

## PREDIÇÃO DA QUALIDADE DE CANA-DE-AÇÚCAR UTILIZANDO IMAGENS ORBITAIS E REDE NEURAL ARTIFICIAL

TATIANA F. CANATA<sup>1</sup>, MARCELO R. BARBOSA JÚNIOR<sup>2</sup>, ROMÁRIO P. DE OLIVEIRA<sup>2</sup>, JOÃO P. P. CHIARELLI<sup>3</sup>, ROUVERSON P. DA SILVA<sup>4</sup>, CARLOS E. A. FURLANI<sup>5</sup>

<sup>1</sup> Eng. Biossistemas, Profa. Doutora, Depto. de Engenharia e Ciências Exatas, Faculdade de Ciências Agrárias e Veterinárias (FCAV/UNESP), Jaboticabal/SP. (16) 3209-7280, tatiana.canata@unesp.br

<sup>2</sup> Mestre em Agronomia (Produção Vegetal), Faculdade de Ciências Agrárias e Veterinárias (FCAV/UNESP), Jaboticabal/SP.

<sup>3</sup> Aluno de graduação em Eng. Agrônoma, Faculdade de Ciências Agrárias e Veterinárias (FCAV/UNESP), Jaboticabal/SP.

<sup>4</sup> Eng. Agrícola, Prof. Adjunto, Depto. de Engenharia e Ciências Exatas, Faculdade de Ciências Agrárias e Veterinárias (FCAV/UNESP), Jaboticabal/SP.

<sup>5</sup> Eng. Agrônomo, Prof. Titular, Depto. de Engenharia e Ciências Exatas, Faculdade de Ciências Agrárias e Veterinárias (FCAV/UNESP), Jaboticabal/SP.

Apresentado no  
LII Congresso Brasileiro de Engenharia Agrícola - CONBEA 2023  
18 a 21 de outubro de 2023 – Ribeirão Preto - SP, Brasil

**RESUMO:** A antecipação dos valores do conteúdo de açúcar presente nos colmos de cana-de-açúcar é essencial para o desenvolvimento de soluções orientadas por dados na agricultura. O objetivo é explorar o potencial de integração entre os dados multiespectrais e os algoritmos de inteligência artificial para a predição de atributos qualitativos da cana-de-açúcar (Brix e Pureza). Estes atributos foram mensurados em laboratório utilizando 240 amostras de colmos de cana-de-açúcar. As imagens de satélite foram adquiridas nas mesmas datas de avaliação em campo a fim de compor a base de dados de refletância do dossel da cultura e o cálculo dos índices de vegetação. Os métodos de Regressão Linear Múltipla (MLR) e Rede Neural Artificial (ANN) foram utilizados no desenvolvimento dos modelos preditivos do conteúdo de Brix e Pureza da cultura. A performance dos modelos foi comparada em relação à acurácia ( $R^2$ ) e ao erro de predição (RMSE-raiz quadrada do erro médio). A MLR apresentou melhor performance na predição de Brix ( $R^2=0,71$ ; RMSE=0,12°Brix). Os valores de Pureza foram preditos com maior acurácia pela ANN ( $R^2=0,92$ ; RMSE=0,11%). Foram obtidos resultados satisfatórios integrando imagens de satélite, dados de biometria da cultura, e algoritmos de inteligência artificial para a predição de atributos qualitativos de cana-de-açúcar.

**PALAVRAS-CHAVE:** agricultura de precisão, inteligência artificial, *Saccharum spp.*

## PREDICTING SUGARCANE QUALITY USING ORBITAL IMAGERY AND ARTIFICIAL NEURAL NETWORK

**ABSTRACT:** Anticipating the values of sugar content from sugarcane stalks is essential to develop data-driven solutions in agriculture. The objective is to explore the potential of integrating multispectral data and artificial intelligence algorithms to predict qualitative attributes of sugarcane (Brix and Purity). Those attributes were measured in laboratory using 240 samples of sugarcane stalks. The satellite imagery was taken for each date of evaluation to compose the database of crop canopy reflectance and calculation of vegetation index. The methods of Multiple Linear Regression (MLR) and Artificial Neural Network (ANN) were

used to develop the predictive models of Brix and Purity crop contents. The performance of the models was compared in terms of accuracy ( $R^2$ ) and error of prediction (RMSE-root mean square error). The MLR predicted Brix content with greater accuracy ( $R^2=0.71$ ;  $RMSE=0.12^\circ\text{Brix}$ ). Purity values were better predicted by ANN modeling ( $R^2=0.92$ ;  $RMSE=0.11\%$ ). Satisfactory results of accuracy were achieved integrating satellite imagery, biometric data, and artificial intelligence-based algorithm on predicting qualitative attributes of sugarcane crop.

**KEYWORDS:** artificial intelligence, precision agriculture, *Saccharum* spp.

**INTRODUÇÃO:** As soluções de agricultura digital têm permitido uma maior aquisição de dados a fim de identificar o potencial produtivo das culturas em larga escala. Tais soluções integradas com as operações mecanizadas viabilizam a otimização de recursos nas lavouras e o aumento de produtividade e qualidade dos produtos de forma sustentável. O uso de dados de sensoriamento remoto no Brasil é focado no diagnóstico precoce de doenças e pragas nas lavouras e na determinação do estresse hídrico de plantas (BOLFE et al., 2020). A cana-de-açúcar é uma das principais culturas com potencial para o aprimoramento de produção utilizando agricultura de precisão e agricultura digital, tendo em vista que a colheita mecanizada representa 33% do custo total de produção para a indústria (ILLUKPITIYA et al., 2022). Assim, é fundamental que sejam desenvolvidas metodologias e ferramentas alternativas em relação à tomada de decisão de colheita da cana-de-açúcar em larga escala. O método tradicional de pagamento aos produtores de cana-de-açúcar é baseado na massa total de matéria-prima e no indicativo de qualidade do material (ATR-açúcar total recuperável). Este indicativo é calculado a partir de uma média amostral dos valores de Brix (sólidos solúveis) e de Pol (sacarose aparente), os quais são mensurados em laboratório utilizando amostragem destrutiva e de baixa resolução, inviabilizando o propósito de intervenções localizadas nas lavouras (FERRAZ et al., 2019). Existem diversas alternativas ao monitoramento de lavouras de cana-de-açúcar baseadas em sensoriamento remoto para a predição de produtividade (CANATA et al., 2021) e no uso de sensores embarcados em colhedoras (MALDANER et al., 2022). Porém, poucos estudos relatam sobre o mapeamento dos atributos de qualidade da cultura. Nawi et al. (2014) observaram relação entre os atributos de qualidade de cana-de-açúcar e os comprimentos de onda no visível e no infravermelho próximo em condições de laboratório. Phetpan et al. (2018) utilizaram a mesma região do espectro eletromagnético com sensores de espectroscopia embarcados no elevador da colhedora de cana-de-açúcar para a predição de graus Brix com assertividade de 78%, aumentando a capacidade de aquisição de dados em campo. Pressupõe-se que o desenvolvimento de metodologias alternativas para a predição de atributos de qualidade da cana-de-açúcar de forma espacializada contribua para o planejamento mais efetivo da colheita mecanizada de cana-de-açúcar e o aumento de sua capacidade operacional. O objetivo do trabalho é explorar a integração entre os dados multiespectrais e os algoritmos de inteligência artificial para a predição de atributos qualitativos da cana-de-açúcar.

**MATERIAL E MÉTODOS:** A metodologia do estudo foi dividida em coleta de dados, pré-processamento dos dados, desenvolvimento dos modelos e avaliação de performance da modelagem. Os modelos preditivos de qualidade de cana-de-açúcar foram desenvolvidos baseando-se na integração de dados de refletância a partir das imagens de satélite, dados de biometria da cultura e medidas georreferenciadas em campo. O estudo foi conduzido em uma área comercial de cana-de-açúcar (21° 18' 55"S, 48° 9' 35"O; 562 m altitude) com a variedade

RB-975201 (6° corte) instalada no campo e espaçamento entrelinhas de 1,50 m. A área total é de 0,32 ha e o tipo de solo é denominado como Oxisol. Foram realizadas oito avaliações considerando amostragem em grade de 9 m x 9 m. Os colmos de cana-de-açúcar foram coletados manualmente a cada 15 dias durante a fase de maturação da cultura nas quatro linhas centrais de cada parcela. Foram coletadas 30 amostras georreferenciadas, totalizando 240 amostras (8 datas de avaliação x 30 amostras por avaliação) encaminhadas ao laboratório para a mensuração de Brix e Pureza de acordo com a metodologia descrita pela Consecana (2006). O conteúdo de Brix foi mensurado por meio de um refratômetro digital (ABBE, Atago Pal-1, Tokyo, Japão). A soma de graus dias de crescimento (GDD - Equação 1) foi utilizada como variável de entrada nos modelos preditivos como dado de biometria da cultura. A temperatura basal (Tb) da cultura foi considerada igual a 18 °C (BACHI e SOUZA, 1978).

$$\text{GDD} = [(T_{\max} - T_{\min})/2] - T_b \quad (1)$$

em que,

GDD - graus dias de crescimento;

Tmax – temperatura máxima do ar (°C);

Tmin – temperatura mínima do ar (°C);

Tb – temperatura basal da cultura (°C).

As imagens de satélite foram selecionadas por meio da plataforma PlanetScope, resolução espacial de 3,0 m, de acordo com a data de avaliação em campo e em função dos dias após o corte (DAC) em relação à data de colheita anterior à coleta de dados. As imagens foram processadas no sistema de informação geográfica denominado QGIS (versão 2.18.26). Um *buffer* interno de 5,0 m foi aplicado ao contorno da área de estudo. As bandas espectrais foram analisadas individualmente (vermelho, verde, azul, e NIR - infravermelho próximo) e foram considerados os índices de vegetação comumente citados na literatura (CANATA et al., 2021) conforme apresentado na Tabela 1.

TABELA 1. Índices de vegetação considerados neste estudo. **Vegetation index considered in the study.**

Índices de vegetação	Equação	Autores
Blue Green Pigment Index (BGI)	Blue/Green	Zarco-Tejada et al. (2005)
Color Index of Vegetation Extraction (CIVE)	$0.441 \times \text{Red} - 0.811 \times \text{Green} + 0.385 \times \text{Blue} + 18.79$	Kataoka et al. (2003)
Green Normalized Difference Vegetation Index (GNDVI)	$(\text{NIR} - \text{Green}) / (\text{NIR} + \text{Green})$	Gitelson et al. (1996)
Normalized Difference Vegetation Index (NDVI)	$(\text{NIR} - \text{Red}) / (\text{NIR} + \text{Red})$	Rouse et al. (1974)

Red: refletância no comprimento de onda do vermelho (590 nm-670 nm); Blue: azul (455 nm-515 nm); Green: verde (500 nm-590 nm); NIR: infravermelho próximo (780 nm-860 nm)

O desenvolvimento dos modelos preditivos foi baseado em nove variáveis independentes: valores de refletância de cada banda espectral (quatro variáveis), índices de vegetação (quatro índices) e a variável GDD. Os atributos de qualidade (Brix e Pureza) foram consideradas como variáveis dependentes na modelagem. Os atributos de entrada e saída foram agrupados e normalizados entre 0 e 1, e posteriormente divididos em treinamento e teste considerando a proporção 70:30. A análise estatística e a elaboração dos modelos preditivos foram executadas na linguagem de programação R 4.2.2 (R Core Team). A seleção das variáveis de entrada relevantes para a predição dos atributos foi realizada através do

pacote *olsrr* (NETER et al., 1996). A Regressão Linear Múltipla (MLR) foi desenvolvida a partir da função *lm* e a Rede Neural Artificial (ANN) por meio do pacote *neuralnet* (GÜNTHER e FRITSCH, 2010). O número de camadas de entrada na ANN corresponde ao número de variáveis selecionadas previamente à execução da modelagem e a camada de saída é o valor predito individualmente (Brix ou Pureza). O número de camadas ocultas foi determinado de acordo com o menor erro obtido entre os valores observado e predito pela ANN (*hidden*=5.000). A performance dos métodos de MLR e ANN foram comparados em relação ao coeficiente de determinação ( $R^2$ ), à raiz quadrada do erro-médio (RMSE), ao erro percentual absoluto médio (MAPE), e ao erro absoluto médio (MAE).

**RESULTADOS E DISCUSSÃO:** O conjunto de dados do conteúdo de Brix variou entre 7,9° e 24,8° e a Pureza entre 40,2% e 99,9%. Os maiores valores de Brix foram observados na sétima avaliação em campo (270 DAC), indicando o momento mais apropriado para a colheita da matéria-prima. Os valores do coeficiente de variação (CV) para Brix (<20%) e Pureza (<21%) indicaram uma baixa-moderada variabilidade espacial dos atributos de qualidade mensurados na área de estudo. A análise de regressão indicou a variável GDD como a mais relevante na predição dos atributos de qualidade, seguido pela banda espectral na região do azul (455 nm - 515 nm). Alguns autores relataram a relação entre a banda espectral na região do azul e o conteúdo de açúcar de diferentes culturas utilizando dados hiper espectrais do dossel das plantas (SIEDLISKA et al., 2021). A Tabela 2 sintetiza os resultados obtidos em relação à performance dos modelos preditivos baseados em MLR e ANN em relação à estimativa de Brix e Pureza.

TABELA 2. Métricas de performance dos modelos preditivos. **Metrics of performance from predictive models.**

	Brix (°)				Pureza (%)			
	RMSE	MAE	MAPE	$R^2$	RMSE	MAE	MAPE	$R^2$
<b>Modelos</b>			%				%	
ANN	0,16	0,12	0,31	0,54	0,11	0,08	0,44	0,92
MLR	0,12	0,09	0,17	0,71	0,10	0,07	0,20	0,86

O menor erro foi obtido na estimativa de Brix considerando a modelagem baseada em MLR (MAPE=0,17%) e maior acurácia na predição deste atributo ( $R^2=0,71$ ). Enquanto que, a ANN apresentou melhores resultados na estimativa do atributo Pureza (MAPE=0,20% e  $R^2=0,92$ ). De modo geral, o modelo preditivo baseado em ANN superestimou os valores de Brix e apresentou menor variação do erro na estimativa de Pureza em comparação ao modelo baseado em MLR (Figura 1). O índice de vegetação GNDVI e a variável GDD apresentaram maior relevância na estimativa dos atributos de qualidade em ambos os modelos de predição. Os modelos desenvolvidos demonstraram capacidade de predição dos atributos qualitativos de cana-de-açúcar de forma antecipada à colheita mecanizada utilizando ferramentas de sensoriamento remoto e, portanto, com alta resolução espacial e temporal de forma não-destrutiva. A acurácia dos modelos preditivos desenvolvidos neste trabalho ressalta a importância da integração de múltiplas fontes de dados em estudos sobre soluções digitais na agricultura, além de proporcionar avanços no geoprocessamento de dados por meio de técnicas de inteligência artificial.

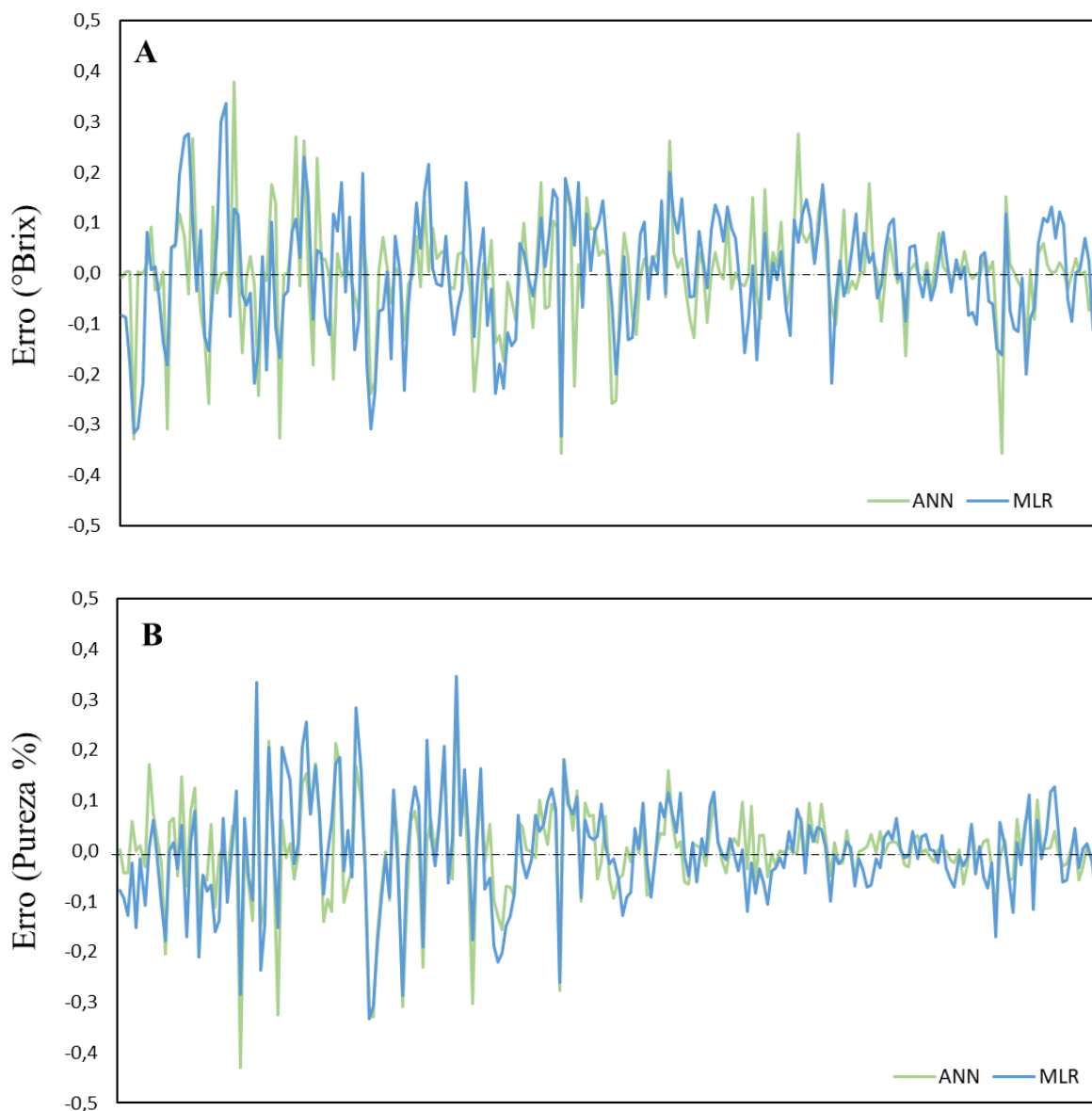


FIGURA 1. Erro de predição dos atributos Brix (A) e Pureza (B) para cada modelo preditivo

**CONCLUSÕES:** Os modelos preditivos desenvolvidos neste estudo apresentaram potencial de estimativa dos atributos de qualidade de cana-de-açúcar, de forma especializada, considerando os dados de refletância do dossel da cultura e a variável GDD. O método de MLR apresentou melhor acurácia na estimativa de Brix e a ANN para a estimativa de Pureza de cana-de-açúcar.

**AGRADECIMENTOS:** À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela disponibilidade de bolsas (código 001) aos autores Marcelo R. Barbosa, Jr. e Romário P. de Oliveira. Ao Laboratório de Máquinas e Mecanização Agrícola (LAMMA) da FCAV/UNESP e ao Laboratório de Processos Industriais da FATEC de Jaboticabal/SP.

## REFERÊNCIAS

- BACHI, O. O. S.; SOUZA, J. A. G. C. Minimum threshold temperature for sugar cane growth. In: International society of Sugar Cane Technologists, São Paulo, SP. *Proceedings...* São Paulo, 1978, v. 2, p. 1733-1741.
- BOLFE, E. L.; JORGE, L. A. de C.; DEL'ARCO SANCHES, I.; et al. Precision and digital agriculture: adoption of technologies and perception of Brazilian farmers. *Agriculture*, v. 10, n. 12, p. 653, 2020. 10.3390/agriculture10120653
- CANATA, T. F.; WEI, M. C.; MALDANER, L. F.; MOLIN, J. P. Sugarcane yield mapping using high-resolution imagery data and machine learning technique. *Remote Sensing*, v. 13, n. 2, p. 232, 2021. 10.3390/rs13020232
- CONSECANA. Manual de Instruções: Conselho dos Produtores de Cana-de-Açúcar, Açúcar e Álcool do Estado de São Paulo. Piracicaba - SP: CONSECANA, 112, 2006.
- FERRAZ, M. N.; CORRÊDO, L. P.; WEI, M. C. F.; MOLIN, J. P. Spatial variability mapping of sugarcane qualitative attributes. *Engenharia Agrícola*, v. 39 (edição especial), p. 109-117, 2019. <https://doi.org/10.1590/1809-4430-Eng.Agric.v39nep109-117/2019>
- GITELSON, A. A.; KAUFMAN, Y. J.; MERZLYAK, M. N. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sens. Environ.* v. 58, p. 289–298, 1996. 10.1016/S0034-4257(96)00072-7
- GÜNTHER, F.; FRITSCH S. *The R Journal*, v. 2, n. 1, p. 30-38, 2010.
- ILLUKPITIYA, P.; YULDASHEV, F.; NASIRU, K. Designing Harvesting and Hauling Cost Models for Energy Cane Production for Biorefineries. *Energies*, v. 15, p. 5403, 2022. <https://doi.org/10.3390/en15155403>
- KATAOKA, T.; KANEKO, T.; OKAMOTO, H.; HATA, S. Crop growth estimation system using machine vision. In: *Proceedings 2003 IEEE/ASME International Conference on Advanced Intelligent Mechatronics*. Kobe, Japan: IEEE b1079–b1083. 2003. 10.1109/AIM.2003.1225492
- MALDANER, L. F.; CANATA, T. F.; MOLIN, J. P. An Approach to Sugarcane Yield Estimation Using Sensors in the Harvester and ZigBee Technology. *Sugar Tech*, v. 24, p. 813–821, 2022. 10.1007/s12355-021-01050-x
- NAWI, N. M.; GUANGNAN, C.; TROY, J. In-field measurement and sampling technologies for monitoring quality in the sugarcane industry: A review. *Precision Agriculture*, v. 15, p. 684–703. 2014. 10.1007/s11119-014-9362-9
- NETER, J.; KUTNER, M. H.; NACHTSHEIM, C. J.; WASSERMAN, W. *Applied linear statistical models*. McGraw-Hill/Irwin, v. 318, 1996.
- PHETPAN, K.; VASU, U.; PANMANAS, S. An online visible and near-infrared spectroscopic technique for the real-time evaluation of the soluble solids content of sugarcane billets on an elevator conveyor. *Computers and Electronics in Agriculture*, v. 154, p. 460–466. 2018. 10.1016/j.compag.2018.09.033
- ROUSE, J. W.; HAAS, R. H.; SCHELL, J. A.; DEERING, D. W. Monitoring vegetation systems in the great plains with ERTS. In: *Proceedings of the Earth Resources Technology Satellite—1 Symposium*, Washington, DC, USA, 10–14 December 1974, 1974. p. 309–317.
- SIEDLIKA, A.; BARANOWSKI, P.; PASTUSZKA-WOŹNIAK, J.; ZUBIK, M.; KRZYSZCZAK, J. Identification of plant leaf phosphorus content at different growth stages based on hyperspectral reflectance. *BMC Plant Biology*, v. 21, n. 28. 2021. 10.1186/s12870-020-02807-4
- ZARCO-TEJADA, P.; BERJÓN, A.; LÓPEZ-LOZANO, R.; MILLER, J. R.; et al. Assessing vineyard condition with hyperspectral indices: Leaf and canopy reflectance simulation in a row-structured discontinuous canopy. *Remote Sens. Environ.*, v. 99, p. 271–287, 2005. 10.1016/j.rse.2005.09.002