

## MODELO DE ANÁLISE ESPECTRAL USANDO FERRAMENTAS COMPUTACIONAIS PARA IDENTIFICAÇÃO DE DATAS DE PLANTIO E COLHEITA DE SAFRAS AGRÍCOLAS NO BRASIL

*Gustavo Eduardo Marcatti<sup>1</sup>, Marconi de Arruda Pereira<sup>2</sup>, José Rogério de Almeida Júnior<sup>2</sup>, Filipe Santos<sup>2</sup>, Pedro Arthur de Azevedo Silva<sup>1</sup>, André Hirsch<sup>1</sup>, Andréa Cristina Rodrigues dos Santos<sup>1</sup>, Fernando Martins Pimenta<sup>1</sup>, Glauciana da Mata Ataíde<sup>4</sup>, Gabriel Pereira<sup>3</sup>, Renato Vinícius Oliveira Castro<sup>4</sup>, Marcela Costa Ferreira<sup>1</sup>, Leonardo Brandão do Prado<sup>3</sup>, Fernando Augusto Teixeira<sup>2</sup>, Webert Saturnino Pinto<sup>1</sup>, Daniel Calbino Pinheiro<sup>1</sup>, Ivair Gomes<sup>3</sup>*

<sup>1</sup> Departamento de Ciências Agrárias – Universidade Federal de São João Del Rei (UFSJ) Caixa Postal 36307-352–SeteLagoas–MG–Brasil {gustavomarcatti@ufsj.edu.br, hirsch\_andre@ufsj.edu.br, Andreacr2010@gmail.com, fernando.m.pimenta@gmail.com, mcferreira.edu@gmail.com, webert.saturnino@gmail.com, dcalbino@ufsj.edu.br};

<sup>2</sup> Departamento de Tecnologias em Engenharia Civil, Computação e Humanidades - Universidade Federal de São João Del Rei (UFSJ) Caixa Postal: 36490-972–Ouro Branco – MG – Brasil {marconi@ufsj.edu.br, filipesantos.lf@gmail.com, teixeira@ufsj.edu.br};

<sup>3</sup> Departamento de Geociências - Universidade Federal de São João Del Rei (UFSJ) Caixa Postal: 35701-970 – São João del Rei– MG – Brasil { pereira@ufsj.edu.br, leonardobprado@gmail.com, ivair@ufsj.edu.br};

<sup>4</sup> Departamento de Engenharia Florestal - Universidade Federal de São João Del Rei (UFSJ) Caixa Postal: 35701-970 – São João del Rei– MG – Brasil {glauciana@ufsj.edu.br, renatocastro@ufsj.edu.br}

### RESUMO

Na busca pela automação dos processos e de algoritmos mais eficientes para auxiliar na tomada de decisões, métodos computacionais tornam-se ferramentas cruciais para o desenvolvimento de diversas áreas das ciências. No presente artigo apresentamos uma metodologia de estudo da análise espectral para localizar datas prováveis de início e fim dos ciclos de culturas, com o auxílio da base de dados do satélite MODIS - MOD13Q1. Observou-se que a ferramenta avaliada se mostrou eficiente, porém passível de um refinamento na definição do comportamento por regiões brasileiras, na obtenção da data de plantio e nos pontos críticos da curva de evolução das culturas.

**Palavras-chave** — Sensoriamento Remoto, Machine Learning, Milho 2ª Safra.

### ABSTRACT

*In the search for the automation of processes and more efficient algorithms to assist in decision making, computational methods become crucial tools for the development of several areas of science. In the present article We present a methodology for studying spectral analysis to locate probable dates for the beginning and end of crop cycles, with the aid of the MODIS - MOD13Q1 satellite database. It was observed that the evaluated tool proved to be efficient, but subject to refinement in the definition of behavior by Brazilian regions, in obtaining the planting date and in the critical points of the crop evolution curve.*

**Keywords** — Remote Sensing, Machine Learning, Second Crop Corn.

### 1. INTRODUÇÃO

Diante do aumento exponencial da população humana, a qual deve alcançar aproximadamente 8 bilhões de pessoas em 2026, a humanidade exercerá uma pressão ainda mais incisiva na produção de alimentos que deverá duplicar nos próximos 30 anos para atender a demanda populacional. Como as terras agricultáveis são limitadas por fatores naturais e biogeográficos, o aumento da produção de alimentos está diretamente relacionado com a necessidade do aumento da produtividade nas mesmas áreas já usadas pela agropecuária, e da recuperação das lavouras e pastagens degradadas em razão do manejo inadequado de culturas agrícolas e/ou de técnicas de manejo e conservação do solo mal aplicadas [1,2,3,4,5,6,7,8].

Na Agricultura 4.0 as imagens de satélite permitem obter uma análise detalhada sobre cobertura vegetal e uso da terra, tipo de solo, topografia, áreas degradadas ou culturas agrícolas afetadas por alguma doença. Entre as aplicações que são possíveis de identificar e que propiciam o aumento da produção podemos citar: estimativa de área plantada; vigor vegetativo das culturas; previsão de safras agrícolas; regiões com maior potencial de produção; determinação de faixas de solo com baixa produtividade; identificação de áreas nativas que devem ser destinadas para preservação (APP e Reserva Legal); áreas destinadas para silvicultura; e identificação de erros em processos de adubação, irrigação e preparo do solo [9].

Atualmente, a tarefa de obtenção de dados de sensoriamento remoto em tempo real é possível através da enorme constelação de satélites em órbita. No total são mais de 20 mil satélites orbitando a Terra simultaneamente, dos quais cerca de uma centena está dedicada ao mapeamento da cobertura terrestre. As imagens geradas possuem uma

resolução espacial que variam de 2 km a menos de 1 m, como os satélites MODIS (EUA - NASA, 2 km a 250 m), Landsat (EUA - NASA, 30 m a 15 m), CBERS (Brasil - China, INPE - CRESDA, 260 m a 2.7 m), Spot (França - ESA, 20 m a 2 m), Ikonos (EUA - DigitalGlobe, 4 m a 1 m), WorldView (EUA - DigitalGlobe, 1.24 m a 31 cm).

Em termos de produção de alimentos a nível global, a soja, o milho e o trigo podem ser consideradas as culturas agrícolas mais importantes para a humanidade. Em muitos países, as commodities fornecidas pelo setor agrícola desempenham um papel importante na economia. Garantir a alimentação da população humana é uma destas metas, a qual pode ser alcançada quando tomadores de decisão são apoiados por ferramentas acessíveis que fornecem respostas rápidas e precisas com alta resolução temporal e cobertura global. Isto vem incentivado os tomadores de decisão a lançar mão destas novas tecnologias da área do sensoriamento remoto [10].

Assim, monitorar a cobertura vegetal e o uso da terra no Brasil é uma ferramenta importante na análise do estágio de desenvolvimento das culturas agrícolas, bem como a sua dinâmica temporal e espacial na paisagem. Além disso, a utilização de imagens multiespectrais auxilia no planejamento agrícola regional, como no caso do monitoramento de propriedades rurais que fazem parte do Programa de Subvenção ao Prêmio do Seguro Rural - PSR, diante da importância que o setor agrícola possui na produção de *commodities* brasileira, o presente trabalho tem como objetivo central avaliar um método de identificação e validação das culturas, neste trabalho, o milho, via assinatura espectral no Mato Grosso do Sul - MS no ano de 2020.

### 1.1. Assinatura espectral

Uma das principais ferramentas para identificar culturas agrícolas são a curva de reflectância, assinatura espectral, e a análise de séries temporais de valores de índices vegetacionais, como no caso do NDVI, entre outros [11,12,13].

Uma assinatura espectral, também chamada de curva de reflectância, pode ser definida como a resposta padrão que uma determinada classe de mapeamento (conjunto de pixels de mesma tonalidade) apresenta ao longo do espectro eletromagnético (bandas) em termos de quantidade de energia refletida (classes de reflectância), captada por um sensor a bordo de um satélite, avião, drone ou VANT [14,13].

Além disso, essas mesmas propriedades da interação da luz do Sol com a vegetação nativa ou culturas agrícolas podem ser aproveitadas em trabalhos ao nível do solo com o emprego de espectrorradiômetros portáteis [14,12].

Desta forma, a análise temporal de séries de culturas agrícolas é capaz de proporcionar parâmetros de sazonalidade de cada um dos anos. Entretanto, a presença de erros inerentes ao processo de imageamento orbital, assim como, a variabilidade da geometria de aquisição do sensor e sua influência na iluminação solar e sombreamento de culturas tem implicações diretas nas estimativas e ajuste das séries temporais, sendo necessária a remoção de dados discrepantes (outliers).

## 2. MATERIAL E MÉTODOS

O presente trabalho utilizou as apólices do banco de dados do Sistema de Subvenção Econômica - PSR para entrecruzar os dados relacionados com a cultura declarada e a resposta espectral padrão esperado para a cultura, com filtros relacionados à área superior a (72 ha) e a seleção das culturas investigadas (Milho 2ª safra).

Para definir o comportamento padrão das respostas espectrais das culturas, o trabalho utilizou imagens do satélite *MODIS* do ano de 2020 através do produto MOD13Q1 disponibilizado a cada 16 dias com 250 metros de resolução espacial.

A identificação da data de plantio pode ser feita utilizando como referência os pontos de mínimos obtidos a partir de uma série temporal de um índice de vegetação (medida de vigor vegetativo), em que os mínimos locais podem estar associados ao solo com pouca ou nenhuma cobertura de vegetação, situação geralmente associada com o pré-plantio ou a pós-colheita.

A metodologia proposta buscou a partir dos pontos de mínimo e máximo da curva espectral, a identificação das datas de plantio e colheita respectivamente. O grande desafio neste método está em desconsiderar pontos de inflexão (mínimos e máximos locais), gerados por fatores como cobertura de nuvens, por exemplo.

A linguagem de programação Python foi utilizada para a identificação de 1664 apólices, e para tal, foi aplicado a ferramenta *Find Peaks* da biblioteca *Scipy*, esta função tem como objetivo destacar os máximos e mínimos locais por simples comparação de valores vizinhos da assinatura espectral.

A identificação de uma determinada cultura, bem como suas datas de plantio e colheita passam pelo reconhecimento espectral do índice vegetativo desta cultura. Assim, este trabalho utilizou a especificação de condição por proeminência, tendo sido avaliado anteriormente através de testes e checagem em massa dos dados, onde foi utilizado o valor de diferença de 0,2 para identificação das culturas, o que foi aplicado nas fórmulas:

**Fórmula 1:** Usada para encontrar a posição específica de todos os picos referentes aos máximos locais no dataframe com as informações de magnitudes de formas de sinais, aplicados para identificação da assinatura espectrográfica.

```
max_peaks,_ = find_peaks(sinalEspectral.to_numpy(), prominence=0.2)
```

**Fórmula 2:** Os índices dos dataframes Pandas[3] de todos os picos referentes aos mínimos locais com mais de 0,2 de proeminência foram coletados com a seguinte fórmula:

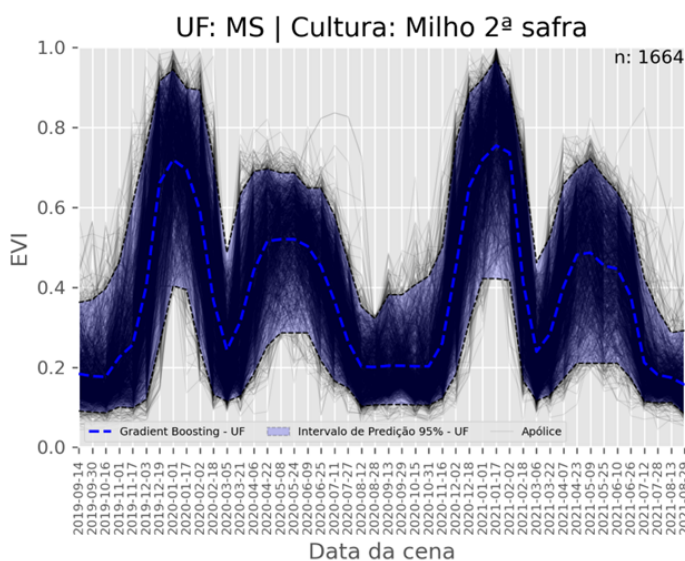
```
max_peaks,_ = find_peaks(-(sinalEspectral.to_numpy()), prominence=0.2)
```

Ao utilizar as Fórmulas 1 e 2 nos sinais de formação dos índices NDVI, EVI, PSRI e PDWI, pode-se identificar as seguintes informações: data de plantio, data de colheita, máximos locais de funções, mínimos locais de função, e o ciclo do plantio. Para identificar o comportamento padrão

(referência) foi aplicado às Apólices do PSR uma subdivisão a partir das Unidades Federativas (UF) do Brasil uma vez que cada UF compartilha diversas características similares, tais como fatores climáticos e possuindo como referência a época ideal de plantio e colheita de acordo com o Zoneamento Agrícola de Risco Climático - ZARC.

### 3. RESULTADOS

Após a conversão e padronização das datas ao longo do ano, referentes às Apólices do PSR e ao ZARC para o mesmo padrão das datas referentes aos índices de vegetação, foi possível iniciar os testes metodológicos, permitindo definir o comportamento padrão da cultura de Milho 2ª safra. A compilação de todas as curvas espectrais desta cultura



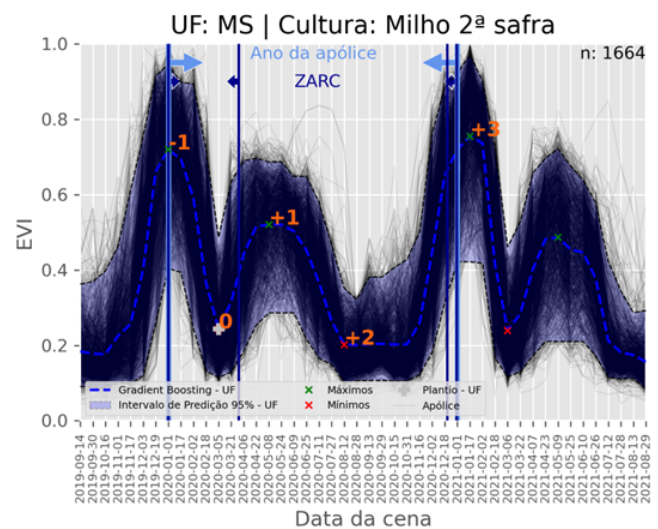
pode ser observada na Figura 1.

**Figura 1: Comportamento espectral padrão para o milho 2ª safra no MS**

Com a aplicação da ferramenta *Find Peaks* em ambiente *Python* foi possível identificar a data de plantio (dia/mês/ano), a partir da identificação dos pontos máximos e mínimos no comportamento espectral padrão.

A identificação dos pontos críticos é importante para o processo de obtenção das métricas de interesse, uma vez que os mesmos correspondem ao valor do índice espectral em datas relevantes para o ciclo da cultura agrícola. Esses pontos são obtidos pelo fato de serem considerados úteis para execução de uma tarefa específica, neste caso identificação dos estágios da cultura.

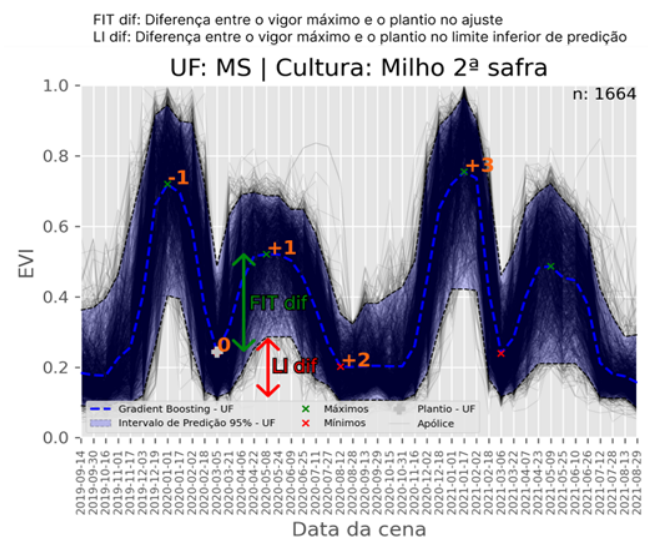
A Figura 2 apresenta exemplos de pontos críticos (data representada no eixo x e valor de índice de vegetação no eixo y), que são obtidos utilizando a data de plantio (Código 0) como referência. O ponto máximo anterior é o Vigor Vegetativo (+1) do Ciclo Atual. Em seguida, o próximo ponto de mínimo é o Ponto de Colheita (+2) da Cultura Agrícola e o próximo ponto máximo é o Máximo Vigor Vegetativo do Ciclo Anterior (-1) e o ponto de máximo posterior é o Máximo Vigor Vegetativo do Ciclo Posterior (+3).



**Figura 2 - Pontos críticos para o Milho 2ª Safra no MS**

### 4. DISCUSSÃO

Foram computadas apenas duas métricas: (1) diferença entre o valor de EVI no ponto de vigor vegetativo máximo e o EVI no ponto de plantio (FIT dif); e (2) a métrica LI dif foi computada utilizando as mesmas datas (ponto de máximo vigor vegetativo e de plantio - figura 3).



**Figura 3 - Pontos Críticos para a cultura do Milho 2ª safra no MS**

Porém, utilizando a curva do Limite Inferior do intervalo de predição do ajuste, os principais pontos críticos a serem identificados para cada apólice são: plantio (Código 0), máximo vigor vegetativo (+1) e colheita (+2). Estes pontos são obtidos através do cálculo de distâncias unidimensionais. Isto é, as distâncias entre as datas dos mínimos / máximos locais da apólice e pontos críticos do comportamento padrão (obtidos em 2 para plantio e 3 para os demais pontos críticos). Sendo assim, a data de plantio da



cultura agrícola na apólice é obtida via análise de proximidade, de modo que o plantio ocorre no ponto de mínimo encontrado na apólice que mais se aproxima (menor distância) do ponto de plantio em relação ao comportamento padrão.

Especificamente, para o plantio deve-se dar preferência para mínimos de acordo com o que é estabelecido pelo ZARC, referente a apólice. A identificação dos demais pontos críticos na apólice segue a mesma lógica, máximo vigor vegetativo para o ponto de máximo que mais se aproxima do máximo vigor vegetativo do comportamento geral, e colheita para o ponto de mínimo da apólice que mais se aproxima do ponto de colheita do comportamento geral.

## 5. CONCLUSÕES

O critério adotado até o momento foi bastante simplificado, em que uma combinação de UF - Cultura com menos de 200 apólices de amostra foram consideradas inadequadas para ajustar um bom modelo. Além disso, a obtenção da data do plantio pode ser melhorada uma vez que, atualmente, são utilizados os pontos de mínimos e máximos. Sugere-se desenvolver procedimentos para fazer um ajuste fino após obtenção dos pontos de mínimos e máximos. O foco deve ser o “pé da curva” do ciclo da cultura agrícola de interesse, e esse ponto é caracterizado pela aceleração do crescimento. Por fim, a obtenção dos pontos críticos pode ser aprimorada, uma vez que o modelo apresentado foi obtido utilizando medida de distância unidimensional, considerando apenas as datas do plantio. Para minimizar a quantidade de erros relacionados à detecção dos pontos críticos, bastaria modificar a forma de cálculo da medida de distância entre os pontos de mínimo/máximo da apólice e pontos críticos com comportamento padrão. Substituir distância unidimensional, que considera apenas as datas, por uma distância bidimensional, que considera as datas e valores de índices.

## 6. REFERÊNCIAS

- [1] WACKERNAGEL, M. et al. Tracking the ecological overshoot of the human economy. *Proceedings of the National Academy of Sciences of the United States of America*, v. 99, n. 14, p. 9266–9271, 2002.
- [2] HIRSCH, A. Habitat fragmentation and priority areas for primate conservation in the Rio Doce Basin, Minas Gerais. *Neotropical Primates*, v. 11, n. 3, p. 195–196, 2003.
- [3] CHOMITZ, K. et al. Viable reserve networks arise from individual landholder responses to Conservation incentives. *Ecology and Society*, Disponível em, v. 11, n. 2, 2006. Acesso: 22 maio 2019.
- [4] DOBROVOLSKI, Ricardo et al. Agricultural expansion can menace Brazilian protected areas during the 21st century. *Natureza & Conservação*, v. 9, n. 2, p. 208–213, 2011.
- [5] BILENCA, David et al. Impactos de la transformación agropecuária sobre la biodiversidad en la Provincia de Buenos Aires. *Revista del Museo Argentino de Ciencias Naturales*, v. 14, n. 2, p. 189–198, 2012.
- [6] ARAÚJO, H.M. et al. Diagnóstico espacial dos agricultores orgânicos certificados na área de atuação do Projeto CVT-MG Guayi - Spatial diagnosis of certified organic farmers on the scope of CVT-MG Guayi. *Cadernos de Agroecologia*, v. 10, n. 3, p. 1–6, 2015.
- [7] LEPSCH, I.F. 19 lições de Pedologia. [S.l.]: São Paulo, Oficina de Textos, 2011. p. 456
- [8] LANDAU, E.C.; L., Moura et al. Dinâmica espaço-temporal da produção de milho, soja e café no Estado de Minas Gerais entre 1990 e 2016. *Boletim de Pesquisa e Desenvolvimento / Embrapa Milho e Sorgo*, v. 162, p. 1–79, 2018.
- [9] TOTVS. Como Funciona o Sensoriamento Remoto na Agricultura. TOTVS, São Paulo.[S.l.: s.n.], 2020. Acesso em: 13/03/2022.
- [10] AWAD, M.M.; ALAWAR, B.; JBEILY, R. A New Crop Spectral Signatures Database Interactive Tool (CSSIT. *Data*, v. 4, n. 77, p. 1–14, 2019.
- [11] JENSEN, J.R. *Introductory Digital Image Processing: a remote sensing perspective*. 2nd. New Jersey: Prentice Hall, 1996. p. 318.
- [12] ZHANG, X. et al. Crop mapping using PROBA-V time series data at the Yucheng and Hongxing Farm in China. *Remote Sensing*, v. 8, n. 915, p. 1–18, 2016
- [13] EMBRAPA. SATVeg - Sistema de Análise Temporal da Vegetação: tutorial. In: EMBRAPA Agricultura Digital, Centro Nacional de Pesquisa em Tecnologia em Informática para a Agricultura – CNPTIA, Empresa Brasileira de Pesquisas Agropecuárias. Embrapa, Campinas: [s.n.], 2022. Acesso em 20/01/2022
- [14] AVERY, T.E.; BERLIN, G.L. *Fundamentals of Remote Sensing and Airphoto Interpretation*. 5th. Upper Saddle River, NJ: Prentice Hall, 1992. p. 472.
- [15] LILLESAND, T.M.; KIEFER, R.W. *Remote Sensing and Image Interpretation*. 3th. New York: John Wiley & Sons, 1994. p. 750.