Oliveira, Gustavo Satyro B. de
Luciano, Ana Claúdia S.

Vegetation index as a tool for monitoring soybean yield in Center-South region of the state of Parana - Brazil

https://doi.org/10.34628/9WKK-FF88
http://repositorio.ulusiada.pt/handle/11067/7378

Metadados

Data de Publicação 2023

Resumo Crop monitoring is an important tool for agricultural management. Soybean is one of the most important crops in Brazil because it can be used as food and has an important role in the economy, especially in the Center-South region. The most common methods to monitor crop yield in Brazil are based on interviews and questionnaires made by the official organizations and answered by the producers. As an alternative to monitoring large areas, remote sensing data has shown the potential to improve crop...

Tipo bookPart

Editora Universidade Lusíada Editora

Vegetation index as a tool for monitoring soybean yield in center – South region of the state of Parana – Brazil

DOI: https://doi.org/10.34628/9WKK-FF88

Gustavo Satyro B. de Oliveira[1], Ana Cláudia S. Luciano[2]

[1] Escola Superior de Agricultura Luiz de Queiroz
13416-030
gustavo.bacelar2010@usp.br

[2] Escola Superior de Agricultura Luiz de Queiroz
13418-260
analuciano@usp.br

Abstract. Crop monitoring is an important tool for agricultural management. Soybean is one of the most important crops in Brazil because it can be used as food and has an important role in the economy, especially in the Center-South region. The most common methods to monitor crop yield in Brazil are based on interviews and questionnaires made by the official organizations and answered by the producers. As an alternative to monitoring large areas, remote sensing data has shown the potential to improve crop yield monitoring. The objective of this study was to evaluate the applicability of using vegetation indices to monitor soybean yield in the Center-South mesoregions in the state of Parana. For this, we used Sentinel-2 imagery from 2019 and 2020 to calculate vegetation indices (NDVI, NDWI, EVI, and SAVI) and the soybean yield from the Brazilian Institute of Geography and Statistics (IBGE). The vegetation indices and soybean yield were correlated using regression analyses. The results were compared with the traditional methods generated by IBGE. Based on the vegetations index used, the EVI and NDWI had the best correlation with the soybean yield. Studies using more extensive time series analyses in different regions are necessary to evaluate which of the vegetation indices are better for monitoring soybean yield to achieve sustainable crop management and production.

Keywords: Remote sensing; Crop monitoring; Yield forecasting; Crop management.

1. Introduction

In Brazil, the beginning of soybean cultivation was in 1882, in the state of Bahia [1]. However, soybean only established itself more definitively in the country in 1914 in Rio Grande do Sul, when the varieties adapted to the country's climatic conditions, mainly with questions related to the photoperiod [2]. Currently, soybean is very important for the Brazilian economy, since the export of the commodity easily exceeds R$ 10 billion annually [3]. In addition, the benefits caused by soybean are even greater, since for one out of every four dollars exported by the country are, in some way, linked to the production chain of the crop [3].

In the Brazilian Center-West region, the soybean production is most significant, where the mainly producers are the states of Mato-Grosso, Goiás and Mato Grosso do Sul, which were responsible for around 49.2% of the production of cereals, oilseeds and legumes in the harvest 2022 [4]. In addition, during the same harvest, it was expected that the produced volume of this commodity will be approximately 118.8 million tons [4].

In Brazil, the monitoring of soybean production is officially carried out through the National Supply Company (CONAB) and the Brazilian Institute of Geography and Statistics (IGBE). However, most of these analyses occur through questionnaires and interviews with the producers that can result, under certain circumstances, in a subjectivity analysis of crop production and productivity. On the other hand, it is important to highlight that remote sensing technologies have been used as a monitoring tool, but still on a smaller scale for crop yield estimation. Therefore, it would be a useful method to monitor soybean yield in an efficient and reliable way to improve the currently used by official organizations. In this context, remote sensing data as satellite images is an important tool to facilitate the monitoring of agricultural areas, such as area identification, yield estimation and agricultural management [5].

The use of satellite images and vegetation indices is a great alternative to the current way of information collecting in large areas, especially in crop production such as soybean. For example, the data from satellite sensors are taken in relatively short periods of time, some being taken daily, which would allow constant monitoring of the crop properties. Additionally, it is possible to monitor large areas, which would make the process of obtaining data more efficient, since it would no longer be carried out by farmers, or some institutions that compiles the data from groups. In addition, the use of vegetation indices is an important variable to monitor soybean areas and could be used as a tool to estimate soybean yield, with relatively high efficiency.

Andrade et al. (2022) [6] used Sentinel-2 and Landsat-8 images, to calculate the NDVI index (Normalized Difference Vegetation Index), in order to propose a linear regression model to estimate soybean production on a farm in the city of *Monte Alegre do Piauí - Brazil*. The authors concluded that the model was more efficient for soybeans in the phenological stages V5, V6 and R2, that is, phases corresponding to vegetative 5, vegetative 6 and reproductive 2 (full flowering). The mean prediction error was 153.9 kg/ha, which corresponds to 4.2% when compared to the data generated with plants obtained in the field. Other studies used vegetation indices and achieved good results in estimating the soybean yield, with R^2 from Monteiro (2019) [10] to Skakun *et al*. (2021) [11].

The monitoring of soybean productivity using remote sensing data and vegetation indices is extremely important to have a sustainable agricultural production and therefore to ensure food security. Moreover, the monitoring of soybeans areas using remote sensing data brings advantages for the agribusiness since it is possible to control the area, yield and management of the crop according to the laws and government policies ensuring environmental sustainability.

2. Methodology

The monitoring of soybean yield consisted of analyzing the relationship between soybean yield and vegetation indices. For this, first were selected the areas of soybean production in the study region. After that were calculated the vegetation indices using images from the Sentinel-2 satellite. The vegetation indices were evaluated based on the correlation with the soybean productivity at a municipal level and by the rainfall variation in the region.

2.1. Dataset analyses

2.1.1. Study area

The study area is in the Central-Southern mesoregion of Paraná (Figure 1), with approximately 2,638,104.8 ha representing 13.20% of the entire extension of the state [7]. In addition, it comprises 29 counties, of which Palmas and Guarapuava stand out due to their economic and population numbers.

Figure 1. Location of the study area in the central-southern region of Paraná. Source: (IPARDES, 2004)

2.1.2. Soybean yield from IBGE

The soybean productivity was obtained from the IBGE platform Automatic Recovery System (SIDRA) by the Systematic Survey of Agricultural Production - LSPA [8]. The LSPA provides estimates on planted areas, harvested areas, amount produced and average yield of crops. The data are available by municipalities. Therefore, files were downloaded in excel format (csv.) of productivity, referring to the years 2019 and 2020 for the cities of the mesoregion.

2.1.3. Soyabean maps

The land use map of the Central-Southern mesoregion of Paraná was used to identify the soybean areas. For this, it used the Map Biomas platform, which is a collaborative network, formed by a group of universities, ONGs and technology startups, producing annually, among several other projects, a mapping of land use coverage throughout the Brazilian territory [9]. For this study, land use data were downloaded for the south-central mesoregion of Paraná between the years 2019 and 2020, using the Google Earth Engine platform. After that, it was selected just the areas within the soybean crop, classified according to the Mapbiomas project.

2.1.4. Rainfall data

In addition to the vegetation indices, it was also necessary to use the rainfall data to understand the productivity data, since the climatic data has an important influence in the crop development. In this study, rainfall data was obtained from the database of the Paraná Water Institute [18]. The data obtained were the average of October, November and December between 2005 and 2019, for each municipality in study areas. It's important to say that it was not possible to obtain accurate data from the municipalities of Campina do Simão, Cantagalo, Foz do Jordão, Mato Rico and Porto Barreiro.

2.1.5. Sentinel-2 imagery

We used images from Sentinel-2 satellite, a sensor created by the European Space Agency, which has a 10 meters of spatial resolution, along the objective of acquiring multispectral and high-resolution images of land use across the planet. The imagery processing was executed in Google Earth Engine platform, in a way that only images with less than 10% of cloud cover were considered. We selected images between the months of October and December of the years 2019 and 2020, according to the soybean development.

The vegetation indices calculated were NDVI [12], NDWI [13], EVI [14] and SAVI [15]. These indices were chosen, primarily because they're often the most used ones to make analysis regarding agricultural aspects and they are used to evaluate biophysical information from crops [19]. After calculating the indices, the average values of NDVI, NDWI, EVI and SAVI were determined for the selected period, from October to December, of the year 2019, for the regions that produced soybeans. For each municipality in the mesoregion, the average values of vegetation indices were calculated in each month of study and analysis. With the values of the vegetation indices, analyses were carried out at a municipal level in relation to soybean productivity and the values of vegetation indices, through correlation analysis, considering the coefficient of determination (R^2).

3. Results and discussion

It was observed that in most cases, the municipalities that had the highest average NDVI values from October to December also had higher productivity (Figure 2). The coefficient of determination (R^2) using NDVI showed a value equal to 0,29. In general, the highest values of NDVI were associated with the higher yields.

Also, it was possible to infer that this increase in productivity in the commented municipalities was due to the higher rainfall between 2005 and 2019 in the study area (Figure 3). However, it is important to consider the accumulated rainfall during the period of analyses, since the crop yield is influenced by climatic conditions considering the harvest cycle.

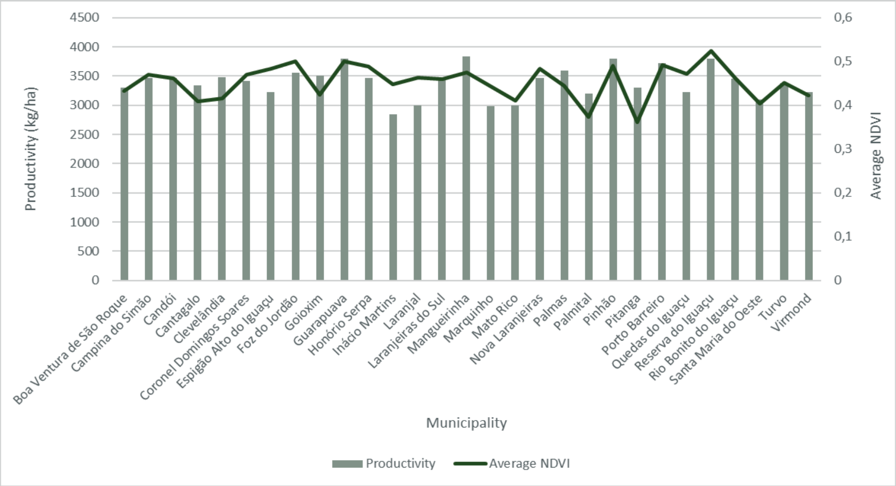

Figure 2: Comparison between productivity and average NDVI.

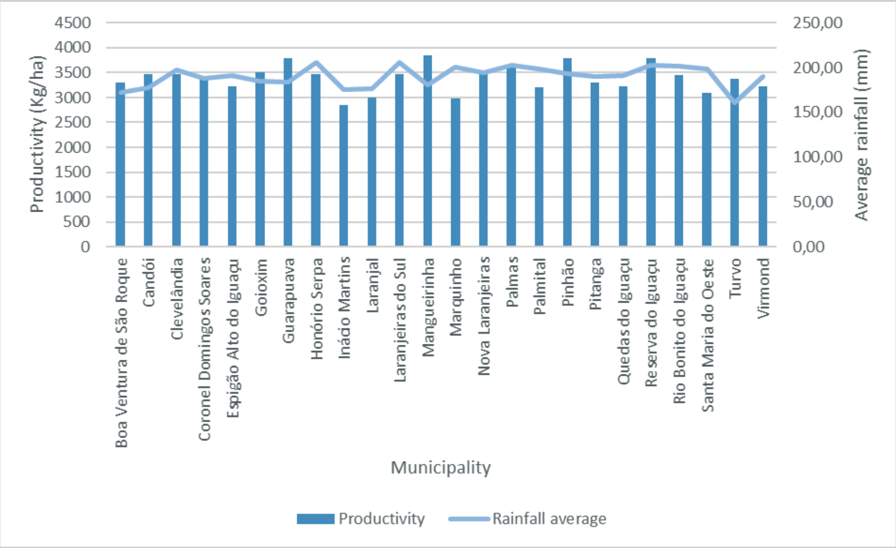

Figure 3: Comparison between rainfall and productivity.

Using the NDWI index it was possible to identify that for the lowest values of NDWI the higher productivity was identified in the municipalities (Figure 4). Such behavior is expected, as this vegetation index is calculated from the green and near-infrared reflectance values, regions of the spectrum where the plant responds to photosynthetic activity and cell structure and development, which are directly affected by water and dry matter values (which vary throughout the crop cycle) [16]. Therefore, during the months in which the study was carried out, soybeans are approaching harvest in the region, therefore, it is natural that the reflectance during this period is greater in the near-infrared. These data could also be confirmed through the R^2, with a value of 0.32, therefore a greater correlation was obtained for NDWI when compared with NDVI to correlate with soybean yield.

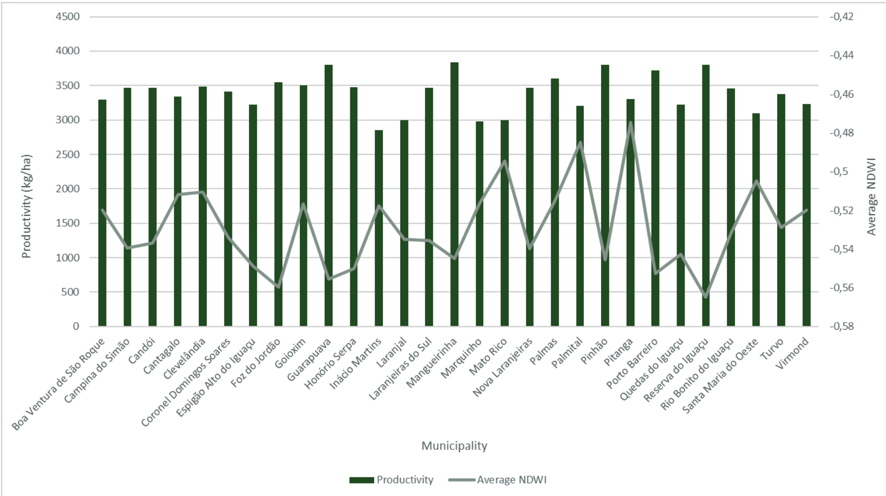

Figure 4: Comparison between productivity and average NDWI values.

The results for the SAVI index showed that the highest SAVI values are associated, in most cases, with the highest values of productivity (Figure 5). It is also interesting to mention that this behavior is very similar to NDVI, a fact that makes a lot of sense. This, considering that the SAVI was created from the NDVI, however having an adjustment of the ground factor, in the equation called L, to correct its interference, especially in places where vegetation cover is low [17]. This line of reasoning can be further corroborated by observing the R^2 value equal to 0.29, a value similar to the coefficient generated by the NDVI ($R^2 = 0,29$).

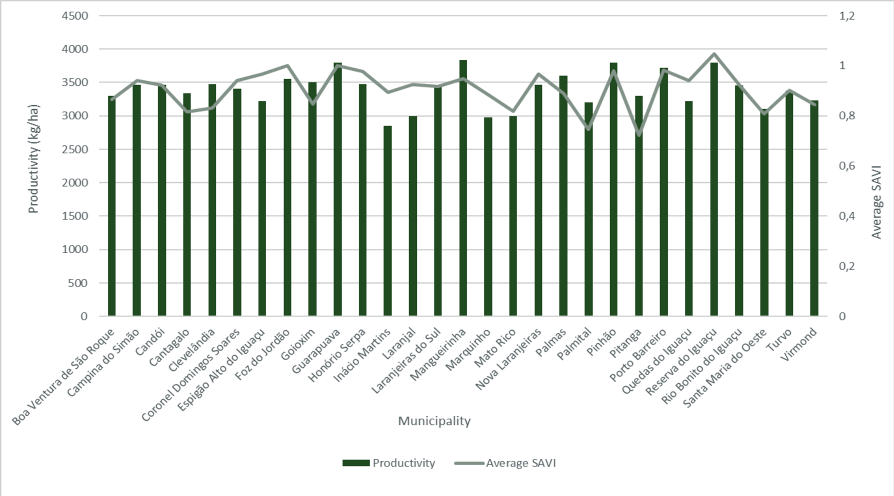

Figure 5: Comparison between productivity and average SAVI values.

For the EVI index, a greater correlation was found between productivity and index values concerning NDVI and NDWI (R^2 equal to 0.45). Therefore, for the majority of municipalities, the highest EVI values are associated with the highest soybean productivity at the municipality level, as shown in Figure 6. The results were more accurate than the other two indices used. This is because the EVI, despite being similar to the NDVI, presents atmospheric and soil correction factors. The dispersion of productivity and EVI data could also be observed, in which the R^2 was equal to 0.45 (Figure 7). This is the highest correlation between productivity and vegetation index value in this study.

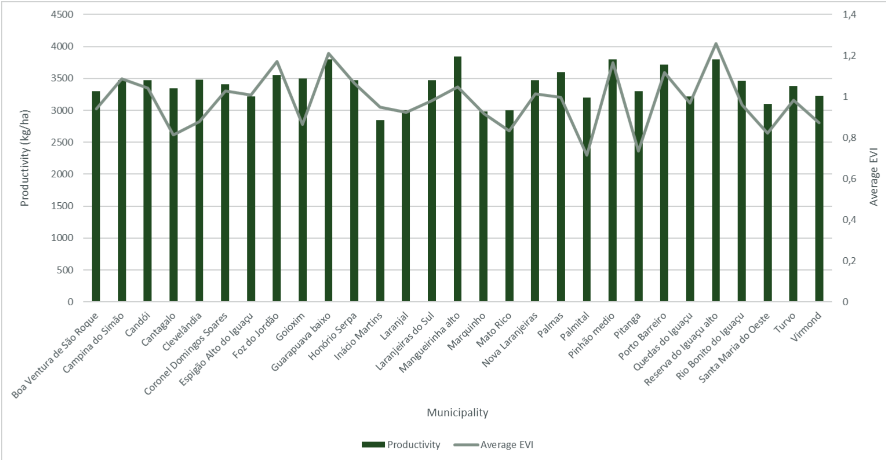

Figure 6: Comparison between productivity and average EVI values.

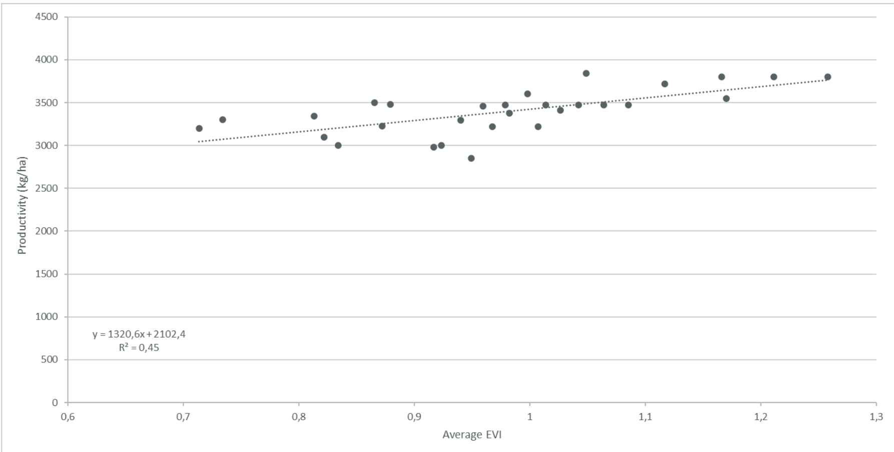

Figure 7. Scatter plot comparing productivity values with average EVI values.

4. Conclusion

The use of vegetation indices to monitor the soybean productivity is an important factor with great potential to forecast crop yield. The EVI was the index that presented the best correlation with productivity, with R^2 equal to 0.45. Other indices, such as NDWI and NDVI, also showed great potential in monitoring the productivity of soybean, showing the importance of integrating and using more than one vegetation index to evaluate crop productivity. However, it's necessary to emphasize the importance of climate monitoring and management for obtaining even more satisfactory results. Finally, the vegetation indices can be an important tool to help the monitoring of food security and sustainability of the crop production.

References

1. EMBRAPA. (1987). *A SOJA NO BRASIL: História e Estatística*. COMITÊ DE PUBLICAÇÕES DO CNPSo. https://www.infoteca.cnptia.embrapa.br/bitstream/doc/446431/1/Doc21.pdf (Origanal published in 1987)
2. Freitas, M. (2011). A cultura da soja no brasil: O crescimento da produção brasileira e o surgimento de uma nova fronteira agrícola. In *Enciclopédia biosfera* (7ª ed., pp. 1–12). Recovered of https://www.conhecer.org.br/enciclop/2011a/agrarias/a%20cultura%20da%20soja.pdf
3. Dall'agno, A. (s.d.). *Importância socioeconômica da soja.* Agência CNpta Embrapa. https://www.agencia.cnptia.embrapa.br/gestor/soja/arvore/CONTAG01_12_271020069131.html.
4. Estatísticas Econômicas. (2022, 11 of august). *Em julho, IBGE prevê safra recorde de 263,4 milhões de toneladas para 2022 | Agência de Notícias*. Agência de Notícias - IBGE. https://agenciadenoticias.ibge.gov.br/agencia-sala-de-imprensa/2013-agencia-de-noticias/releases/34626-em-julho-ibge-preve-safra-recorde-de-263-4-milhoes-de-toneladas-para-2022
5. JOHNSON, D. M et. al. (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, v. 141, p. 116–128.
6. Andrade, T. G., Andrade Junior, A. S. D., Souza, M. O., Lopes, J. W. B., & Vieira, P. F. D. M. J. (2022). Soybean yield prediction using remote sensing in southwestern piauí state, brazil. *Revista Caatinga*, *35*(1), 105–116. https://doi.org/10.1590/1983-21252022v35n111rc
7. IPARDES. (2004). *Leituras regionais: Messoregião geográfica centro-sul paranaense*. http://www.ipardes.gov.br/biblioteca/docs/leituras_reg_meso_centro_sul.pdf
8. IBGE. (2016, 6 of december). *Novo SIDRA permite consultar facilmente dados de estudos e pesquisas do IBGE também em dispositivos móveis | Agência de Notícias*. Agência de Notícias - IBGE. https://agenciadenoticias.ibge.gov.br/agencia-sala-de-imprensa/2013-agencia-de-noticias/releases/9481-novo-sidra-permite-consultar-facilmente-dados-de-estudos-e-pesquisas-do-ibge-tambem-em-dispositivos-moveis#:~:text=O%20SIDRA%20permite%20consultar%20informações,PIB)%20e%20suas%20séries%20históricas.

9. MAPBIOMAS. (2022). *Mapbiomas Brasil | O projeto*. Mapbiomas Brasil. https://mapbiomas.org/o-projeto
10. Monteiro, P. (2021). *Comportamento espectrotemporal da soja utilizando sensores orbitais e não orbital e correlação dos índices de vegetação com a produtividade* [Doctoral dissertation, Universidade Tecnológica do Paraná]. https://repositorio.utfpr.edu.br/jspui/bitstream/1/24912/1/comportamentoespectrotemporasoja.pdf
11. Skakun, S., Kalecinski, N. I., Brown, M. G. L., Johnson, D. M., Vermote, E. F., Roger, J.-C., & Franch, B. (2021). Assessing within-field corn and soybean yield variability from worldview-3, planet, sentinel-2, and landsat 8 satellite imagery. *Remote Sensing*, *13*(5), 872. https://doi.org/10.3390/rs13050872
12. Rouse, J. W. (1973). *Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation*. IDB - Index DataBase. https://www.indexdatabase.de/db/r-single.php?id=675
13. 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
14. Huete, A., Justice, C., & Leeuwen, W. (1999). *Modis vegetation index (mod 13)*. MODIS (Moderate Resolution Imaging Spectroradiometer). https://modis.gsfc.nasa.gov/data/atbd/atbd_mod13.pdf
15. Rondeaux, G., Steven, M., & Baret, F. (1996). Optimization of soil-adjusted vegetation indices. *Remote Sensing of Environment*, *55*(2), 95–107. https://doi.org/10.1016/0034-4257(95)00186-7
16. Castro, B., Sartori, A., & Moraes, D. (2017). Índices de vegetação (IV's) correlacionados com parâmetros do solo em área de cana-de-açúcar. (pp. 4104-4111). Universidade do Sagrado Coração. https://proceedings.science/proceedings/59/_papers/59943/download/abstract_file1
17. United States Geological Survey (USGS). (2023). *Landsat soil adjusted vegetation index*. USGS. https://www.usgs.gov/landsat-missions/landsat-soil-adjusted-vegetation-index
18. Paraná Water Institute. (2023). *Relatório de Alturas Mensais de Precipitação*. http://www.sih-web.aguasparana.pr.gov.br/sih-web/gerarRelatorioAlturasMensaisPrecipitacao.do?action=carregarInterfaceInicial
19. Thenkabail, P. S., Smith, R. B., & De Pauw, E. (2000). Hyperspectral Vegetation Indices and Their Relationships with Agricultural Crop Characteristics. *Remote Sensing of Environment*, *71*(2), 158–182. https://doi.org/10.1016/s0034-4257(99)00067-x