

ESPECTROSCOPIA NIR APLICADA AO MAPEAMENTO DA VARIABILIDADE ESPACIAL DA QUALIDADE DE CANA-DE-AÇÚCAR

Lucas de Paula Corrêdo¹; José Paulo Molin²

¹Engenheiro Agrícola, Professor - Depto. de Agronomia, Universidade Federal de Viçosa, UFV, Viçosa ? MG. lucas.corredo@ufv.br / (31) 9 8581-8748; ²Engenheiro Agrícola, Professor - Depto. de Engenharia de Biossistemas, ESALQ, USP, Piracicaba ? SP

Apresentado no
Congresso Brasileiro de Agricultura de Precisão- ConBAP 2022
Campinas, SP, 09 a 11 de agosto de 2022

RESUMO: Um nova vertente tem se aberto para gestão da variabilidade espacial de lavouras agrícolas a partir de pesquisas buscando mapear a qualidade de produtos agrícolas. A cana-de-açúcar é comercializada com base em atributos de qualidade, como o teor de sólidos solúveis (Brix). Entretanto, não existe um equipamento destinado a esse fim. O objetivo dessa pesquisa foi avaliar a espectroscopia de infravermelho próximo (NIRS) como sensor proximal para estimativa de Brix e aplicações em campo. No primeiro estudo, avaliou-se o potencial da NIRS para prever e mapear o Brix em uma lavoura comercial de cana-de-açúcar. Foram desenvolvidos modelos de regressão por mínimos quadrados parciais (PLS) a partir de amostras de cana desfibrada, e utilizados para produção de mapas de qualidade de uma lavoura comercial a partir de amostras coletadas em campo. No segundo, determinou-se a melhor condição de preparo mínimo de amostras para predição de Brix a partir de NIRS. Foram avaliadas amostras de colmos inteiro, cana desfibrada e caldo cru. No terceiro, avaliou-se um método para mensurar a qualidade da cana-de-açúcar a partir de um sensor NIRS embarcado. O Brix apresentou dependência espacial moderada. Além disso, a cana desfibrada apresentou melhor potencial preditivo do atributo. Por fim, um método de transferência de calibração baseado em amostras de cana desfibrada permitiu produzir modelos para estimativa de Brix com um sensor NIR embarcado no elevador de uma colhedora.

PALAVRAS-CHAVE: Agricultura digital; Espectrômetro portátil; Qualidade tecnológica

NIR SPECTROSCOPY APPLIED TO SPATIAL VARIABILITY MAPPING OF SUGARCANE QUALITY

ABSTRACT: New opportunities have been appearing for spatial variability management of agricultural crops from research aiming to map the quality of agricultural products. Sugarcane is marketed based on quality attributes, such as soluble solids content (Brix). However, there is no equipment designed for this purpose. This research aimed to evaluate the near-infrared spectroscopy (NIRS) as a proximal sensor for Brix estimation and field applications. In the first study, we evaluated the potential of NIRS to predict and map Brix in a commercial sugarcane crop. Partial least squares (PLS) regression models were developed from defibrated cane samples and used to produce quality maps of a commercial crop from field-collected samples. In the second, we determined the best minimum sample preparation condition for Brix prediction from NIRS. Billets, defibrated cane, and raw juice samples were evaluated. In the third, we evaluated a method to measure sugarcane quality from an on-board NIRS sensor. Brix showed moderate spatial dependence. Moreover, the defibrated cane showed a better predictive potential of the attribute. Finally, a calibration transfer method based on defibrated cane samples allowed to produce models for Brix estimation with an on-board NIR sensor on a harvester elevator.

KEYWORDS: Digital agriculture; Portable spectrometer; Technological quality

INTRODUÇÃO: Sensores para estimativa de atributos relacionados à qualidade de produtos agrícolas têm sido avaliados e implementados desde linhas de produção em setores industriais até iniciativas em âmbito de produção agrícola (WALSH *et al.*, 2020). Com aplicações avançadas no setor industrial, a espectroscopia no infravermelho próximo (NIRS) apresenta-se como a melhor alternativa devido à precisão das medições, realização de análises não destrutivas e rápidas, facilidade de operação, custo reduzido e sustentabilidade da análise (CORTÉS *et al.*, 2019). Produtos cujo valor agregado está diretamente atrelado a atributos de

qualidade, como a cana-de-açúcar, justificam o esforço para monitorar a variabilidade espacial desses parâmetros nas lavouras a fim de realizar a gestão localizada visando obter maiores retornos financeiros. Assim, o objetivo principal desse trabalho foi investigar métodos para monitorar a variabilidade espacial do teor de sólidos solúveis (Brix) da cana-de-açúcar com base em espectroscopia de infravermelho próximo (NIRS). A espectroscopia na região infravermelho próximo é uma técnica analítica de espectroscopia vibracional a partir da qual a interação da radiação ocorre principalmente através de excitação, combinações e ressonâncias de modos vibracionais fundamentais de grupos funcionais moleculares contendo átomos como C, N, O e S ligados a um átomo de hidrogênio, os principais átomos e moléculas presentes em compostos orgânicos (PASQUINI, 2018). Sensores baseados neste princípio físico foram amplamente adaptados para aplicação em linhas na produção industrial de alimentos, para mensuração de atributos de qualidade de diversas matrizes tais como amostras sólidas intactas, pastosas e fluidas (POREP; KAMMERER; CARLE, 2015). Em usinas de cana-de-açúcar são utilizados em análises de rotina para medir o teor de sacarose aparente (Pol) e concentração de sólidos solúveis (Brix) do caldo (SEXTON *et al.*, 2020). Assim, essa tecnologia tem sido objeto de estudo de diversos trabalhos aplicados à cultura da cana-de-açúcar, para aplicações em condições controladas avaliando formas de mensuração para viabilizar a tecnologia como ferramenta de mensuração para gestão da variabilidade espacial da qualidade (NAWI *et al.*, 2013; NAWI; CHEN; JENSEN, 2013; PHETPAN; UDOMPETAIKUL; SIRISOMBOON, 2018; UDOMPETAIKUL; PHETPAN; SIRISOMBOON, 2021) ou em programas de melhoramento de cana (MARAPHUM *et al.*, 2018; PHUPHAPHUD *et al.*, 2020, 2019). Entretanto, poucos estudos avaliaram a técnica de NIRS para aplicação direta em campo, com o objetivo de avaliar potencial de aplicação frente às condições hostis da colheita mecanizada para mensuração da variabilidade espacial da qualidade. Sendo assim, esse estudo teve por objetivos: (i) investigar a dependência espacial de Brix ao longo de uma lavoura comercial com predição baseada em modelos de NIRS; (ii) avaliar diferentes níveis de processamento mínimo de amostras para predição de Brix; e (iii) avaliar um método de mensuração com um sensor NIRS embarcado em uma colhedora para medições espectrais em tempo real.

MATERIAIS E MÉTODOS: A pesquisa foi dividida em três estudos consecutivos. No primeiro, avaliou-se o potencial da NIRS para prever e mapear o Brix em uma lavoura comercial de cana-de-açúcar. No segundo, determinou-se a melhor condição de preparo mínimo de amostras para predição de Brix a partir de NIRS. No terceiro, avaliou-se um método para mensurar a qualidade da cana-de-açúcar a partir de um sensor NIRS embarcado.

A obtenção dos dados de referência foi realizada em um laboratório de qualidade industrial, seguindo os métodos convencionais de análise de qualidade propostos pelo CONSECANA (CONSECANA, 2015). O procedimento de coleta de amostras para obtenção das leituras espectrais correspondentes aos dois primeiros estudos foi realizado simultaneamente à preparação da amostra para análise convencional de qualidade, conduzida tal como descrito:

- Uma sonda oblíqua coletou uma amostra de colmos de cana-de-açúcar na carga transportada por cada caminhão para a indústria da usina para proceder às análises convencionais (Figura 1a);
- A amostra foi moída em um desfibrador mecânico e homogeneizada em uma betoneira adaptada. Em seguida, foram coletadas amostras com 500 g da cana-de-açúcar desfibrada homogeneizada (Figura 1b);
- A amostra de 500 g foi prensada em uma prensa hidráulica sob pressão constante, a 24,5 MPa (250 kgf cm⁻²) durante 60 s, para obter o caldo cru para análise convencional (Figura 1c);
- O teor de sólidos solúveis (Brix) foi determinado vertendo caldo bruto em um refratômetro (RX-5000 α , ATAGO Co Ltd., Tokyo, Japão).

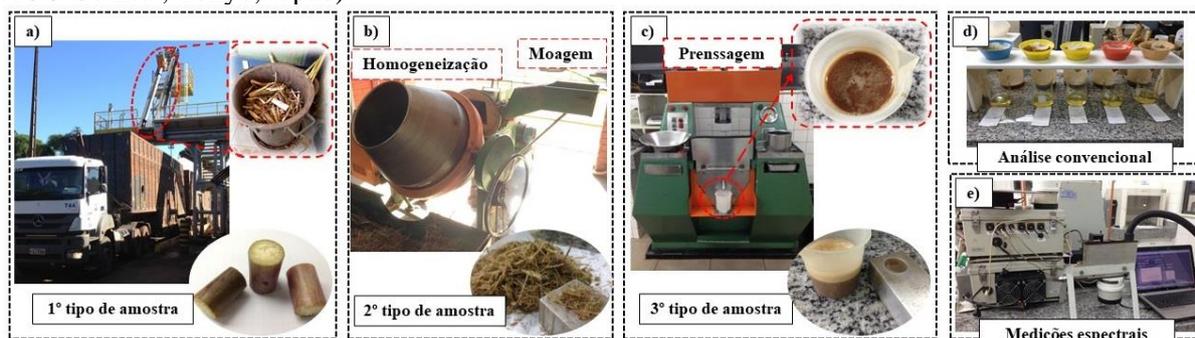


FIGURA 1. Sequência de aquisição de cada tipo de amostra durante as etapas de preparação das amostras para análise convencional: Amostragem de colmos de cana de açúcar por sonda oblíqua no caminhão de carga (a) e colmos de cana de açúcar para medições na superfície e no corte transversal transversal (b); moagem e homogeneização da cana de açúcar e obtenção de amostra de cana desfibrada; prensagem de amostra desfibrada para extração de caldo e caldo bruto extraído (c); análise convencional (d) e obtenção de dados espectrais de cada tipo de amostra (e). **Sequence of acquisition of each sample type for during the sample preparation steps for conventional analysis: Sampling of sugarcane stalks by oblique probe in the cargo truck (a) and sugarcane stalks for surface and cross-sectional measurements (b); milling and homogenization of sugar cane and obtaining shredded cane sample; pressing of shredded sample for juice extraction and extracted raw juice (c); conventional analysis (d) and obtaining spectral data for each type of sample (e).**

Para o primeiro estudo, conduzido na safra de 2018/19, foram coletadas 350 amostras de cana desfibrada homogeneizada (Figura 1b) para desenvolvimento dos modelos de calibração. Cada amostra foi composta por três subamostras de 50 g de cana desfibrada, as quais foram submetidas a leituras espectrais, com cinco repetições, a fim de obter um espectro médio representativo. As leituras espectrais foram registradas como absorvância medida com um espectrômetro portátil AgriNIR (Dynamica Generale S.p.A., Poggio Rusco, Mantova, Itália), operando na faixa de 1100-1800 nm com resolução espectral de 10 nm. Para validação deste estudo, foram coletadas 90 amostras georreferenciadas ao longo de uma área comercial, de forma aleatória (Figura 2), provenientes de um banco de dados previamente coletado (FERRAZ *et al.*, 2019). Foram coletados quatro colmos, cortados ao nível do solo, e removidas as folhas. As amostras foram levadas ao laboratório e processadas seguindo o protocolo de análise do CONSECANA, conforme descrito anteriormente. O mesmo equipamento NIR e procedimentos de análise foram utilizados para coleta de dados de calibração e validação.

Para o segundo estudo, conduzido durante seis meses distribuídos ao longo da safra 2019/20, foram coletados três tipos de amostra. Buscou-se inserir variação dos teores nos modelos de calibração por meio de dados coletados em diferentes estações ao longo da safra (PAGANI *et al.*, 2017), de junho a dezembro. Inicialmente, após a etapa de sondagem, foram coletados três colmos selecionados aleatoriamente, seccionados e levemente limpos (Figura 1a). Em seguida, após a moagem e homogeneização, foi coletada uma subamostra de 50 g de cana desfibrada (Figura 1b). Por fim, após a prensagem, foi coletada uma subamostra de 50 mL de caldo cru (Figura 1c). As leituras foram realizadas, em triplicata, em três pontos equidistantes na casca, de acordo com método adaptador a partir do proposto por Nawi *et al.* (2013), e na secção transversal de cada colmo, conforme proposto por Nawi, Chen e Jensen (2013), e em três subamostras de cana desfibrada e caldo cru. Foram coletados 302 espectros médios para cada tipo de amostra, correspondentes a amostras coletadas em cada carga, divididas em conjunto de calibração (227 amostras, 75%) e validação (75 amostras, 25%) pelo método de Kennard-Stone (KENNARD; STONE, 1969). As leituras espectrais do segundo estudo foram obtidas como absorvância mensurada com um espectrômetro Veris vis-NIR (Veris Technologies Inc., Salina, KS, EUA), em modo de operação em bancada, operando na faixa espectral de 373 a 2222 nm, com resolução média de aproximadamente 5 nm (Figura 1e). Toda a aquisição de dados espectrais de ambos os estudos foi realizada em ambiente com temperatura controlada (20 ± 5 °C), imediatamente após a obtenção de cada amostra, com o intuito de minimizar os efeitos da degradação do açúcar por microrganismos.

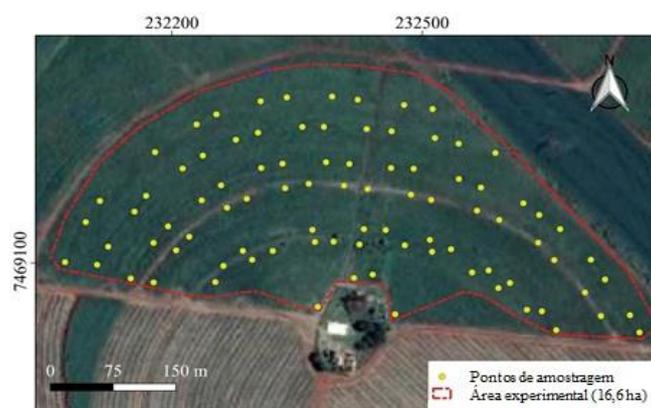


FIGURA 2. Pontos de coleta de amostras georreferenciadas em lavoura comercial.

O terceiro estudo foi executado durante a safra 2020/21, em três dias consecutivos, em três áreas de uma lavoura comercial, com a cultivar CTC 4 em segundo corte. Foi desenvolvida uma plataforma para obtenção de dados espectrais com um micro-espectrômetro (microNIR 1700 spectrometer, Viavi Solutions, JDSU Corporation, Milpitas, CA) embarcado no elevador de uma colhedora de cana-de-açúcar, imediatamente antes do extrator secundário, conforme descrito por NAWI, CHEN e JENSEN (2014). Como fonte de radiação eletromagnética, foram utilizadas quatro lâmpadas halógenas (tipo AR111) com tensão de 12 V e potência de 50 W (PHETPAN; UDOMPETAIKUL; SIRISOMBOON, 2018). As lâmpadas foram montadas de forma a irradiar a região imediatamente sob o sensor. Todo o sistema embarcado foi coberto por uma caixa de aço galvanizado com o interior pintado em preto fosco, e o exterior pintado em branco, para minimizar as interferências da iluminação externa. A plataforma foi fixada ao elevador por meio de quatro amortecedores de vibração de borracha.

Antes das medições em tempo real, foi realizada a calibração do sensor NIR em condições de campo e otimização da altura em relação ao alvo. Um aumento da distância operacional poderia causar uma queda na relação sinal/ruído do espectro devido a uma diminuição da intensidade luminosa que atinge a amostra. Para isso, o tempo de integração do espectrômetro precisa ser devidamente otimizado para contabilizar a perda em intensidade fotométrica. Uma placa de sulfato de bário branco quadrado (BaSO_4), com 0,40 m, foi utilizada como padrão de referência para calibrar o sensor, assumindo uma reflectância uniforme igual a um (100%) (MAYRINK *et al.*, 2019), enquanto uma referência escura foi obtida a partir do ambiente de medição da câmara. As luzes foram ligadas para calibração branca e desligadas para a referência escura (ambiente da câmara). A placa branca foi posicionada sobre a esteira do elevador, e a calibração foi otimizada para 50 leituras médias e tempo de integração de 35 ms (0,7 ms por ponto amostral), resultando num tempo de medição de 1,75 s para obter cerca de 50000 a 60000 contagens brutas de intensidade fotométrica, como recomendado pelo fabricante. Sendo assim, o sensor foi fixado a uma altura de 200 mm entre a janela de safira e o fundo da correia transportadora e 40 mm sobre as aletas da correia transportadora do elevador. A obtenção do posicionamento de cada espectro coletado foi realizada com um receptor GNSS SMART6-L™ (Novatel Inc., Calgary, Alberta, Canadá) com correção TerraStar-C, que permite a precisão de $\pm 0,09$ m. O GNSS foi fixado sobre a cabine da colhedora na direção do corte de base. Todo o sistema foi alimentado pela bateria da colhedora.

A colheita foi realizada a uma velocidade de 1 m s^{-1} , e a velocidade da correia transportadora do elevador a uma velocidade de 2 m s^{-1} . O tempo de alimentação foi avaliado visualmente para correção e *offset* da posição das leituras espectrais ao centro de cada linha colhida. Durante a colheita, foram coletadas amostras aleatoriamente nas três áreas durante as medições com o sistema embarcado para desenvolvimento dos modelos de calibração e validação. Para realizar a amostragem, o transbordo avançou à frente da colhedora e a amostra foi lançada diretamente sob uma lona no solo. Em seguida, o ponto colheita foi registado com um receptor GNSS portátil (Garmin 62 s, precisão de 5 m). A amostragem foi realizada com um espaçamento mínimo de duas fileiras de cana colhida. Foram coletadas 66 amostras para as três áreas, armazenadas em sacos plásticos e enviadas para o laboratório, onde foram armazenadas em congelador a $2 \text{ }^\circ\text{C}$ para minimizar os efeitos da degradação dos compostos orgânicos por microorganismos e analisados no dia seguinte para determinar o Brix seguindo o protocolo do CONSECANA, como descrito para os estudos anteriores. Além disso, foram adquiridos os espectros de subamostras de cana desfibrada em bancada para as mesmas 66 amostras, registrados como a média de três mensurações em cada subamostra. As amostras foram segmentadas em conjunto de calibração (47 amostras) e validação (19 amostras), compostos por 75% e 25% de amostras de cada área de coleta, respectivamente.

Diferentes técnicas de pré-processamento de dados foram aplicadas aos espectros antes da calibração. O ajuste do pré-processamento foi avaliado pelo valor mais baixo do erro quadrático médio (RMSE) de validação cruzada. Adicionalmente, para o terceiro estudo, foi aplicado o método de transferência de calibração *piecewise direct standardization* (PDS) (WANG; VELTKAMP; KOWALSKI, 1991), com o intuito de desenvolver uma matriz de correção dos espectros coletados em tempo real (JI; VISCARRA ROSSEL; SHI, 2015; NAWAR; MUNNAF; MOUAZEN, 2020). Após a realização de testes preliminares, foi pré-definida uma janela de 9 comprimentos de onda para execução do algoritmo PDS. Em seguida, os modelos de calibração foram executados baseados em regressões por mínimos quadrados parciais (PLS). Nos dois primeiros estudos, a calibração foi desenvolvida utilizando o método de *venetian blinds* com 10 *splits*, e no terceiro estudo foi utilizado o método *leave-one-out*, devido ao menor número de amostras. O ajuste do modelo PLS foi

determinado com base no menor número de variáveis latentes (LV), em que o erro quadrático médio (RMSE) de validação cruzada não foi substancialmente superior ao RMSE mínimo observado.

Para avaliação de desempenho dos modelos, nos dois primeiros estudos, além do RMSE para avaliação da acurácia, foi utilizado o coeficiente de determinação (R^2). No terceiro estudo, também foram avaliados o RMSE relativo (RRMSE) e o desempenho para o intervalo interquartil (RPIQ).

Para o primeiro e terceiro estudo, foram realizadas análises geoestatísticas para a base de dados georreferenciados contendo valores observados e preditos por espectroscopia NIR. Os semivariogramas foram ajustados para estimar a estrutura da variação espacial, selecionados com base no menor valor de RMSE. A dependência espacial foi classificada de acordo com método descrito por Cambardella et al. (1994). Os mapas foram produzidos através da interpolação dos dados utilizando krigagem ordinária. Em seguida, os atributos de qualidade da cana de açúcar foram mapeados. A similaridade entre os mapas observado e predito, no primeiro estudo, foi avaliada por comparação visual e pelo coeficiente de correlação (COLAÇO *et al.*, 2019).

Para o segundo estudo, foi realizado um teste de aleatorização (VAN DER VOET, 1994) com nível de significância de 0,05. O objetivo foi comparar a precisão dos modelos de regressão utilizando diferentes tipos de amostras de cana-de-açúcar no conjunto de validação. As hipóteses avaliadas foram nula (H_0) para acurácia (RMSE) similar, e hipótese alternativa (H_1) para acurácia não similares.

RESULTADOS E DISCUSSÃO: Os dados espectrais de mensurações realizadas em amostras de cana desfibrada no primeiro estudo apresentaram melhor acurácia na validação cruzada (menor RMSE) quando foram pré-processados por meio do método *standard normalize variation* (SNV) (BARNES; DHANOA; LISTER, 1989) seguido pela segunda derivação usando o algoritmo Savitzky-Golay (SAVITZKY; GOLAY, 1964), com janela de nove bandas espectrais e polinômio de grau dois.

Os valores para Brix correspondentes às amostras de calibração possibilitaram representar as amostras obtidas no campo, com range de variação daquelas (valores máximos e mínimos) superior ao destas (Tabela 1). Além disso, os valores de validação externa concentraram-se mais próximo dos valores mínimos observados para o conjunto de calibração.

TABELA 1. Estatística descritiva para as amostras dos conjuntos de calibração e validação externa. **Descriptive statistics of the samples for the calibration and external validation sets.**

	N	Máximo	Média	Mínimo	Mediana	SD	CV
Calibração	350	23,80	20,46	13,80	20,73	1,47	7,22
Validação	90	19,20	17,48	15,70	17,60	0,69	3,99

N – número de amostras; SD – desvio padrão; CV – coeficiente de variação.

Um modelo apropriado deve apresentar baixo valor de RMSE e elevado R^2 , para etapas de calibração e validação externa (NAWI; CHEN; JENSEN, 2013). Os resultados obtidos para os modelos de predição de Brix a partir de amostras de cana desfibrada foram satisfatórios (Figura 3a), apresentando RMSE de 0,66 e 0,80, e R^2 de 0,80 e 0,71 para validação cruzada e validação externa (predição). Usando espectros de absorvância da superfície externa de colmos de cana-de-açúcar, Nawi *et al.* (2013) obtiveram resultados de RMSE de predição de 1,51%, valor maior do que os obtidos no presente estudo.

Os mapas resultantes da interpolação dos resultados obtidos a partir de análises convencionais (Figura 3b) e os mapas resultantes da interpolação dos resultados preditos pelo modelo desenvolvido baseado em NIRS (Figura 3c) mostraram similaridade espacial substancial, baseado em análise visual, resultado do desempenho satisfatório apresentado pelos modelos de predição. Distingue-se uma região com baixo potencial para Brix quando comparada com outras regiões do campo, estendendo-se praticamente por toda a região leste em direção à área central superior. Por outro lado, existe uma região, na área central inferior do talhão, que apresentou um elevado teor de Brix em relação às outras áreas do talhão. Os resultados evidenciam a importância para o mapeamento de atributos de qualidade da cana-de-açúcar, uma vez que existe uma distribuição desigual dos valores Brix ao longo da área. Os modelos geoestatísticos apresentaram valores de

efeito pepita igual a 0,30 e 0,28, alcance de 132,4 e 122,5, e índice de dependência espacial de 62,9 e 52,6% para valores observados e preditos, respectivamente, apresentando dependência espacial moderada (CAMBARDELLA *et al.*, 1994).

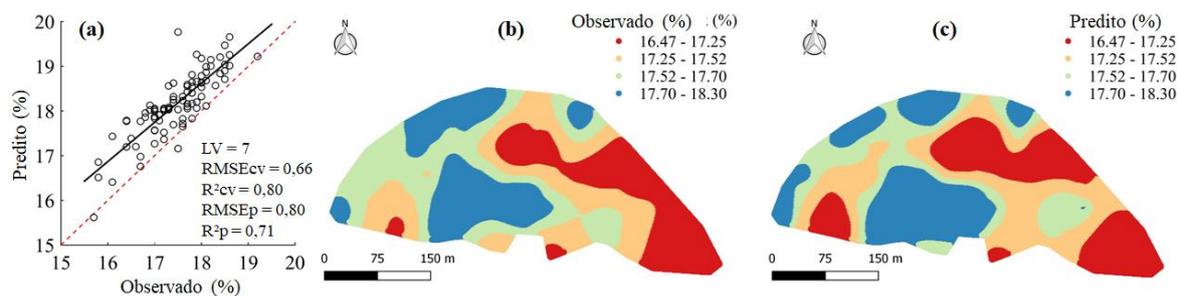


FIGURA 3. Gráfico de dispersão de teores de conteúdo de sólidos solúveis (Brix) do conjunto de dados de validação externa (a). Mapas de Brix observado (b) e predito (c) baseado em regressão por mínimos quadrados parciais a partir de dados de NIRS. **Scatter plot of measured versus predicted solids soluble content (Brix) of external validation data set. Maps of Brix measured (b) and predicted (c) based on partial least squares (PLS) regression from NIRS data.**

A utilização de NIRS proporcionou um método prático, rápido, preciso e confiável para predição de Brix. Acredita-se que a preparação mínima deve resolver a questão da variabilidade do conteúdo ao longo do caule da cana, uma vez que proporciona uma amostragem mais uniforme e homogênea. Embora a colheita da cana-de-açúcar seja um ambiente particularmente desafiador e a ausência de preparação de amostras seja um aspecto desejável para o desenvolvimento de um sensor embarcado. Alguns pesquisadores alcançaram resultados satisfatórios na predição do Brix em condições controladas baseados em mensurações de colmos (PHETPAN; UDOMPETAIKUL; SIRISOMBOON, 2018; UDOMPETAIKUL; PHETPAN; SIRISOMBOON, 2021). No entanto, é necessário avançar para aplicações no campo, principalmente quanto ao método de amostragem e técnicas de processamento dos dados espectrais coletados.

No segundo estudo, os resultados para análise convencional das amostras obtidas ao longo de uma safra, onde buscou-se avaliar diferentes tipos de amostra quanto ao potencial preditivo, permitiram obter grande variação nos teores de Brix (Tabela 2). O algoritmo de Kennard-Stone permitiu a seleção de um conjunto de dados representativo da população de estudo quanto à distribuição de teores mínimos e máximos.

TABELA 2. Estatística descritiva dos resultados de referência de Brix de todas as amostras, conjuntos de calibração e validação externa. **Descriptive statistics of the reference results for Brix of all samples, calibration, and external validation data sets.**

	N	Máximo	Média	Mínimo	Mediana	SD	CV
Todos	302	23,42	18,95	13,08	18,99	1,71	9,02
Calibração	227	23,42	18,86	13,08	18,80	1,66	8,80
Validação externa	75	23,05	19,24	13,55	19,59	1,85	9,62

N – número de amostras; SD – desvio padrão; CV – coeficiente de variação.

Houve uma subestimativa de valores elevados de Brix e uma superestimativa de valores mais baixos para todos os tipos de amostras avaliados (Figura 4). Entretanto, este efeito foi mais intenso para as condições de amostras sem preparo, tais como para dados obtidos por mensuração na casca (SS) e na secção transversal (CSS) dos colmos, quando comparado às amostras preparadas (cana desfibrada – DF, e caldo cru – RJ).

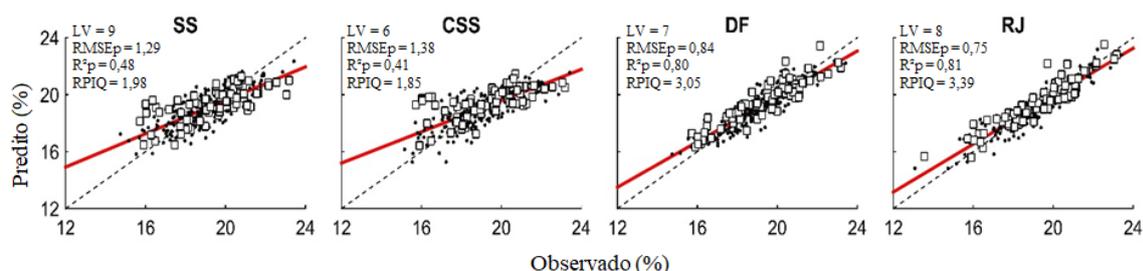


FIGURA 4. Gráficos de valores observados versus preditos de Brix de cana-de-açúcar a partir de espectroscopia vis-NIR com medições na casca (SS) e secção transversal (CSS), cana desfibrada (DF) e caldo cru (RJ). **Plots of observed values versus predicted sugarcane Brix from vis-NIR by skin (SS) and cross-sectional (CSS) scanning of billets, defibrated (DF), and raw juice (RJ) samples.**

Não houve diferença significativa entre a acurácia dos modelos (p -valor $> 0,05$) ao avaliar-se comparativamente todas os tipos de amostra, exceto para comparação entre SS e CSS (Tabela 3). Além disso, o desempenho apresentado para predição de Brix a partir de amostras de DF e RJ foi similar, com valores de R^2_p e RPIQ próximos, e menor número de variáveis latentes (conferindo menor complexidade aos modelos), evidenciando que o preparo da cana por moagem e homogeneização é o suficiente para determinação do atributo de qualidade. Amostras sem preparo, como SS e CSS, não apresentaram diferença significativa na acurácia quando comparadas a amostras de DF e RJ (Tabela 3). Entretanto, o desempenho caracterizado pelos valores de R^2 e RPIQ foi inferior (Figura 4).

TABELA 3. p -Valores obtidos com o teste de aleatorização para o conjunto de dados de validação externa para comparação binária entre os tipos de amostra avaliados. **p -Values of randomization test of external validation set for all compared sugarcane sample types.**

Comparação binária entre tipos de amostras						
	SS vs. CSS	SS vs. DF	SS vs. RJ	CSS vs. DF	CSS vs. RJ	DF vs. RJ
p -valor	0,104	1,00	1,00	1,00	1,00	0,879

Em estudo precursor, Nawi et al. (2013) obtiveram valores de 0,87 para o R^2_p . O desempenho satisfatório desta métrica pode ser explicado pelo método de aquisição de dados adotado pelos autores, com amostras individualizadas de acordo com a porção do caule (porção inferior, média e superior). O acúmulo de sacarose ocorre de forma ascendente, com mais concentração de açúcares na porção inferior e menos nos entrenós da porção superior, perto das folhas (WANG, *et al.*, 2017). Contudo, se analisarmos as características de uma colhedora de cana-de-açúcar, após os caules passarem pelo sistema de rolos picadores, a distinção entre as porções do caule não é viável. Por outro lado, o preparo de amostras desfibradas constitui-se um método destrutivo e que pode ser impraticável para uma adaptação visando obter mensurações em tempo real. Assim, é necessário avançar em estudos aplicando técnicas de processamento de dados multivariados, como transferência de calibração, a fim de possibilitar a obtenção de dados de qualidade sem preparo de amostras.

No terceiro estudo, o algoritmo de Kennard-Stone para segmentação de dados permitiu obter um conjunto de calibração cujo intervalo de valores máximo e mínimo foi representativo do conjunto de dados completo (Tabela 4). Este efeito é desejável para que o conjunto de dados de calibração caracterize a variação existente em todo o conjunto de dados que caracteriza a população (talhões) de estudo.

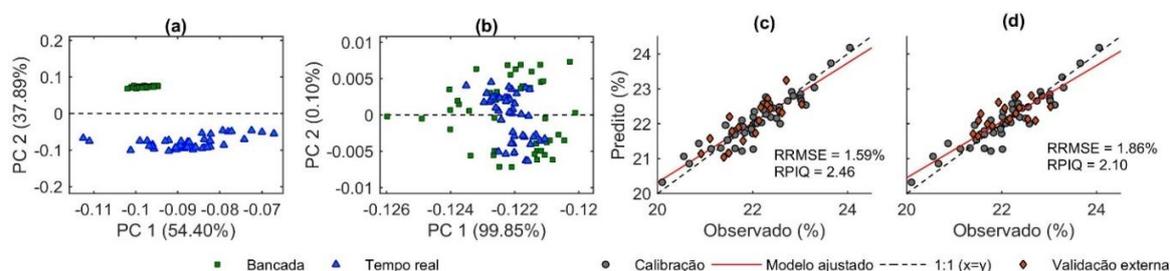
TABELA 4. Estatística descritiva dos resultados de referência de Brix de todas as amostras, conjuntos de calibração e validação externa. **Descriptive statistics of the sugarcane quality parameters for all collected samples, for the cross-validation and prediction data set.**

	N	Máximo	Média	Mínimo	Mediana	SD	CV
Todos	66	24,05	22,09	20,10	22,09	0,73	3,30

Calibração	47	24,05	22,14	20,10	22,11	0,79	3,57
Validação	19	23,03	21,98	20,92	21,94	0,57	2,59

N – número de amostras; SD – desvio padrão; CV – coeficiente de variação.

Por meio de uma análise exploratória de componentes principais (PCA) é possível observar o efeito do método PDS de transferência de calibração sobre os espectros obtidos pelo sensor embarcado (Figura 5). Antes de aplicar o PDS, espectros de amostras de cana desfibrada obtidos em condições controladas apresentaram diferenças evidentes quanto à variância dos dados explicada pela PCA (Figura 5a). A variância dos dados obtidos nas medições em tempo real possivelmente ocorreu devido a fatores ambientais adversos intrínsecos da colheita mecanizada sob medições espectrais quando comparadas às medições em condições controladas. Por outro lado, após a aplicação do modelo PDS nas medições em tempo real, os escores são exibidos no espaço do conjunto mestre, ou seja, os espectros de cana desfibrada (Figura 5b). Além disso, os dois primeiros componentes explicaram 90,73% da variância dos dados antes da transferência de calibração, e após a transferência de calibração, 99,85% da variância de ambos os conjuntos de dados passou a ser explicada apenas pela primeira componente principal. Este resultado indica que a variância devida à diferença entre os dados espectrais correspondentes às amostras de cana desfibrada medidas sob condições controladas e as amostras obtidas com as medições a bordo foi substancialmente reduzida.



Gráfico, Gráfico de dispersão Descrição gerada automaticamente

FIGURA 5. Os escores dos componentes principais para predição por sensor embarcado (em tempo real), e amostras de cana desfibradas medidas em laboratório (bancada), antes (a) e depois (b) da transferência de calibração por PDS (*Piecewise Direct Standardization*), e porcentagem de variação explicada por cada componente mostrada nos eixos entre parênteses. Valores preditos versus valores observados a partir de cana desfibrada (c) e a partir de modelos PLS-PDS para mensuração em tempo real (d). **The principal component scores for on-board field prediction (real time), and defibrated cane samples measured in the laboratory (benchtop), before (a) and after (b) calibration transfer by Piecewise Direct Standardization (PDS), and the percentage of variance explained by each component is shown on the axes in parentheses. Predicted vs. measured values from defibrated cane (c) and from PLS-PDS models from on-board measurements (d).**

Houve substancial similaridade entre a performance de ambos os modelos de predição, utilizando condições controladas, a partir da mensuração de amostras de cana desfibrada em bancada (Figura 5c) e com mensurações em tempo real e utilizando o método PDS para transferência de calibração (Figura 5d), com valores de RPIQ iguais a 2,46 e 2,10 respectivamente. Além disso, a acurácia relativa de ambos os modelos apresentou valores próximos, com RRMSE iguais a 1,86% e 1,59%, para mensurações na colhedora e em bancada, respectivamente. Em estudo precursor, Nawi et al. (2013) alcançaram RRMSEP igual 8,47% e 7,91% para a predição de Brix por meio de espectros de absorvância NIR e de reflectância, respectivamente, para medições realizadas na superfície exterior da casca em condições controladas. Em outro estudo, baseado em medições da secção transversal de amostras de colmos em condições controladas, Nawi, Chen e Jensen (2013) atingiram valores RRMSEP de até 8,13% para a predição de Brix. Os resultados apresentados foram inferiores aos obtidos no presente estudo, embora realizados em condições controladas, evidenciando que o método de transferência de calibração pode ser eficaz para realizar o pós-processamento de dados espectrais coletados em tempo real com sensor embarcado em colhedoras de cana-de-açúcar.

Os teores de Brix apresentaram dependência espacial moderada (CAMBARDELLA *et al.*, 1994) para as três áreas de estudo (Tabela 5). No primeiro estudo, também foi observada dependência espacial moderada a partir do mapa gerado por interpolação de amostras coletadas em campo e analisadas em laboratório. Rodrigues,

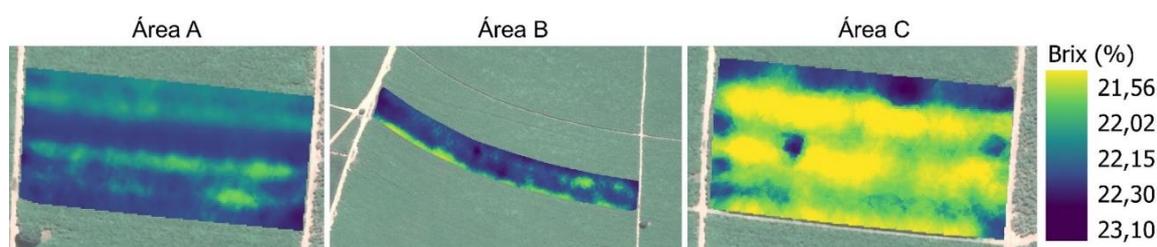
Magalhães e Franco (2013) observaram dependência espacial fraca a moderada para Brix predito a partir de modelos matemáticos ajustados a partir de parâmetros físico-químicos do solo e nitrogênio foliar como variáveis preditoras.

TABELA 5. Parâmetros do semivariograma dos modelos ajustados para análise da variabilidade espacial de Brix das três áreas avaliadas. **Semivariogram parameters of the models fitted to analyze the spatial variability of Brix in the three fields evaluated.**

Área	Modelo	C0	C0 + C1	A (m)	C0 / (C0 + C1)
A	Exponencial	0,060	0,098	33,62	0,61
B	Exponencial	0,252	0,363	64,48	0,69
C	Exponencial	0,381	0,517	38,76	0,74

C0 – efeito pepita; C1 – semivariância estrutural; A – alcance

Os mapas de qualidade da cana-de-açúcar, obtidos a partir do modelo de predição PLS-PDS gerado, mostraram efeitos evidentes da variabilidade nos campos experimentais (Figura 6). Além disso, a variabilidade entre as diferentes áreas é ainda mais evidente. O campo C apresentou uma maior área com baixas concentrações de Brix do que os campos A e B. A modelagem permitiu discriminar valores altos e baixos de Brix, mesmo em áreas pequenas, e especialmente em distâncias maiores. Foi possível observar variações relativas de 7,14% entre valores máximos e mínimos para os campos experimentais. Para avançar na aplicação da técnica de mensuração de qualidade em tempo real torna-se essencial a avaliação em áreas extensas, em diferentes momentos ao longo de uma mesma safra, e entre diferentes safras. A aquisição de uma tal base de dados permitirá o desenvolvimento de modelos de calibração mais robustos, inserindo maior variabilidade dos teores no desenvolvimento dos modelos de predição, e permitindo a avaliação de técnicas mais robustas para processamento de dados como métodos não-lineares (SEXTON *et al.*, 2020). Além disso, recomenda-se a avaliação de outras técnicas de transferência de calibração e ortogonalização de dados multivariados. Por fim, é importante a avaliação da interferência da temperatura do ambiente ao longo do dia, uma vez que esse parâmetro pode contribuir com vibrações moleculares dos compostos orgânicos (PASQUINI, 2018) afetando a qualidade do sinal mensurado. A utilização deste atributo combinado com dados de produtividade, pode permitir a aquisição de mapas de variabilidade da quantidade de açúcares produzidos por área. Constitui uma informação valiosa para a gestão agrônômica dos insumos agrícolas, visando a qualidade do produto fornecido às usinas e para a gestão financeira de toda a operação.



Tela de computador com texto preto sobre fundo branco Descrição gerada automaticamente com confiança média

FIGURA 6. Mapas de variabilidade espacial de Brix da cana-de-açúcar para os três campos experimentais medidos a partir de sensor NIR embarcado na colhedora e preditos pelo modelo de regressão PLS combinado com o método de transferência de calibração *piecewise direct standardization*. **Sugarcane Brix spatial variability maps for the three experimental fields measured by on-board NIR sensor on the harvester and predicted by PLS regression model combined with piecewise direct standardization calibration transfer method.**

CONCLUSÃO: (i) O teor de sólidos solúveis da cana-de-açúcar apresentou dependência espacial moderada, quando estimado por meio de amostragem e mensuração por NIRS em cana desfibrada; (ii) Não houve diferença estatística significativa na acurácia (RMSE) de mensuração de Brix para amostras sem preparo, como

colmo inteiro, quando comparado a amostras desfibradas e de caldo, entretanto, o potencial preditivo (RPIQ e R²) de amostras desfibradas foi superior ao das demais formas; e (iii) O método de transferência de calibração *piecewise direct standardization* permitiu realizar o pós-processamento de mensurações espectrais realizadas em tempo real com sensor embarcado, baseado em mensurações realizadas em condições controladas com amostras de cana desfibrada, identifica a dependência espacial moderada do Brix, e assim mapear a variabilidade espacial do atributo de qualidade em áreas de uma lavoura comercial.

AGRADECIMENTOS:

O estudo foi financiado por meio do processo nº 2018/25008-8, Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), referente ao doutoramento do primeiro autor. Agradecemos ao grupo São Martinho e à Spectral Solutions por prover a infraestrutura e equipamentos para a aquisição de dados do experimento.

REFERÊNCIAS

BARNES, R. J.; DHANOA, M. S.; LISTER, S. J. Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. **Applied Spectroscopy**, vol. 43, no. 5, p. 772–777, 1989. <https://doi.org/10.1366/0003702894202201>.

CAMBARDELLA, C. A.; MOORMAN, T. B.; NOVAK, J. M.; PARKIN, T. B.; KARLEN, D. L.; TURCO, R. F.; KONOPKA, A. E. Field-Scale Variability of Soil Properties in Central Iowa Soils. **Soil Science Society of America Journal**, vol. 58, no. 5, p. 1501–1511, 1994. <https://doi.org/10.2136/sssaj1994.03615995005800050033x>.

COLAÇO, A. F.; MOLIN, J. P.; ROSELL-POLO, J. R.; ESCOLÀ, A. Spatial variability in commercial orange groves. Part 2: relating canopy geometry to soil attributes and historical yield. **Precision Agriculture**, vol. 20, no. 4, p. 805–822, 2019. DOI 10.1007/s11119-018-9615-0. Available at: <https://doi.org/10.1007/s11119-018-9615-0>.

CONSECANA. National Concill of Sugarcane Producers of Sa?o Paulo State. vol. 6, p. 81, 2015. .

CORTÉS, V.; BLASCO, J.; ALEIXOS, N.; CUBERO, S.; TALENS, P. Monitoring strategies for quality control of agricultural products using visible and near-infrared spectroscopy: A review. **Trends in Food Science and Technology**, vol. 85, p. 138–148, 1 Mar. 2019. <https://doi.org/10.1016/j.tifs.2019.01.015>.

FERRAZ, M. N.; CORRÊDO, L. P.; WEI, M. C. F.; MOLIN, J. P. Spatial variability mapping of sugarcane qualitative attributes. **Engenharia Agrícola**, vol. 39, no. spe, p. 109–117, 2019. DOI 10.1590/1809-4430-eng.agric.v39nep109-117/2019. Available at: <http://dx.doi.org/10.1590/1809-4430-Eng.Agric.v39nep109-117/2019>.

JI, W.; VISCARRA ROSSEL, R. A.; SHI, Z. Improved estimates of organic carbon using proximally sensed vis-NIR spectra corrected by piecewise direct standardization. **European Journal of Soil Science**, vol. 66, no. 4, p. 670–678, 1 Jul. 2015. DOI 10.1111/ejss.12271. Available at: <https://onlinelibrary.wiley.com/doi/full/10.1111/ejss.12271>.

KENNARD, R. W.; STONE, L. A. Computer Aided Design of Experiments. **Technometrics**, vol. 11, no. 1, p. 137–148, 1969. <https://doi.org/10.1080/00401706.1969.10490666>.

MARAPHUM, K.; CHUAN-UDOM, S.; SAENGPRACHATANARUG, K.; WONGPICHET, S.; POSOM, J.; PHUPHAPHUD, A.; TAIRA, E. Effect of waxy material and measurement position of a sugarcane stalk on the rapid determination of Pol value using a portable near infrared instrument. **Journal of Near Infrared Spectroscopy**, vol. 26, no. 5, p. 287–296, 2018. DOI 10.1177/0967033518795810. Available at: <https://journals.sagepub.com/doi/pdf/10.1177/0967033518795810>.

- MAYRINK, G. O.; VALENTE, D. S. M.; QUEIROZ, D. M.; PINTO, F. A. C.; TEOFILLO, R. F. Determination of chemical soil properties using diffuse reflectance and ion-exchange resins. **Precision Agriculture**, vol. 20, p. 541–561, 2019. DOI 10.1007/s11119-018-9597-y. Available at: <https://doi.org/10.1007/s11119-018-9597-y>.
- NAWAR, S.; MUNNAF, M. A.; MOUAZEN, A. M. Machine learning based on-line prediction of soil organic carbon after removal of soil moisture effect. **Remote Sensing**, vol. 12, no. 8, p. 1308, 1 Apr. 2020. DOI 10.3390/RS12081308. Available at: www.mdpi.com/journal/remotesensing.
- NAWI, N. M.; CHEN, G.; JENSEN, T. In-field measurement and sampling technologies for monitoring quality in the sugarcane industry: a review. **Precision Agriculture**, vol. 15, no. 6, p. 684–703, 2014. <https://doi.org/10.1007/s11119-014-9362-9>.
- NAWI, N. M.; CHEN, G.; JENSEN, T. Visible and shortwave near infrared spectroscopy for predicting sugar content of sugarcane based on a cross-sectional scanning method. **Journal of Near Infrared Spectroscopy**, vol. 21, no. 4, p. 289–297, 2013. <https://doi.org/10.1255/jnirs.1060>.
- NAWI, N. M.; CHEN, G.; JENSEN, T.; MEHDIZADEH, S. A. Prediction and classification of sugar content of sugarcane based on skin scanning using visible and shortwave near infrared. **Biosystems Engineering**, vol. 115, p. 154–161, 2013. DOI 10.1016/j.biosystemseng.2013.03.005. Available at: <http://dx.doi.org/10.1016/j.biosystemseng.2013.03.005>.
- PAGANI, V.; STELLA, T.; GUARNERI, T.; FINOTTO, G.; VAN DEN BERG, M.; MARIN, F. R.; ACUTIS, M.; CONFALONIERI, R. Forecasting sugarcane yields using agro-climatic indicators and Canegro model: A case study in the main production region in Brazil. **Agricultural Systems**, vol. 154, p. 45–52, 1 Jun. 2017. <https://doi.org/10.1016/j.agsy.2017.03.002>.
- PASQUINI, C. Near infrared spectroscopy: A mature analytical technique with new perspectives – A review. **Analytica Chimica Acta**, vol. 1026, p. 8–36, Oct. 2018. DOI 10.1016/j.aca.2018.04.004. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0003267018304793>.
- PHETPAN, K.; UDOMPETAIKUL, V.; SIRISOMBOON, P. 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**, vol. 154, p. 460–466, Nov. 2018. DOI 10.1016/j.compag.2018.09.033. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0168169918302096>.
- PHUPHAPHUD, A.; SAENGPRACHATANARUG, K.; POSOM, J.; MARAPHUM, K.; TAIRA, E. Non-destructive and rapid measurement of sugar content in growing cane stalks for breeding programmes using visible-near infrared spectroscopy. **Biosystems Engineering**, vol. 197, p. 76–90, 1 Sep. 2020. <https://doi.org/10.1016/j.biosystemseng.2020.06.012>.
- PHUPHAPHUD, A.; SAENGPRACHATANARUG, K.; POSOM, J.; MARAPHUM, K.; TAIRA, E. Prediction of the fibre content of sugarcane stalk by direct scanning using visible-shortwave near infrared spectroscopy. **Vibrational Spectroscopy**, vol. 101, p. 71–80, Mar. 2019. DOI 10.1016/j.vibspec.2019.02.005. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0924203118302613>.
- POREP, J. U.; KAMMERER, D. R.; CARLE, R. On-line application of near infrared (NIR) spectroscopy in food production. **Trends in Food Science and Technology**, vol. 46, no. 2, p. 211–230, 1 Dec. 2015. <https://doi.org/10.1016/j.tifs.2015.10.002>.
- RODRIGUES, F. A.; MAGALHÃES, P. S. G.; FRANCO, H. C. J. Soil attributes and leaf nitrogen estimating sugar cane quality parameters: Brix, pol and fibre. **Precision Agriculture**, 2013. <https://doi.org/10.1007/s11119-012-9294-1>.

SAVITZKY, A.; GOLAY, M. J. E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. **Analytical Chemistry**, vol. 36, no. 8, p. 1627–1639, 1964. DOI 10.1021/ac60214a047. Available at: <https://pubs.acs.org/sharingguidelines>.

SEXTON, J.; EVERINGHAM, Y.; DONALD, D.; STAUNTON, S.; WHITE, R. Investigating the identification of atypical sugarcane using NIR analysis of online mill data. **Computers and Electronics in Agriculture**, vol. 168, 1 Jan. 2020. DOI 10.1016/j.compag.2019.105111. Available at: <https://doi.org/10.1016/j.compag.2019.105111>.

UDOMPETAIKUL, V.; PHETPAN, K.; SIRISOMBOON, P. Development of the partial least-squares model to determine the soluble solids content of sugarcane billets on an elevator conveyor. **Measurement**, vol. 167, p. 107898, 1 Jan. 2021. DOI 10.1016/j.measurement.2020.107898. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S026322412030436X>.

VAN DER VOET, H. Comparing the predictive accuracy of models using a simple randomization test. **Chemometrics and Intelligent Laboratory Systems**, vol. 25, no. 2, p. 313–323, 1994. [https://doi.org/10.1016/0169-7439\(94\)85050-X](https://doi.org/10.1016/0169-7439(94)85050-X).

WALSH, K. B.; BLASCO, J.; ZUDE-SASSE, M.; SUN, X. Visible-NIR ‘point’ spectroscopy in postharvest fruit and vegetable assessment: The science behind three decades of commercial use. **Postharvest Biology and Technology**, vol. 168, p. 111246, 1 Oct. 2020. <https://doi.org/10.1016/j.postharvbio.2020.111246>.

WANG, J.; ZHAO, T.; YANG, B.; ZHANG, S. Sucrose Metabolism and Regulation in Sugarcane. **Journal of Plant Physiology & Pathology**, vol. 05, no. 04, 7 Feb. 2017. <https://doi.org/10.4172/2329-955x.1000167>.

WANG, Y.; VELTKAMP, D. J.; KOWALSKI, B. R. Multivariate Instrument Standardization. **Analytical Chemistry**, vol. 63, no. 23, p. 2750–2756, 1 Dec. 1991. DOI 10.1021/ac00023a016. Available at: <https://pubs.acs.org/doi/abs/10.1021/ac00023a016>.