página (4895–4904), 2017.
- [9] Papon, J., Abramov, A., Schoeler, M., & Worgotter, F. *Voxel cloud connectivity segmentation - Supervoxels for point clouds. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, página (2027–2034), (2013).
- [10] Rish, I.; Hellerstein, J. e Jayram, T., “An analysis of data characteristics that affect naive Bayes performance”. *Technical report*, 2001.
- [11] IBGE, Instituto Brasileiro de Geografia e Estatística., “Parana”. 2017, Disponível em: <<https://cidades.ibge.gov.br/brasil/pr/panorama>>. Acesso em: 2018.
- [12] Congalton, R. G.; Green, K. “Assessing the Accuracy of Remotely Sensed Data: Principles and Practices”. 2nd edition. ed. Boca Raton, FL, USA: CRC Press, 2009.
- [13] H. Huang et al., “Mapping major land cover dynamics in Beijing using all Landsat images in Google Earth Engine,” *Remote Sens. Environ.*, volume (202), páginas (166–176), 2017.
- [14] Silva L. C. D. A., Richetti J., Becker W. R., Verica W. R., A. Paludo, and J. A. Johann, “Algoritmos de Aprendizado de Máquina para classificação de soja em imagens Landsat-8,” páginas (65–78), 2017.
- [15] Rodrigues J. A., Libonati R., Peres L. F., and Setzer A., “Burned Area Mapping on Conservation Units of Mountains Region of Rio de Janeiro Using Landsat-8 Data During the 2014 Drought,” *Anuário do Inst. Geociências - UFRJ*, volume (41), número (1), páginas (318–327), 2018.
- [16] Roy D. P. et al., “Characterization of Landsat-7 to Landsat-8 reflective wavelength and normalized difference vegetation index continuity,” *Remote Sens. Environ.*, volume (185), páginas (57–70), 2016.
- [17] Câmara G., Valeriano D. de M., and Soares J. V., “Metodologia para o Cálculo da Taxa Anual de Desmatamento na Amazônia Legal”, 2006.
- [18] Rizzi R. and Rudorff B. F. T., “Soybean crop area estimation in Rio Grande do Sul through Landsat images,” *Rev. Bras. Cartogr.*, volume (3), páginas (226–234), 2005.