

Uso de geotecnologias na obtenção de área plantada de soja e milho em Santa Catarina*Use of geotechnologies to obtain planted area of soybeans and corn in Santa Catarina**Uso de geotecnologías en la obtención de superficie sembrada de soja y maíz en Santa Catarina***Tarik Cuchi**

Universidade Federal de Viçosa. Departamento de Engenharia Florestal, Viçosa, Minas Gerais, Brasil.

 0000-0002-8930-2341**Haroldo Tavares Elias**

Empresa de Pesquisa Agropecuária e Extensão Rural de Santa Catarina. Centro de Socioeconomia e Planejamento Agrícola, Florianópolis, SC, Brasil.

 0000-0003-2863-8105**João Rogério Alves**

Empresa de Pesquisa Agropecuária e Extensão Rural de Santa Catarina. Centro de Socioeconomia e Planejamento Agrícola, Florianópolis, SC, Brasil.

 0009-0004-5285-9973**Bruna Parente Porto**

Empresa de Pesquisa Agropecuária e Extensão Rural de Santa Catarina. Centro de Socioeconomia e Planejamento Agrícola, Florianópolis, SC, Brasil.

 0009-0003-6492-0260

Resumo: O monitoramento das culturas de soja e milho é essencial para o agronegócio catarinense. Este estudo, um trabalho em andamento que apresenta resultados promissores, treinou um algoritmo de inteligência artificial (IA) para classificar soja, milho-grão e milho-silagem e estimar a área plantada líquida na safra 2023/2024. Utilizaram-se séries temporais de imagens orbitais, índices espectrais e dados do terreno. O classificador atingiu acurácia global de 92% e Kappa de 0,88, mostrando-se eficaz, auditável e promissor para o monitoramento agrícola do Estado, servindo como base para aprimoramentos futuros.

Palavras-chave: sensoriamento remoto; inteligência artificial; aprendizagem de máquina; Google Earth Engine.

Abstract: Monitoring soybean and corn crops is essential for agribusiness in Santa Catarina. This study, an ongoing work presenting promising results, trained an artificial intelligence (AI) algorithm to classify soybeans, grain corn, and silage corn and estimate the net planted area for the 2023/2024 harvest. Time series of orbital images, spectral indices, and terrain data were used. The classifier achieved an overall accuracy of 92% and a Kappa of 0.88, proving to be effective, auditable, and promising for agricultural monitoring in the state, serving as a basis for future improvements.

Keywords: remote sensing; artificial intelligence; machine learning; Google Earth Engine.

Resumen: El monitoreo de los cultivos de soja y maíz es esencial para la agroindustria en Santa Catarina. Este estudio, un proyecto en curso con resultados prometedores, entrenó un algoritmo de inteligencia artificial (IA) para clasificar la soja, el maíz de grano y el maíz para ensilaje, y estimar la superficie sembrada neta en la temporada agrícola 2023/2024. Se utilizaron series temporales de imágenes orbitales, índices espectrales y datos topográficos. El clasificador alcanzó una precisión general del 92 % y un coeficiente Kappa de 0,88, demostrando su eficacia, fiabilidad y potencial para el monitoreo agrícola en el Estado, sirviendo como base para futuras mejoras.

Palabras clave: teledetección; inteligencia artificial; aprendizaje automático; Google Earth Engine.

O estado de Santa Catarina é conhecido por sua diversidade na produção agropecuária, que inclui culturas agrícolas, produtos de origem animal e florestal. Essas atividades desempenham um papel crucial na economia estadual, sendo fontes de renda e emprego nos meios rural e urbano. Entre as principais culturas temporárias cultivadas no verão catarinense, destacam-se a soja, o milho-grão e o milho-silagem, por sua participação significativa na composição do valor da produção agropecuária do Estado em mais de 15% (Epagri/Cepa, 2025a).

Atualmente, a Epagri/Cepa dispõe de uma metodologia consolidada para estimar a área plantada, baseada na interlocução entre assistentes de pesquisa, extensionistas e analistas, que consolidam dados obtidos junto a produtores e cooperativas (Epagri/Cepa, 2025b). Apesar de robusto, esse método requer complementação por meio de mapeamentos por Sensoriamento Remoto (SR). Nos últimos anos, o avanço dessas geotecnologias tornou essa alternativa ágil, com ferramentas como o *Google Earth Engine* (GEE) permitindo a análise rápida de grandes áreas (Gorelick *et al.*, 2017), e algoritmos de inteligência artificial (IA) demonstrando elevada eficácia na classificação de dados orbitais (Elbasi *et al.*, 2023; Espinel *et al.*, 2024).

Diante do exposto, este estudo teve como objetivo desenvolver e aplicar uma metodologia complementar à da Epagri/Cepa, baseada no treinamento de um algoritmo de IA para classificar soja, milho-grão e milho-silagem em Santa Catarina, a partir de séries temporais de imagens orbitais processadas no GEE e na linguagem R. A proposta buscou gerar um mapeamento direto e auditável da área plantada líquida, ampliando o suporte às análises e ao planejamento territorial, configurando-se como um estudo preliminar para validação da técnica no Estado.

O estudo ocorreu em todo o território catarinense, utilizando uma série temporal de imagens *Harmonized Landsat Sentinel-2* (HLS) de outubro de 2023 a fevereiro de 2024, com resolução espacial de 30m e temporal de 2 a 3 dias (Neigh *et al.*, 2021). A metodologia, com base no algoritmo *Random Forest* (RF), foi aplicada em duas etapas: primeira, uma classificação binária pixel a pixel no *Google Earth Engine* (GEE) para mapear a presença das culturas-alvo. Em seguida, sobre a área resultante, uma Classificação Orientada a Objetos (OBIA) foi executada, utilizando a linguagem de programação R, para a distinção das culturas-alvo. O modelo foi treinado com 185 variáveis preditoras, extraídas de imagens filtradas por um limiar de até 65% de cobertura de nuvens (parâmetro CLOUD_COVERAGE) e submetidas ao mascaramento de qualidade pela banda Fmask. Nesta etapa, foram removidos os pixels identificados como nuvem (bit 1) e sombra de nuvem (bit 3), garantindo a integridade das observações. As séries temporais resultantes foram reduzidas a estatísticas temporais (média, mediana, desvio padrão, mínimo, máximo e percentis — 5, 25, 75 e 95). Foram utilizadas bandas espectrais, índices espectrais, dia juliano do pico de NDVI (máximo vigor), e dados de relevo do programa NASA SRTM (*Shuttle Radar Topography Mission*), conforme Tabela 1.

Tabela 1 – Bandas e índices espectrais selecionados para o processo de classificação

Abreviação	Nome Completo	Fórmula	Função Principal
Red	Banda Vermelha	-	Reflete a saúde da vegetação; usada em diversos índices.
NIR	Banda Infravermelho Próximo	-	Sensível à biomassa vegetal; essencial para índices de vegetação.
SWIR1	Banda Infravermelho Curto 1	-	Relacionada à umidade do solo e da vegetação.
SWIR2	Banda Infravermelho Curto 2	-	Indicador de estrutura da vegetação e cobertura do solo.
BI	<i>Brightness Index</i>	$\sqrt{((\text{Red}^2 + \text{Green}^2 + \text{Blue}^2)/3)}$	Mede o brilho da superfície; útil para diferenciar tipos de cobertura.
SR	<i>Simple Ratio</i>	NIR / Red	Mede diretamente o vigor da vegetação.
NDVI	<i>Normalized Difference Vegetation Index</i>	$(\text{NIR} - \text{Red})/(\text{NIR} + \text{Red})$	Avalia a densidade e o vigor da vegetação.
DOYNDVI_{max}	Dia Juliano do pico de NDVI	-	Indica o dia do ano com maior vigor vegetativo.
EVI2	<i>Enhanced Vegetation Index 2</i>	$2,5 \times ((\text{NIR} - \text{Red})/(\text{NIR} - \text{Red} + 1))$	Melhora a sensibilidade em áreas com vegetação densa.
GCVI	<i>Green Chlorophyll Vegetation Index</i>	$(\text{NIR} / \text{Green}) - 1$	Estima o teor de clorofila na vegetação.
GWCCI	<i>Greenness and Water Content Composite Index</i>	$((\text{NIR} - \text{Red})/(\text{NIR} + \text{Red})) \times \text{SWIR1}$	Combina vigor da vegetação com conteúdo hídrico.
LSWI	<i>Land Surface Water Index</i>	$(\text{NIR} - \text{SWIR1})/(\text{NIR} + \text{SWIR1})$	Detecta umidade da vegetação e presença de água.
NDSVI	<i>Normalized Difference Senescent Vegetation Index</i>	$(\text{SWIR1} - \text{Red})/(\text{SWIR1} + \text{Red})$	Identifica vegetação senescente ou em declínio.
NDWI	<i>Normalized Difference Water Index</i>	$(\text{NIR} - \text{Green})/(\text{NIR} + \text{Green})$	Indica presença de água na vegetação ou solo.
MNDWI	<i>Modified Normalized Difference Water Index</i>	$(\text{Green} - \text{SWIR1})/(\text{Green} + \text{SWIR1})$	Detecta corpos d'água com maior precisão.
MCRC	<i>Modified Chlorophyll Reflectance Index</i>	$(\text{SWIR1} - \text{Green})/(\text{SWIR1} + \text{Green})$	Estima teor de clorofila com sensibilidade à umidade.
NDRI	<i>Normalized Difference Residue Index</i>	$(\text{Red} - \text{SWIR2})/(\text{Red} + \text{SWIR2})$	Detecta resíduos de colheita ou cobertura morta.
NDTI	<i>Normalized Difference Tillage Index</i>	$(\text{SWIR1} - \text{SWIR2})/(\text{SWIR1} + \text{SWIR2})$	Diferencia áreas com palha daquelas com solo exposto.
STI	<i>Simple Tillage Index</i>	SWIR1 / SWIR2	Alternativa simples para identificar solo exposto versus cobertura vegetal.
BUI	<i>Built-up Index</i>	$(\text{SWIR1} - \text{NIR})/(\text{SWIR1} + \text{NIR} - \text{Red})/(\text{SWIR1} + \text{NIR} + \text{Red})$	Detecta áreas urbanizadas.
NDBI	<i>Normalized Difference Built-up Index</i>	$(\text{SWIR1} - \text{NIR})/(\text{SWIR1} + \text{NIR})$	Identifica áreas construídas.

Fonte: Elaborado pelos autores (2026).

O conjunto de dados de referência (ground truth) foi constituído por 481 amostras georreferenciadas, estratificadas nas classes de soja ($n = 237$), milho silagem ($n = 147$) e milho grão ($n = 97$). A aquisição das coordenadas foi realizada por equipes técnicas da Epagri/Cepa em cooperação com empresas do setor produtivo. Ressalta-se que a distribuição espacial das amostras contemplou todas as mesorregiões do Estado, assegurando a representatividade da variabilidade espectro-temporal das culturas frente às distintas condições edafoclimáticas da área de estudo.

Para a classificação binária, foram coletadas 18.650 amostras de pixels, divididas igualmente entre “presença” e “ausência”. As amostras de “presença” foram criadas a partir de polígonos desenhados em talhões homogêneos, usando as coordenadas de campo como referência. As de “ausência” foram obtidas por fotointerpretação sobre o basemap do Google e a mediana da série temporal HLS, com apoio complementar do MapBiomias — Coleção 9 para casos duvidosos, abrangendo diversas classes como vegetação arbórea, campos, água e outras culturas não-alvo. O conjunto de dados foi dividido em 70% para treino e 30% para teste, resultando em uma máscara de alta fidelidade das culturas-alvo, que serviu como base para a posterior classificação OBIA.

Em posse dos dados HLS mascarados, procedeu-se à segmentação para a classificação OBIA (distinção das culturas) utilizando o algoritmo de Crescimento de Regiões da biblioteca TerraLib, implementado via pacote SegOptim no ambiente R. Foram adotados como parâmetros um tamanho mínimo de segmento de 1 pixel e um limiar de similaridade de 0,025.

É importante ressaltar que, nesta metodologia, a segmentação não teve por escopo a delimitação cadastral exata dos limites dos talhões agrícolas, mas sim a agregação de pixels adjacentes com comportamento espectral similar. Essa estratégia de geração de objetos homogêneos (superpixels) foi empregada fundamentalmente para mitigar a variabilidade espectral intraclasse e reduzir o ruído de alta frequência (“efeito sal e pimenta”), frequente em abordagens pixel a pixel, garantindo assim maior consistência radiométrica às unidades classificadas.

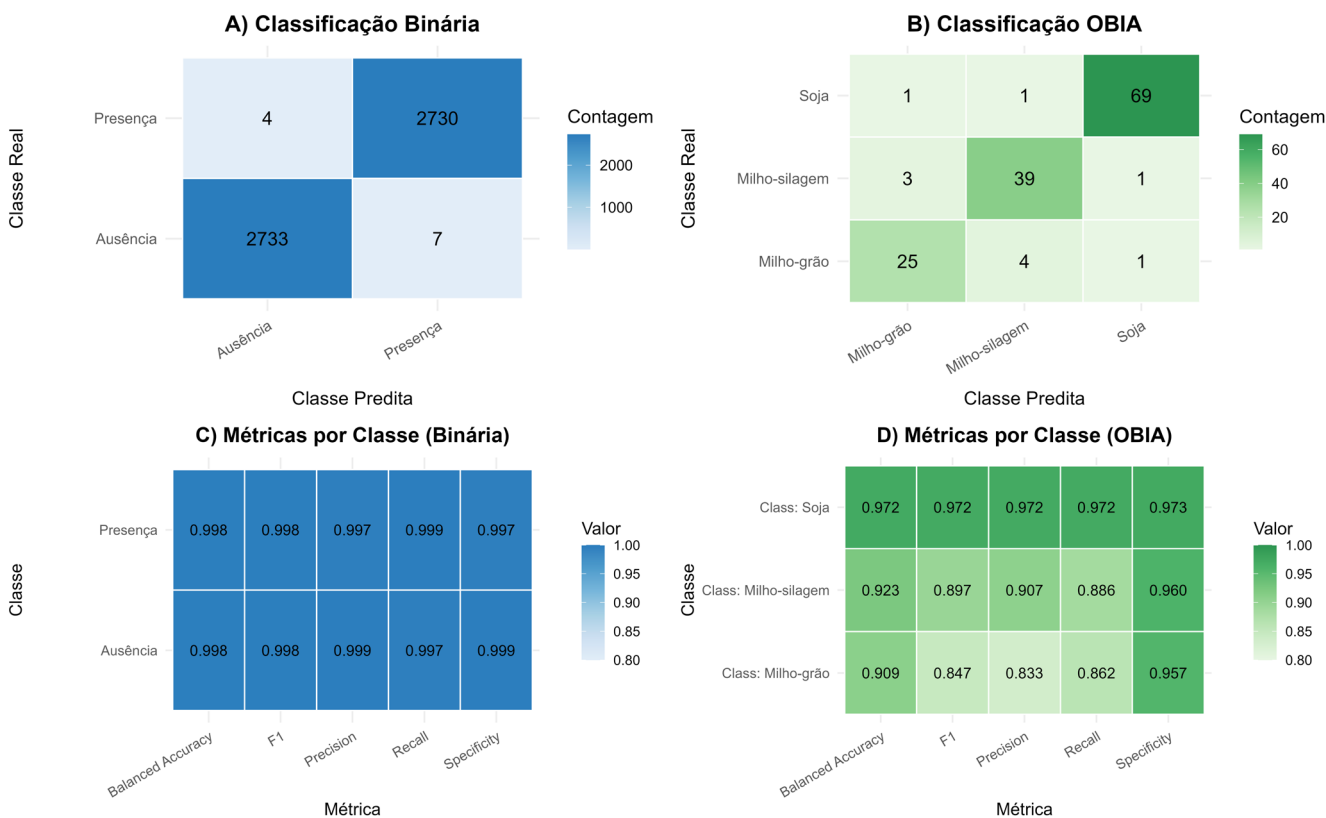
Na classificação OBIA, a atribuição das classes de uso e cobertura aos segmentos foi realizada mediante operação espacial de contenção (*spatial join*) entre os vetores de amostras e a segmentação gerada. O processamento foi conduzido em ambiente R, utilizando a estrutura do *tidymodels* integrada ao *ranger* e *terra*. Os dados foram particionados aleatoriamente em 70% para treinamento e 30% para teste. Para mitigar o desbalanceamento entre classes, aplicou-se o algoritmo SMOTE (*Synthetic Minority Over-sampling Technique*) com uma razão de *oversampling* de 2 (*over_ratio* = 2). Ressalta-se que o SMOTE foi aplicado estritamente sobre os dados de treinamento dentro do fluxo de processamento (*recipe*), mantendo o conjunto de teste composto exclusivamente por dados originais de campo para garantir a independência da validação.

O modelo Random Forest foi configurado com 500 árvores (*num.trees* = 500) e teve seus hiperparâmetros otimizados via busca em grade (*grid search*) com 4 níveis de regularidade. Foram ajustados simultaneamente o número de variáveis para divisão a cada nó (*mtry*) e o tamanho mínimo do nó (*min_n*), utilizando validação cruzada estratificada em 5 partes (*5-fold cross-validation*). O critério de seleção do melhor modelo foi a maximização da acurácia. A avaliação final sobre os 30% de teste incluiu a análise da matriz de confusão e o cálculo de métricas globais e por classe. A classificação binária alcançou 99% de acurácia global (Kappa 0,99), enquanto a classificação OBIA obteve 92% de acurácia global (Kappa = 0,88), indicando um modelo preditivo robusto. Na análise por classe (Figura 1), a soja apresentou melhor desempenho. As classes de milho tiveram alta especificidade, mas o milho-grão registrou menor recall e precisão, provavelmente devido ao número reduzido de amostras, limitando a generalização da variabilidade espectral.

A metodologia resultou no mapa da safra 2023/2024 (Figura 2), estimando 878.398ha de soja, 202.122ha de milho-grão e 353.937ha de milho-silagem. Esses resultados contribuem ao implementar uma nova tecnologia e, consequentemente, complementar o atual método de estimativas da Epagri/Cepa.

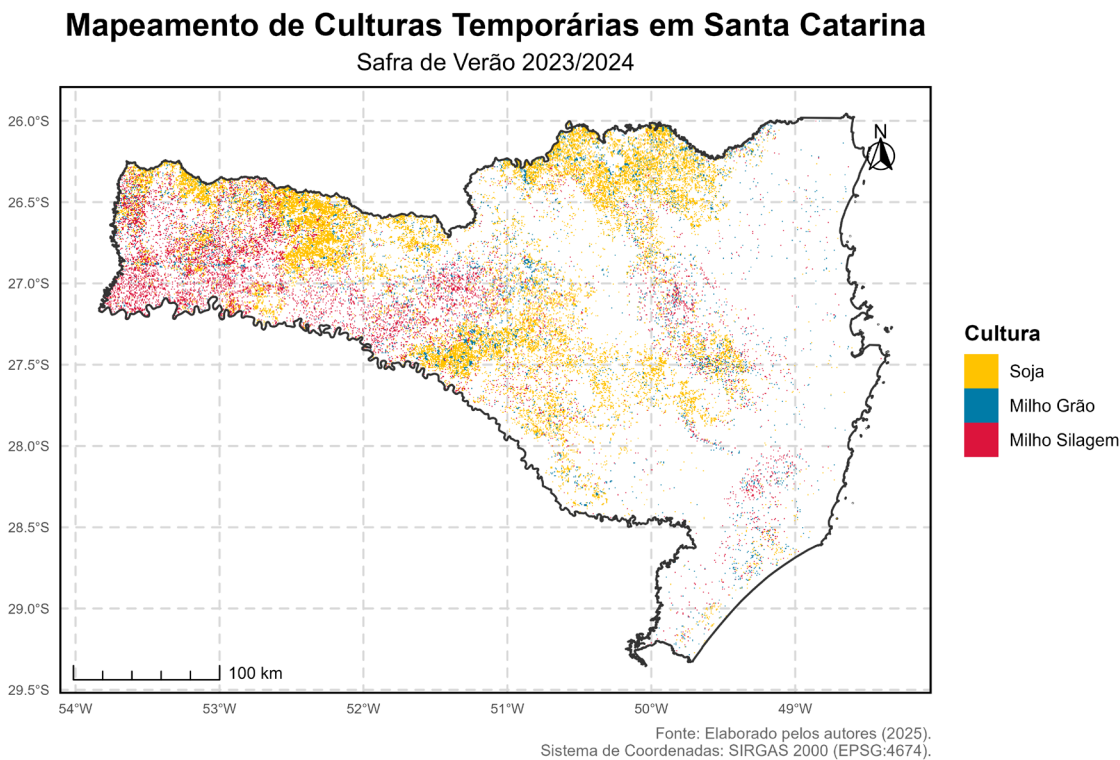
Embora a confusão espectral entre milho-grão e milho-silagem seja um desafio metodológico, o uso de séries temporais densas (HLS) mostra-se eficaz para essa distinção. Conforme validado por pesquisas recentes (Lin *et al.*, 2025; Shahrabi *et al.*, 2020), a abordagem captura eventos fenológicos chave, distinguindo a colheita da biomassa ainda verde (silagem) da senescência completa (grão). Adicionalmente, a alta resolução temporal permite acomodar as variações nas datas de semeadura que, mesmo balizadas pelo zoneamento agrícola estadual, oscilam naturalmente devido a fatores climáticos e operacionais locais fora do controle experimental.

Figura 1 – Resultados da validação das classificações (binária e OBIA) realizadas em amostras de teste



Fonte: Elaborado pelos autores (2026).
Nota: A — Matriz de confusão da classificação binária; B — Matriz de confusão da classificação OBIA; C — Métricas de avaliação por classe da classificação binária; D — Métricas de avaliação por classe da classificação OBIA

Figura 2 – Mapa da distribuição das culturas soja, milho-grão e milho-silagem da safra 2023/2024 no estado de Santa Catarina (primeira safra)



Fonte: Elaborado pelos autores (2026).

O desbalanceamento de amostras, especialmente para milho-grão, é outra restrição, mesmo com a aplicação da técnica SMOTE. Este continua sendo um desafio central em classificações, onde o desempenho de técnicas de *oversampling* é um tópico de pesquisa ativa para garantir que classes minoritárias sejam corretamente identificadas sem comprometer a generalização do modelo (Fonseca; Douzas; Bacao, 2021).

Observou-se a presença de ruídos de classificação (erros de comissão) em áreas não-agrícolas, como zonas urbanas litorâneas (e.g., Florianópolis), decorrentes de confusão espectral em pixels mistos. Tais erros, inerentes a classificações automáticas nesta escala, não comprometem a estimativa global nas principais regiões produtoras e serão mitigados em etapas futuras com a aplicação de máscaras de exclusão urbana e de uso do solo.

Adicionalmente, a janela temporal analisada foca exclusivamente na primeira safra, não abrangendo cultivos de segunda safra, delimitando o escopo do mapa. A pesquisa segue em andamento, e o aprimoramento contínuo do modelo depende de um ciclo de validação constante em campo, em sinergia com a competência dos técnicos da Epagri, para fortalecer a calibração do modelo e sua aplicabilidade operacional em safras futuras com a técnica de transferência de aprendizagem.

REFERÊNCIAS

- ELBASI, E.; ZAKI, C.; TOPCU, A. E.; ABDELBAKI, W.; ZREIKAT, A. I.; CINA, E.; SHDEFAT, A.; SAKER, L. Crop Prediction Model Using Machine Learning Algorithms. **Applied Sciences**, v. 13, n. 16, p. 9288, 16 ago. 2023. DOI: <http://dx.doi.org/10.3390/app13169288>.
- EMPRESA DE PESQUISA AGROPECUÁRIA E EXTENSÃO RURAL DE SANTA CATARINA. CENTRO DE SOCIOECONOMIA E PLANEJAMENTO AGRÍCOLA. **Síntese Anual da Agricultura de Santa Catarina**: 2024. Florianópolis: Epagri/Cepa, 2025a. Disponível em: https://docweb.epagri.sc.gov.br/website_cepa/publicacoes/Sintese_2024.pdf. Acesso em: 12 nov. 2025.
- EMPRESA DE PESQUISA AGROPECUÁRIA E EXTENSÃO RURAL DE SANTA CATARINA. CENTRO DE SOCIOECONOMIA E PLANEJAMENTO AGRÍCOLA. **Metodologia e estratégia de ação**: Projeto Safras. Florianópolis: Epagri/Cepa, 2025b. Disponível em: <https://drive.google.com/file/d/1ELcn2GrhjvCD5lXbskU1od1woS18kKbw/view>. Acesso em: 15 jan. 2026.
- ESPINEL, R.; HERRERA-FRANCO, G.; RIVADENEIRA GARCÍA, J. L.; ESCANDÓN-PANCHANA, P. Artificial Intelligence in Agricultural Mapping: A Review. **Agriculture**, v. 14, n. 7, p. 1071, 2024. DOI: <http://dx.doi.org/10.3390/agriculture14071071>.
- FONSECA, J.; DOUZAS, G.; BACAO, F. Improving Imbalanced Land Cover Classification with K-Means SMOTE: Detecting and Oversampling Distinctive Minority Spectral Signatures. **Information**, v. 12, n. 7, p. 266, 2021. DOI: <http://dx.doi.org/10.3390/info12070266>.
- GORELICK, N.; HANCHER, M.; DIXON, M.; ILYUSHCHENKO, S.; THAU, D.; MOORE, R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. **Remote Sensing of Environment**, v. 202, p. 18–27, 2017. DOI: <http://dx.doi.org/10.1016/j.rse.2017.06.031>.
- LIN, Z.; HUANG, R.; TAN, S.; YANG, L.; HUANG, J.; SU, L.; HU, Z. Silage Maize Identification Using a Temporal Difference-Based Model with Sentinel-2 Data: Insights from a Harvest-Based and Temporally Transferable Approach. **Agronomy**, v. 15, n. 6, p. 1438, 2025. DOI: <http://dx.doi.org/10.3390/agronomy15061438>.
- NEIGH, C.; JU, J.; ROGER, J.-C.; SKAKUN, S.; VERMOTE, E.; CLAVERIE, M.; DUNGAN, J.; YIN, Z.; FREITAG, B.; JUSTICE, C. **HLS Operational Land Imager Surface Reflectance and TOA Brightness Daily Global 30m v2.0**: NASA Land Processes Distributed Active Archive Center, 2021. DOI: <https://doi.org/10.5067/HLS/HLSL30.002>.

INFORMAÇÕES ADICIONAIS

Breve currículo dos autores:

Tarik Cuchi

Engenheiro florestal (UFSC/FW). Mestre em Engenharia Florestal (Udesc/CAV). Doutor em Ciências Florestais (Unicentro/Irati). Possui trajetória na área de Recursos Florestais, tendo experiência em Floresta Urbana, Ecologia de Paisagens e Geoinformática, na análise da vegetação no gradiente urbano-rural e gestão territorial. Atualmente, é professor na Universidade Federal de Viçosa - UFV, no Departamento de Engenharia Florestal - DEF. E-mail: tarik.cuchi@ufv.br

Haroldo Tavares Elias

Engenheiro-agrônomo, formado pela Universidade Federal de Santa Catarina (UFSC) em 1986. Possui mestrado em Genética e Melhoramento de Plantas pela Universidade Federal de Lavras (UFLA) e doutorado em Agronomia pela Universidade Estadual de Maringá (UEM). Atualmente, é analista de pesquisa em socioeconomia na Empresa de Pesquisa Agropecuária e Extensão Rural de Santa Catarina (Epagri) trabalhando com temas relacionados à análise de mercado, produção vegetal e agricultura familiar. E-mail: htelias@epagri.sc.gov.br

João Rogério Alves

Engenheiro-agrônomo (UFSC/SC). Mestre em Agroecossistemas (UFSC/SC). Especialista em Biossegurança (UFSC/SC). Atualmente, é analista de Socioeconomia e Desenvolvimento Rural do Centro de Socioeconomia e Planejamento Agrícola da Empresa de Pesquisa Agropecuária e Extensão Rural de Santa Catarina (Epagri/Cepa). Atuando nos seguintes temas: agricultura familiar, levantamento de preços agropecuários, custos de produção, análise de mercado e acompanhamento de safras agrícolas. E-mail: joaoalves@epagri.sc.gov.br

Bruna Parente Porto

Graduação em Agronomia pela Universidade Federal de Santa Catarina, 2012. Atualmente, é extensionista rural na Empresa de Pesquisa Agropecuária e Extensão Rural de Santa Catarina (Epagri), atuando no projeto Levantamento das expectativas e projeções da produção agrícola catarinense, desenvolvido pelo Centro de Socioeconomia e Planejamento agrícola (Epagri/Cepa). E-mail: brunaporto@epagri.sc.gov.br

Contribuição de autoria - Credit:

Tarik Cuchi: Conceituação, curadoria dos dados, análise formal, administração do projeto, investigação, metodologia, validação dos dados, redação do manuscrito original, Revisão e edição.

Haroldo Tavares Elias: Conceituação, curadoria dos dados, investigação, metodologia, validação dos dados, Revisão e edição.

João Rogério Alves: Conceituação, curadoria dos dados, investigação, metodologia, validação dos dados, Revisão e edição.

Bruna Parente Porto: Conceituação, curadoria dos dados, investigação, metodologia, validação dos dados, Revisão e edição.

Conflitos de interesse: não há conflito.

Licença de uso: conforme normas da revista CCBY 4.0.

Preprints: não se aplica.

Aprovação do Comitê de Ética na Investigação: não se aplica.

Consentimento para o uso de imagens: não se aplica.

Acordo com a revisão aberta: não se aplica.

Disponibilidade dos dados de investigação e outros materiais: dados serão disponibilizados pelos autores por solicitação;

Agradecimentos: os autores agradecem à Fundação de Amparo à Pesquisa e Inovação do Estado de Santa Catarina (Fapesc) pelo apoio financeiro e institucional, essencial para a realização desta pesquisa e para o alcance de seus objetivos.

Financiamento: Fapesc/TO — 2023TR001539

Agradecimentos: À Fapesc no atendimento á pesquisa pelo Edital Fapesc/Epagri n.º 17/2023 — Fapesc/TO — 2023TR001539.