REMOTE SENSING FOR IDENTIFYING SOYBEAN CULTIVARS AND ESTIMATING CROP YIELDS

Bárbara Carolline Matwijou[1*], José Dimas Rodrigues Oliveira[1], Marco Antônio da Silva[1], Gabriel Geovane Scheidt[1], Armando Lopes de Brito Filho[2], Maura Gabriela da Silva Brochado[1], Franciele Morlin Carneiro[1]

[1]Federal Technological University of Paraná (UTFPR), Santa Helena *campus*, Paraná State, Brazil.
[2]Western Paraná State University (UNIOESTE), Cascavel *campus*, Paraná State, Brazil.
*e-mail: barbaram@alunos.utfpr.edu.br

Abstract.

Traditional methods for cultivar identification and agricultural productivity estimation have been increasingly complemented or replaced by innovative approaches using geotechnologies and Artificial Intelligence (AI). These modern techniques offer greater efficiency, speed, and sustainability in agricultural production systems. Among these advances, Remote Sensing (RS) has stood out as an effective tool for agricultural monitoring. It allows the collection of spectral and biophysical information about crops in a non-destructive and large-scale manner. This study aimed to identify soybean cultivars and estimate their productivity by integrating RS and AI techniques. It was conducted in an experimental area of the Federal Technological University of Paraná (UTFPR), located on the Santa Helena campus, in the state of Paraná, Brazil. In total, 20 soybean cultivars were evaluated in nine replicates, resulting in a dataset of 180 experimental samples. The analyses focused on biophysical crop variables and vegetation indices derived from proximal remote sensing using the Rapidscan active optical sensor and the CS-45 model. They also used orbital remote sensing from the PlanetScope satellite constellation, which provides 3-meter-resolution imagery. The imagery includes spectral bands (red, green, blue, yellow, red edge, and NIR) and vegetation indices, with a focus on NDVI (Normalized Difference Vegetation Index) and NDRE (Normalized Difference Red Edge Index). To identify the cultivars, we used Pearson's correlation coefficient, hierarchical clustering, and a dendrogram to evaluate the interactions between vegetation indices, spectral bands, and their relationships with soybean cultivation. The dataset was divided into 70% for model training and 30% for validation, with soybean productivity as the output variable and biophysical variables, vegetation indices, and spectral bands as input variables. The machine learning algorithm used was the Random Forest Regressor. The results indicated strong spectral interactions between the evaluated cultivars. The best-performing model achieved an accuracy of 80.49% using biophysical variables associated with the NDRE index derived from proximal remote sensing. In contrast, the orbital remote sensing analysis showed an accuracy of 77.05% when using spectral bands. This study concludes that the integration of remote sensing and artificial intelligence is an effective strategy for identifying cultivars and estimating soybean productivity, thus promoting precision agriculture.

Keywords.

Artificial Intelligence, Precision Agriculture, Vegetation indices.

Introduction

Soybean (*Glycine max* (L.) Merrill) productivity is influenced by biotic and abiotic factors, including water availability, solar radiation, soil compaction, insect pests and diseases, plant nutritional status, and farm management practices (Felema et al., 2013). Given this complexity, Precision Agriculture is an important strategy for increasing the productive efficiency and sustainability of agricultural systems by managing spatial and temporal variability across production areas (Brasil, 2012). Remote Sensing is a tool that provides information on vegetation without direct contact with the plants, using various data acquisition platforms (Shiratsuchi et al., 2014). Despite advances in integrating Precision Agriculture, Remote Sensing, and Artificial Intelligence, challenges remain in the automated identification of cultivars under real field conditions. Therefore, the present study aimed to identify soybean cultivars, estimate their agricultural productivity by integrating remote sensing data and Artificial Intelligence techniques, and evaluate and compare the performance of predictive models generated from orbital and proximal remote sensing data.

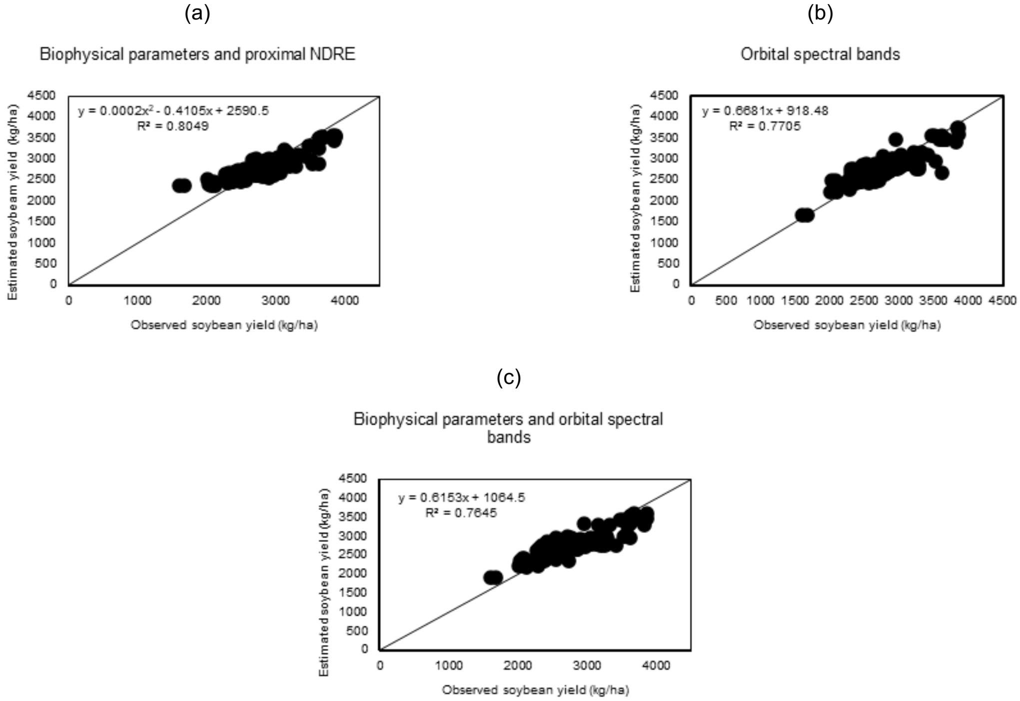
Material e Methods

The study was conducted in the experimental area of the Federal Technological University of Paraná (UTFPR), Santa Helena *campus*, located at coordinates 24°50'55.22'' S and 54°20'52.83'' E, at an altitude of 239 meters, on Red Latossolo (IBGE, 2025) and in a Cfa climate according to the Köppen-Geiger classification (Beck et al., 2018). The experimental design adopted was a completely randomized block design with 20 soybean cultivars and 9 replicates, for a total of 180 experimental units. Crop biophysical variables were collected, including canopy area, plant height, canopy width, chlorophyll a and b, and yield. Proximal remote sensing was performed using the CS-45 active optical sensor to obtain the NDVI and NDRE indices. Orbital remote sensing data were obtained from the PlanetScope constellation, covering spectral bands of blue, green, red, yellow, red edge, and near-infrared. Initially, Pearson's correlation coefficient was applied to verify the relationship between cultivars and spectral variables. Subsequently, the Random Forest Regressor algorithm was used to estimate productivity. The dataset was split into 70% for training and 30% for validation, and model performance was evaluated using the following metrics: Coefficient of Determination (R^2), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE).

Results and Discussion

The best-performing predictive models are shown in Fig 1. The scenario comprising biophysical parameters and the NDRE index obtained through near-surface remote sensing yielded the best result, with $R^2 = 0.8049$ (Fig 1a). For orbital remote sensing, the exclusive use of spectral bands yielded $R^2 = 0.7705$ (Fig 1b), whereas the combination of spectral bands and biophysical parameters yielded $R^2 = 0.7645$ (Fig 1c). The results demonstrate that integrating remote sensing and machine learning algorithms is an efficient strategy for estimating soybean yield, enabling the identification of relevant agronomic characteristics and demonstrating potential for application in sustainable agricultural systems (Gava et al., 2022).

Fig. 1. Performance of validated predictive models for estimating soybean yield under optimal conditions: (a) biophysical parameters and proximal NDRE, (b) orbital spectral bands, and (c) biophysical parameters associated with spectral bands.



Conclusion

The integration of remote sensing and artificial intelligence techniques constitutes an effective approach for identifying soybean cultivars and estimating agricultural productivity. The best performance was achieved by combining biophysical parameters with NDRE indices derived from near-infrared remote sensing, yielding $R^2 = 0.8049$. The results also demonstrated that orbital remote sensing performs satisfactorily for large-scale applications, especially when using spectral bands exclusively. Thus, both approaches show high potential for applications in agricultural monitoring systems and decision support.

References

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., & Wood, E. F. (2018). Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Scientific Data, 5*, 180214.

BRASIL. (2012). Ministério da Agricultura, Pecuária e Abastecimento. Projeções do 293 Agronegócio 2011/2012 a 2021/2022. Brasília. 50 p.

Felema, J., Raiher, A. P., & Ferreira, C. R. (2013). Agropecuária brasileira: desempenho regional e determinantes de produtividade. *Revista de Economia e Sociologia Rural*, *51*, 555–573.

IBGE - Instituto Brasileiro de Geografia e Estatística. (2025). *Banco de dados e informações ambientais (BDiA)*. IBGE. https://www.ibge.gov.br/geociencias/informacoes-ambientais/geologia/23382-banco-de-informacoes-ambientais.html.

Shiratsuchi, L. S., Brandão, Z. N., Vicente, L. E., Victoria, D. C., Ducati, J. R., Oliveira, R. P., & Vilela, M. F. (2014). Sensoriamento remoto: conceitos básicos e aplicações na agricultura de precisão. In A. C. C. Bernardi, J. M. Naime, Á. V. Resende, L. H. Bassoi, & R. Y. Inamasu (Eds.), *Agricultura de precisão: resultados de um novo olhar* (2nd ed., pp. 132–159). Embrapa.

Gava, R., Santana, D. C., Cotrim, M. F., Rossi, F. S. Teodoro, L. P. R., Silva Júnior, C. A. da, & Teodoro, P. E. (2022). Soybean Cultivars Identification Using Remotely Sensed Image 307 and Machine Learning Models. *Sustainability*, *14*, 7125–7125.