

Estimativa de produtividade da cultura da soja na safra 2021/22: Índices de vegetação e *Machine Learning*

Victor Messias Moreira¹ & Daniel Noe Coaguila Nuñez¹

¹ Centro Universitário do Sudoeste Goiano, UniBRAS, Rio Verde, Goiás, Brasil

Correspondência: Victor Messias Moreira, Centro Universitário do Sudoeste Goiano, UniBras, Rio Verde, Estado de Goiás, Brazil. E-mail: messiasmor94@gmail.com

Recebido: Dezembro 05, 2022

Aceito: Dezembro 12, 2022

Publicado: Janeiro 01, 2023

DOI: 10.14295/bjs.v2i1.247

URL: <https://doi.org/10.14295/bjs.v2i1.247>

Resumo

A soja é uma das culturas de maior importância econômica no mundo sendo o Brasil é o maior produtor mundial desse grão. Conhecer a produtividade nem sempre é possível já que estes vem atrelados ao tipo de tecnologia que a fazenda possui e permite avaliar de forma indireta a qualidade do manejo. Assim, objetivou-se estimar a produtividade da cultura da soja na safra 21/22 no sudoeste goiano com o uso de índices de vegetação e *Machine Learning*. Foram utilizados os índices de vegetação EVI, NDRE, NDVI, NDWI e os valores de reflectância da composição RGB do satélite Sentinel 2A e 2B, imagens harmonizadas, livres de nuvens com uma antes da semeadura, durante a lavoura e uma imagem após a colheita. Foram obtidos pontos aleatórios para cada uma das seis classes de produtividade e atribuídos os valores dos índices de vegetação para cada data e classe. A matriz de dados foi processada na plataforma do *Google Colaboratory* utilizando o classificador *Random Forest* do pacote *Scikit-Learn*. Avaliando todos os parâmetros permitidos pelo *Random Forest*, o melhor Score (0,6825), para estimar a produtividade da soja, foi obtido utilizando os parâmetros critério gini, 85% de amostras e 120 estimadores, utilizando todas as imagens recorrentes do período da safra 21/22 e imagens antes da semeadura e após da colheita.

Palavras-chave: *Glycine max*, *Random Forest*, *Google Colaboratory*, sensoriamento remoto.

Estimate of soybean crop productivity in the 2021/22 season: Vegetation indices and *Machine Learning*

Abstract

Soybean is one of the most economically important crops in the world, with Brazil being the world's largest producer of this grain. Knowing the productivity is not always possible since these are linked to the type of technology that the farm has and allows an indirect evaluation of the quality of management. Thus, the objective was to estimate the productivity of the soybean crop in the 21/22 season in southwest Goiás using vegetation indices and machine learning. The vegetation indices EVI, NDRE, NDVI, NDWI and the reflectance values of the RGB composition of the Sentinel 2A and 2B satellite were used, harmonized images, free of clouds with one before sowing, during plowing and one image after harvesting. Random points were obtained for each of the six productivity classes and vegetation index values were assigned for each date and class. The data matrix was processed on the Google Collaboratory platform using the Random Forest classifier from the Scikit-learn package. Evaluating all parameters allowed by Random Forest, the best score (0.6825) to estimate soybean productivity was obtained using the gini criteria, 85% of samples and 120 estimators, using all recurrent images of the harvest period 21/22 and images before sowing and after harvesting.

Keywords: *Glycine max*, random forest, *Google Colaboratory*, remote sensing.

1. Introdução

O Brasil é um dos maiores produtores de soja (*Glycine max* (L.) Merrill) do mundo, a produtividade média nacional no ano safra 2021/2022 dessa oleaginosa, foi de 3.029 kg ha⁻¹, em uma área de 41.452,0 milhões de ha⁻¹, com uma produção de 125.552,3 milhões de toneladas (Conab, 2022).

O conhecimento sobre a produtividade de uma lavoura em específico, se torna importante devido a que esta reflete em grande medida sobre os efeitos controlados, no caso o manejo da lavoura desde o preparo do solo até a colheita (Trindade et al., 2019; Santos et al., 2021). A colheita é o momento em que o produtor irá avaliar o desempenho das tecnologias utilizadas, que vão desde a implantação da lavoura, manejo e tratos culturais, realizados durante todo o desenvolvimento da cultura, visando obter a maior lucratividade (Chioderoli et al., 2012).

O cálculo da produtividade pode ser realizado pelo próprio sistema instalado no maquinário que realiza a colheita como os das marcas Massey Ferguson®, New Holland®, Case-IH®, John Deere®, Valtra® entre outras. Todavia, estes dados de produtividade devem ser tratados antes de trabalhar com eles, pois, alguns fatores inerentes ao maquinário como a largura da plataforma de corte, entrada e saída da lavoura, número de manobras e a limpeza obrigatória, que em muitas situações não ocorrem, além do próprio sensor de rendimento, estão entre os mais importantes para a obtenção de dados confiáveis (Amado et al., 2007; Acosta et al., 2018).

Existem uma outra forma de quantificar a produtividade de uma lavoura. Estamos falando do sensoriamento remoto que a partir de sensores a bordo de diferentes posicionamentos de satélites é possível estimar a produtividade em campo (Zanzarini et al., 2013; Maldaner; Molin, 2019; Trindade et al., 2019), uma vez que, com base nas características da reflectância podem ser utilizadas mediante cálculos matemáticos de bandas espectrais diferentes (Atzberger, 2013; Romansini et al., 2020; Santos et al., 2021).

Os primeiros satélites que orbitam e monitoram o planeta foram lançados em meados do século passado, mas apenas a partir da década dos 80 que possuímos um acervo de imagens disponíveis de forma gratuita. Existem duas missões principais das quais podemos obter imagens, a missão Landsat dos EEUU e a missão Sentinel-2 da Agência Espacial Europeia (ESA, das siglas em inglês). A vantagem do satélite Sentinel-2 em comparação com a família Landsat é a resolução espacial (10 m). Ambas as missões fornecem imagens de forma gratuita, basta apenas acessar o site específico de cada projeto (Zanzarini et al., 2013; Maldaner; Molin, 2019; Romansini et al., 2020; Santos et al., 2021).

O uso de índices de vegetação a nível de escala regional tem se popularizado com o uso de imagens de satélite, estes índices fornecem um histórico do comportamento da lavoura. Índices como o NDVI (*Normalized Difference Vegetation Index* - Índice de Vegetação por Diferença Normalizado) o qual se caracteriza por ser sensível à presença de clorofila e pigmentos clorofilianos absorvedores da radiação solar na faixa do vermelho (Ponzoni et al., 2012; Ponzoni et al., 2015).

O EVI (*Enhanced Vegetation Index* - Índice de Vegetação Aprimorado) sensível a variação da estrutura e arquitetura do dossel das plantas, o qual inclui o índice de área foliar e a estrutura da planta; NDRE (*Normalized Difference Red Edge* - Diferença Normalizada na Banda de Borda Vermelha) é capaz de fornecer análise mais aprimorada, além de apresentar a possibilidade de (diferentemente do NDVI) medir mais a fundo no dossel (Kuechly et al., 2020; Jiang et al., 2021); NDWI (*Normalized difference Water Index* - Índice de diferença normalizada da água) é derivado dos canais NIR e infravermelho de ondas curtas (SWIR), responde a mudanças tanto no conteúdo de água (absorção da radiação SWIR) quanto no mesofilo (refletância da radiação NIR) em copas de vegetação (Gao, 1996; Kuechly et al., 2020; Jiang et al., 2021).

Processar dados espaciais nos fornecem informações de vital importância para a estimativa da produtividade, porém a demanda computacional é muito alta. Assim, surge uma alternativa para os usuários que não possuem um poder de processamento a altura dos dados espaciais, o *Google Colaboratory* consiste em uma plataforma baseada em nuvem que fornece acesso irrestrito aos serviços de nuvem do Google e unidades de processamento gráfico (Garcia; Heras; Sanz-Sáez, 2020).

Inclusive com poder computacional de nuvem, a estimativa de produtividade se tornaria muito onerosa senão utilizar-mos algoritmos sofisticados de *Machine Learning* como o *Random Forest*. O *Random Forest*, fundamentalmente esta baseado em *árvores* de decisão, cada um deles com respostas únicas, assim no resultado se utiliza a sabedoria das massas (das *árvores* utilizadas) para entregar um resultado; e pode ser utilizado para comparar diferentes índices comumente utilizados na agricultura, tendo como base a correlação entre os parâmetros analisados e o índice mais adequado, tornando assim, uma avaliação mais precisa (Zarei; Mahmoudi; Moghimi, 2022).

Perante o exposto objetivou-se estimar neste estudo, a produtividade da cultura da soja na safra 21/22 no Sudoeste goiano brasileiro, com o uso de índices de vegetação e *Machine Learning*.

2. Material e Métodos

2.1 Localização e caracterização da área experimental

O estudo foi realizado na Fazenda Bom Jesus com as coordenadas (17°12'24.37" S e 51°15'19.63" O), altitude de 849 m, no município de Rio Verde, Estado de Goiás, Brasil. O solo da área experimental é classificado como Latossolo Vermelho distroférrico (LVdf), típico, textura muito argilosa, fase Cerrado (Santos et al., 2018).

Segundo a classificação de Peel et al. (2007) e Alvares et al. (2014), o clima da região é classificado como Aw (tropical), com chuva nos meses de outubro a maio, e com seca nos meses de Junho a Setembro. A temperatura média anual varia entre 20 e 35 °C e as precipitações variam entre 1.500 e 1.800 mm anuais e o relevo é suave ondulado (6% de declividade).

2.2 Imagens e índices de vegetação

Foram utilizadas imagens harmonizadas do satélite Sentinel-2A e 2B para a área descrita acima, no período de duração da safra. Foram utilizadas imagens livres de nuvens desde uma semana antes da semeadura até uma semana após a colheita.

Os índices de vegetação utilizados foram NDVI dado pela equação 1; EVI pela equação 2; NDRE pela equação 3, e NDWI pela equação 4:

$$\text{NDVI} = (\text{IVP} - \text{V}) / (\text{IVP} + \text{V}) \text{ Eq. (1)}$$

$$\text{EVI} = \text{G} \cdot (\text{IVP} - \text{V}) / (\text{IVP} + \text{C}_1 \cdot \text{C}_2 \cdot \text{A} + \text{L}) \text{ Eq. (2)}$$

$$\text{NDRE} = (\text{IVP} - \text{RE}) / (\text{IVP} + \text{RE}) \text{ Eq. (3)}$$

$$\text{NDWI} = (\text{pvd} - \text{pivp}) / (\text{pvd} + \text{pivp}) \text{ Eq. (4)}$$

Onde:

IVP – reflectância no infravermelho próximo;

V – Reflectância no vermelho;

A – Reflectância no azul;

C₁ – Coeficiente de correção dos efeitos atmosféricos para a banda do vermelho (6);

C₂ – Coeficiente de correção dos efeitos atmosféricos para a banda do azul (7,5);

L – Fator de correção para a interferência do solo (1);

G – Fator de ganho (2,5);

RE – Energia refletida no intervalo de onda “red edge” do espectro eletromagnético;

pvd - Reflectância na faixa da luz verde,

pivp - Reflectância na faixa do infravermelho próximo.

2.3 Processamento de dados e parâmetros do classificador Random Forest

Os dados de produtividade aferidos pela colheitadeira foram classificados conforme (Tabela 1).

Tabela 1. Classificação dos dados de produtividade aferidos pela colheitadeira, safra 2021/22, Rio Verde, GO, Brasil.

Produtividade da colheitadeira (sacas por hectare)	Classificação
45-50	0
50-55	1
55-60	2
60-65	3
65-70	4
70-75	5

75-80	6
80-85	7

Fonte: Autores, 2022.

A classificação dos dados se torna necessária para facilitar o cálculo do algoritmo. A partir do mapa de colheita, foram alocados pontos aleatórios (100) por cada classe, na sequência esses pontos foram utilizados para extrair os valores dos índices de vegetação utilizados (Figura 1).

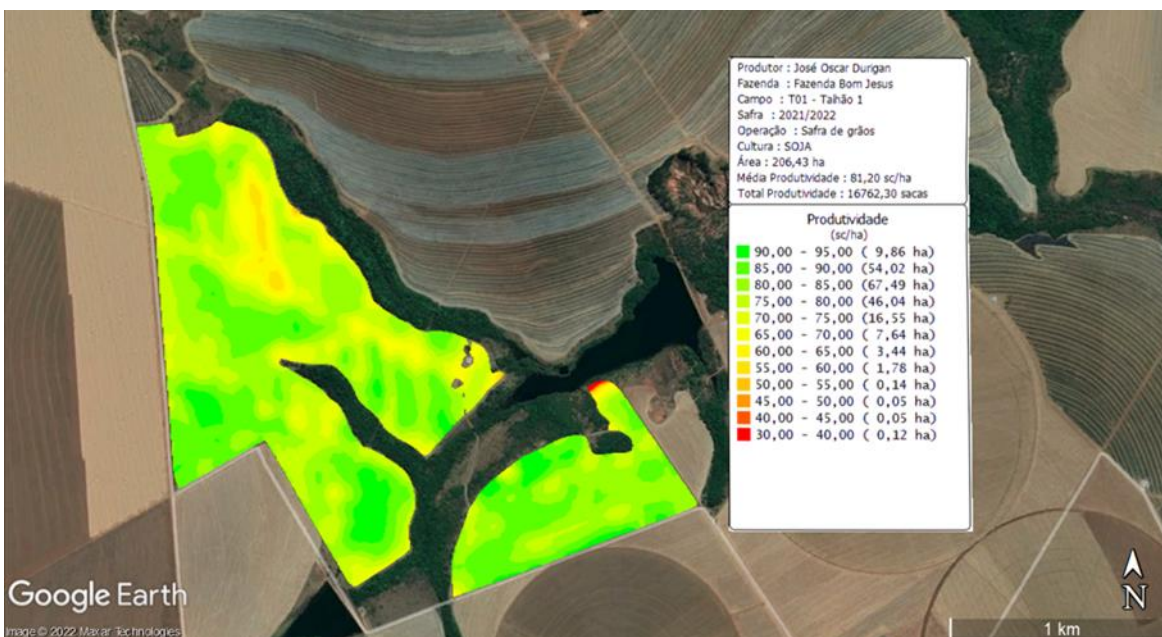


Figura 1. Mapa de produtividade, Rio Verde, GO, Brasil, safra 2021/22. Fonte: Autores, 2022.

Foram utilizados dois modelos matemáticos, o primeiro com índices de vegetação antes e depois da colheita e o segundo com índices de vegetação apenas durante o período da lavoura. Sendo que, para ambos os modelos foram mantidos os parâmetros do *Random Forest* na procura do melhor *Score*.

Com a matriz de dados pronta, ao todo 800 pontos foram utilizados (100 por classe), nesse sentido, foi utilizado a plataforma do *Google Colaboratory*, com o seguinte *Script*:

Contexto e dados

```
import pandas as pd
ivs = pd.read_excel('/content/drive/MyDrive/Dados_Python_2.xlsx')
print (ivs.head())
print (ivs.tail())
```

Estimador: Random Forest

```
from sklearn.ensemble import RandomForestClassifier

bosque = RandomForestClassifier()
bosque.get_params()
```

Busca da grade

```
from sklearn.model_selection import GridSearchCV

parametros = {"criterion": ("gini", "entropy", "log_loss"),
              "n_estimators": (10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160,
                               170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300),
              "max_samples": (0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5, 0.55, 0.6, 0.65, 0.7, 0.75, 0.8, 0.85)}

#Exemplos de scoring = "accuracy", "recall", "roc_auc", etc.

rejilla = GridSearchCV(bosque,
                      parametros,
                      scoring="accuracy",
                      cv=10)

rejilla.fit(ivs[ivs.columns[:-1]].values,
           ivs["Classe"].values)
```

Informação da busca da grade

```
sorted(rejilla.cv_results_.keys())
```

Grade de parâmetros explorados

```
rejilla.cv_results_["params"]
```

Resultados da busca em grade

```
print('ranking de modelo =', rejilla.cv_results_["rank_test_score"])
print('score dos modelos =', rejilla.cv_results_["mean_test_score"])
print('mejor score = ', rejilla.best_score_)
print('mejores parâmetros = ', rejilla.best_params_)
```

Utilizando e extraindo o melhor modelo

```
print(rejilla.predict([[0.131883,0.0549264,0.192003,-0.296583,0.108768, 0.0463271,
0.146554,-0.300479,0.809327,0.0759759,0.741172,0.247573,0.143157,0.857533,0.38574,
0.146591,0.882096,0.431933,0.988897,0.135764,0.831759,0.41955,0.99645,0.15253,0.871917,
0.422176,0.128624,0.0357779,0.110189,-0.185902]])) #colocar aqui os valores que desejam testar

mejor_bosque = rejilla.best_estimator_
print(mejor_bosque.predict([[0.131883,0.0549264,0.192003,-0.296583,0.108768, 0.0463271,
0.146554,-0.300479,0.809327,0.0759759,0.741172,0.247573,0.143157,0.857533,0.38574,
0.146591,0.882096,0.431933,0.988897,0.135764,0.831759,0.41955,0.99645,0.15253,0.871917,
0.422176,0.128624,0.0357779,0.110189,-0.185902]]))
```

Fonte: Autores, 2022.

3. Resultados e Discussão

A estimativa da produtividade foi obtida utilizando *Random Forest*, através da plataforma *Google Colaboratory*, o satélite utilizado foram o Sentinel 2A e 2B. Os dados utilizados foram os índices: EVI, NDRE, NDVI, NDWI, RGB, nas datas: 17-out, 22-out, 06-dez, 21-dez, 26-dez, 15-jan, 20-jan e 01-mar.

Os parâmetros usados no classificador *Random Forest* consistiram em:

criterion: gini, entropy, log_loss;

n_estimators: 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300;

max_samples: 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5, 0.55, 0.6, 0.65, 0.7, 0.75, 0.8, 0.85;

scoring: accuracy;

cv=10.

Os resultados e modelos obtidos estão descritos na tabela 2.

Tabela 2. Modelos matemáticos, safra 2021/22, Rio Verde, GO, Brasil.

		EVI	NDRE	NDVI	NDWI	RGB	
1	17-out	criterion: gini max_samples: 0,85 n_estimators: 120 score: 0,6825 tempo: 1h30min					Modelo com índices de vegetação antes da semeadura e depois da colheita.
	22-out						
	06-dez						
	21-dez						
	26-dez						
	15-jan						
	20-jan						
	01-mar						
2	06-dez	criterion: entropy max_samples: 0,85 n_estimators: 170 score: 0,6675 tempo: 1h20min					Modelo com índices de vegetação apenas durante o período da lavoura.
	21-dez						
	26-dez						
	15-jan						
	20-jan						

Fonte: Autores, 2022.

No presente trabalho, conseguimos um *Score* de 68,3% na estimativa da produtividade, sendo que, quando utilizamos mais dados (índices de vegetação), mesmo da época sem vegetação na área, o modelo (1) aumentou o *Score* quando comparado com o modelo (2), com dados apenas durante o período da lavoura (Tabela 2). Sakamoto et al. (2014) encontraram alta precisão na estimativa de produtividade nos EUA, todavia os autores usaram imagens desde a semeadura, o que resultou em diminuição do erro à medida que mais imagens foram geradas e adicionadas durante a progressão da temporada de cultivo.

Neste trabalho apenas foram incluídos dados de índices de vegetação, acredita-se que com um maior acervo de dados (declividade, teor de argila etc.) o *Score* do modelo poderia aumentar. Todavia, as estimativas da produtividade são altamente afetadas por fatores climáticos, como elevadas precipitações, períodos de estresse e seca, além das decisões agrônômicas. O clima tem grande e importante contribuição para a variabilidade da produção, os modelos de previsão de produtividade possuem um desempenho melhor quando todos os fatores são levados em consideração (Johnson, 2014; Shao et al., 2015).

Além de todos os desafios, encontrar imagens que são consideradas adequadas para a estimativa de produção, ou seja, imagens sem cobertura de nuvens, para o acompanhamento periódico da lavoura são os principais gargalos (Gorelick et al., 2017). Schirmbeck et al. (2020) e Trindade et al. (2019), concluem que o melhor estágio fenológico para registro de imagens da soja é de R3 a R5. Sharma et al. (2015) também mostram que as diferenças entre NDVI e NDRE são mais pronunciadas nos estágios fenológicos iniciais da cultura. Portanto, se considera que são necessários mais estudos visando verificar se a resposta destas relações apresentam semelhanças e ligações.

As etapas mais críticas para a estimativa da produção, é obter informações confiáveis sobre o manejo e tratamentos culturais, aferição dos equipamentos, além da distribuição geográfica (Sakamoto et al., 2014; Shao et al., 2015; Jin et al., 2017). O Serviço Nacional de Estatísticas Agrícolas (NASS) dos EUA por meio da agência de estatísticas do Departamento de Agricultura dos Estados Unidos (USDA), publica uma série de informações detalhadas sobre os campos de produção de todo o país desde 2008, o que corrobora para uma informação com maior veracidade (Lobell, 2013).

Estudos recentes, sobre novas ferramentas e o surgimento do compartilhamento de informações em nuvem, tem possibilitado avanços no sensoriamento remoto, e na exploração de técnicas visando à classificação de culturas com base em imagens de satélite (Sakamoto et al., 2014; Cunha et al., 2018). Este é um passo essencial no desenvolvimento de modelos de previsão quase em tempo real, bem como, os métodos de classificação para mapeamento de culturas, apresentando alta eficiência computacional e robustez contra *Overfitting* (Azzari et al., 2017; Shelestov et al., 2017; Michels et al., 2018).

4. Conclusão

Nas condições do presente estudo com o classificador *Random Forest* (gini, amostras de 0,85 e 120 estimadores) na plataforma *Google Colaboratory* utilizando índices de vegetação na lavoura de soja conseguimos atingir um *Score* de 68,3% para a estimativa da produtividade da soja comparado com os dados de campo coletados pela colheitadeira na safra 2021/22 utilizando índices de vegetação e *Machine Learning*.

5. Referências

- Acosta, J. J. B., Cabrera, M. G., Ibras, R. F., González, J. D., Chamorro, S. M. & Escobar, J. (2018). Variabilidade espacial da produtividade, perdas na colheita e lucratividade da cultura de soja. *Revista Agrogeoambiental*, 10(1), 27-46. <http://dx.doi.org/10.18406/2316-1817v10n120181050>
- Alvares, C. A., Stape, J. L., Sentelhas, P. C. Gonçalves, J. D. M. & Sparovek, G. (2013). Koppen's climate classification map for Brazil. *Meteorologische Zeitschrift*, 22(6), 711-728. <http://dx.doi.org/10.1127/0941-2948/2013/0507>
- Amado, T. J. C., Santi, A. L., Moreira, J. H. & Souza, L. A. (2007). Variabilidade espacial e temporal da produtividade de culturas sob sistema plantio direto. *Pesquisa Agropecuária Brasileira*, 42(8), 1101-1110. <https://doi.org/10.1590/S0100-204X2007000800006>
- Atzberger, C. & Rembold, F. (2013). Mapping the spatial distribution of winter crops at sub-pixel level using AVHRR NDVI time series and neural nets. *Remote Sensing*, 5(3), 1335-1354. <https://doi.org/10.3390/rs5031335>
- Azzari, G., Jain, M. & Lobell, D. B. (2017). Towards fine resolution global maps of crop yields: Testing multiple methods and satellites in three countries. *Remote Sensing of Environment*, 202, 129-141. <https://doi.org/10.1016/j.rse.2017.04.014>
- Casado-García, Á., Heras, J. & Sanz-Sáez, A. (2019). *Google Colaboratory for Quantifying Stomata in Images*. In: International Conference on Computer Aided Systems Theory. Springer, Cham., 231-238.
- Chioderoli, C. A., Silva, R. P., Noronha, R. H. F., Cassia, M. T. & Santos, E. P. (2012). Perdas de grãos e distribuição de palha na colheita mecanizada de soja. *Revista Bragantia*, 71(1), 112-121. <https://doi.org/10.1590/S0006-87052012005000003>
- Conab - Companhia Nacional de Abastecimento. Acompanhamento da safra brasileira de grãos, v. 12 – Safra 2021/22, n. 12 – Décimo segundo levantamento, setembro de 2022. Brasília, 2022.
- Cunha, R. L. F., Silva, B. & Netto, M. A. S. (2018). A scalable machine learning system for preseason agriculture yield forecast. *Proceedings - IEEE 14th International Conference on e-Science*, e-Science 2018,

423-430.

- Gao, B. C. (1996). NdwI - A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote sensing of environment*, 58(3), 257-266. [https://doi.org/10.1016/S0034-4257\(96\)00067-3](https://doi.org/10.1016/S0034-4257(96)00067-3)
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D. & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202, 18-27. <https://doi.org/10.1016/j.rse.2017.06.031>
- Jiang, X., Fang, S., Huang, X., Liu, Y. & Guo, L. (2021). Rice mapping and growth monitoring based on time series GF-6 images and red-edge bands. *Remote Sensing*, 13(4), 579. <https://doi.org/10.3390/rs13040579>
- Jin, Z., Azzari, G. & Lobell, D. B. (2017). Improving the accuracy of satellite-based high-resolution yield estimation: A test of multiple scalable approaches. *Agricultural and Forest Meteorology*, 247, 207-220. <https://doi.org/10.1016/j.agrformet.2017.08.001>
- Johnson, D. M. (2014). An assessment of pre- and within-season remotely sensed variables for forecasting corn and soybean yields in the United States. *Remote Sensing of Environment*, 141, 116-128. <https://doi.org/10.1016/j.rse.2013.10.027>
- Kuechly, H. U., Cozacu, A., Kodl, G., Nicolai, C. & Vallentin, C. (2020) Grundlagen der Fernerkundung. Infreihe SAPIENS: Satellitendaten für Planung, Industrie, Energiewirtschaft und Naturschutz, Deutsches Geo Forschungs Zentrum GFZ. <https://doi.org/10.2312/sapiens.2020.001>
- Lobell, D. B. (2013). The use of satellite data for crop yield gap analysis. *Field Crops Research*, 43, 56-64. <https://doi.org/10.1016/j.fcr.2012.08.008>
- Maldaner, L. F. & Molin, J. P. (2019). Data processing within rows for sugarcane yield mapping. *Scientia Agricola*, 77(5), 1-8. <https://doi.org/10.1590/1678-992X-2018-0391>
- McFeeters, S. K. (1996). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing*, 17(7), 1425-1432. <https://doi.org/10.1080/01431169608948714>
- Michels, R. N., Canteri, M. G., De Aguiar, M. A., Bertozzi, J. & Dal Bosco, T. C. (2018). Reflectance as a decision-making tool for the application of fungicide to the Asian Soybean Rust. *Revista de Ciências Agrárias*, 41(3), 832-840. <https://doi.org/10.19084/RCA17333>
- Peel, M. C., Finlayson, B. L. & McMahon, T. A. (2007). Updated world map of the Köppen-Geiger climate classification. *Hydrology and Earth System Sciences*, 11(5), 1633-1644. <https://doi.org/10.5194/hess-11-1633-2007>
- Pereira, L. M., Santos, S. C. C. C., Do Nascimento, J. M. & Secretti, M. L. (2021). Correlação da produtividade da soja com os atributos químicos do solo e o NDVI. *Revista Brasileira de Agropecuária Sustentável*, 11(1), 430-438. <https://doi.org/10.21206/rbas.v11i1.12711>
- Ponzoni, F. J., Shimabukuro, Y. E.; Kuplich, T. M. (2012). *Sensoriamento remoto aplicado ao estudo da vegetação*. 2. ed. São José Dos Campos: Saraiva.
- Ponzoni, F. J., Pacheco, L. R. F., Santos, S. B. & Andrades-Filho, C. O. (2015). Caracterização espectral-temporal de dosséis de Eucalyptus spp. mediante dados radiométricos TM/Landsat5. *Cerne*, 2(2), 267-275.
- Romansini, V.A., Martins, J. A., Pivetta, L. G. & et al. (2020). *Índices de vegetação derivados de imagens orbitais como indicadores de produtividade da cultura da soja*. In: Ribeiro, J.C. (Org.). Desenvolvimento Social e Sustentável das Ciências Agrárias 2. Ponta Grossa: Atena, 2020.
- Sakamoto, T., Gitelson, A. A. & Arkebauer, T. J. (2014). Near real-time prediction of U.S. corn yields based on time-series MODIS data. *Remote Sensing of Environment*, 147, 219-231. <https://doi.org/10.1016/j.rse.2014.03.008>
- Santos, H. G., Jacomine, P. K. T., Anjos, L. H. C., Oliveira, V. A., Lumberras, J. F., Coelho, M. R., Almeida, J. A., Araujo Filho, J. C., Oliveira, J. B. & Cunha, T. J. F. (2018). *Sistema brasileiro de classificação de solos*. Brasília, DF: Embrapa.
- Schirmbeck, L. W., Fontana, D. C., Dalmago, G. A., Schirmbeck, J., Vargas, P. R. & Fernandes, J. M. C. (2020). Condições hídricas de lavoura de soja usando sensoriamento remoto terrestre. *Agrometeoros*, 27(1), 173-181. <http://dx.doi.org/10.31062/agrom.v27i1.26565>

- Shao, Y., Campbell, J. B., Taff, G. N. & Zheng, B. (2015). An analysis of cropland mask choice and ancillary data for annual corn yield forecasting using MODIS data. *International Journal of Applied Earth Observation and Geoinformation*, 38, 78-87. <https://doi.org/10.1016/j.jag.2014.12.017>
- Sharma, L. K., Bu, H., Denton, A. & Franzen, D. W. (2015). Active-optical sensors using red NDVI compared to red edge NDVI for Prediction of corn grain yield in North Dakota, USA. *Sensors*, 15, 27832–27853. <https://doi.org/10.3390/s151127832>
- Shelestov, A.; Lavreniuk, M.; Kussul, N.; Novikov, A.; Skakun, S. (2017). Exploring Google Earth Engine platform for big data processing: Classification of multi-temporal satellite imagery for crop mapping. *Frontiers in Earth Science*, 5(17), 1-10. <https://doi.org/10.3389/feart.2017.00017>
- Trindade, F. S., De Carvalho Alves, M., Noetzold, R., De Andrade, I. C. & Pozza, A. A. A. (2019). Relação espectro-temporal de índices de vegetação com atributos do solo e produtividade da soja. *Revista de Ciências Agrárias*, 62, 1-11. <http://200.129.150.26/index.php/ajaes/article/view/2928>
- Zanzarini, F. V., Pissarra, T. C., Brandão, F. J. & Teixeira, D. D. (2013). Correlação espacial do índice de vegetação (NDVI) de imagem Landsat/ETM+ com atributos do solo. *Revista Brasileira de Engenharia Agrícola e Ambiental*, 17, 608-614. <https://doi.org/10.1590/S1415-43662013000600006>
- Zarei, A. R., Mahmoudi, M. R. & Moghimi, M. M. (2022). Determining the most appropriate drought index using the random forest algorithm with an emphasis on agricultural drought. *Natural Hazards*, 1-24. <https://doi.org/10.1007/s11069-022-05579-2>

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).