EEAOC RIAT
Revista Industrial y Agrícola de Tucumán

ISSN 0370-5404

En línea 1851-3018

Tomo 104 (1): 16-20; 2027

EEAOC
ESTACION EXPERIMENTAL AGROINDUSTRIAL OBISPO COLOMBRES
Tucumán | Argentina

Av. William Cross 3150
T4101XAC - Las Talitas.
Tucumán, Argentina.

Trabajo ya publicado

Modelos FT-NIRS para la predicción de variables de calidad en caña desfibrada mediante sistema Bruker CPS

Sorol, Natalia*; Diez, Paula*; Jiménez, Rubén*; Aballay Cristian*; Rementería Álvaro*; Henríquez, Diego**; Ruiz, Marcelo*; Franck Colombres, Federico*** y Ostengo Santiago**

*Sección Química, **Programa de Mejoramiento Genético de Caña de Azúcar, ***Sección Ingeniería. Estación Experimental Agroindustrial Obispo Colombres (EEAOC), Las Talitas, Tucumán, R. Argentina, T4101XAC. Email: nrsorol@eeaoc.org.ar

RESUMEN

La espectroscopía de infrarrojo cercano (NIRS, siglas del inglés de Near Infrared Spectroscopy) constituye una alternativa rápida y no destructiva para la evaluación de la calidad industrial de la caña de azúcar. El objetivo de este trabajo fue desarrollar y validar modelos de calibración NIRS para estimar las principales variables de calidad industrial en caña desfibrada mediante un sistema de presentación de caña (CPS, siglas del inglés de Cane Presentation System) acoplado a un espectrómetro FT-NIR Bruker. Se analizaron 453 muestras correspondientes a 15 genotipos (variedades comerciales y clones promisorios) de caña de azúcar, cosechadas en diferentes zonas de Tucumán durante 2025. Se determinaron mediante métodos analíticos de referencia Brix y Pol en jugo, Pol y fibra en caña, cenizas conductimétricas y azúcares recuperables. Los modelos se desarrollaron mediante regresión por mínimos cuadrados parciales (PLS, siglas del inglés de Partial Least Squares), utilizando un conjunto de muestras para calibración y otro para validación externa. Los mejores resultados predictivos se obtuvieron para Brix y Pol en jugo y Pol en caña, con coeficientes de determinación (R^2) superiores a 0,94 y valores de desvío de predicción residual, RPD (siglas en inglés para Residual Prediction Deviation) mayores a 4. Los modelos de fibra y azúcares recuperables también mostraron un buen desempeño, con valores de RPD entre 2,3 y 3,8, mientras que el modelo de cenizas conductimétricas presentó la menor capacidad predictiva. Los resultados indican que el sistema FT-NIR Bruker CPS constituye una herramienta con potencial para la determinación rápida y no destructiva de variables de calidad industrial en caña desfibrada, y su aplicación como apoyo en programas de mejoramiento genético y evaluación varietal.

Palabras clave: espectroscopía de infrarrojo cercano, calibración multivariada, regresión PLS, mejoramiento genético, análisis no destructivo.

ABSTRACT

FT-NIRS Models for Predicting Quality Parameters in Shredded Sugarcane Using the Bruker CPS System

Near-infrared spectroscopy (NIRS) is a rapid and non-destructive alternative for assessing the industrial quality of sugarcane. The objective of this study was to develop and validate NIRS calibration models to estimate the main industrial quality parameters in shredded sugarcane using a Cane Presentation System (CPS) coupled to a Bruker FT-NIR spectrometer. A total of 453 samples representing 15 sugarcane genotypes (commercial varieties and promising clones), harvested from different locations in Tucumán during 2025, were analyzed. Juice Brix and Pol, cane Pol and fiber, conductivity ash, and recoverable sugars were determined using reference analytical methods. The models were developed using partial least squares (PLS) regression, with separate sample sets for calibration and

Trabajo presentado en la XXIV Reunión Técnica Nacional de la Sociedad Argentina de Técnicos de la Caña de Azúcar (SATCA), 19 y 20 de marzo de 2026, Tucumán, Argentina.

Fecha de recepción: 22/10/2026

Fecha de aceptación: 30/10/2026

external validation. The best predictive results were obtained for juice Brix and Pol and cane Pol, with coefficients of determination (R^2) greater than 0.94 and residual prediction deviation (RPD) values above 4. The fiber and recoverable sugar models also showed good performance, with RPD values ranging from 2.3 to 3.8, whereas the conductivity ash model exhibited the lowest predictive ability. The results indicate that the Bruker FT-NIR CPS system has potential as a tool for the rapid and non-destructive determination of industrial quality parameters in shredded sugarcane and for supporting sugarcane breeding programs and varietal evaluation.

Key words: near-infrared spectroscopy, multivariate calibration, PLS regression, genetic improvement, non-destructive analysis.

INTRODUCCIÓN

La caracterización de la calidad de la caña de azúcar es fundamental en los programas de mejoramiento, constituyendo un pilar clave en la evaluación varietal y el control industrial. La espectroscopía de infrarrojo cercano (NIRS) ofrece una alternativa rápida, no destructiva y sin requerimiento de reactivos frente a los métodos de referencia convencionales, aunque requiere el desarrollo previo de modelos de calibración.

A fines de 2024, la Estación Experimental Agroindustrial Obispo Colombres (EEAOC) adquirió un espectrómetro de infrarrojo cercano por transformada de Fourier que emplea el sistema de presentación de caña (FT-NIR Bruker CPS, por sus siglas en inglés) mediante el subsidio Equipar Ciencia 2, destinado a fortalecer la capacidad analítica del Laboratorio de Investigaciones Azucareras (LIA). Este sistema permite analizar muestras de caña desfibrada de gran volumen (5 a 18 kg), mejorando significativamente la representatividad y reduciendo la variabilidad asociada al muestreo.

En este trabajo se presenta el desarrollo de modelos de calibración y validación para las principales variables de calidad industrial, obtenidos a partir de muestras procesadas en el marco del Subprograma de Mejoramiento Genético de Caña de Azúcar de la EEAOC. El estudio busca evaluar el desempeño del Bruker CPS como herramienta de apoyo para la selección y evaluación de nuevas variedades.

El presente trabajo tiene como objetivo desarrollar modelos de calibración NIRS para estimar variables clave de calidad de caña de azúcar a partir de una base de datos amplia y diversa, y validar su capacidad predictiva en comparación con métodos químicos de referencia.

MATERIALES Y MÉTODOS

Preparación de la muestra

Las determinaciones se realizaron en el LIA entre mayo y septiembre de 2025 en 453 muestras correspondientes a 15 genotipos de caña de azúcar (variedades comerciales y clones promisorios) cosechados manualmente en diferentes zonas de la provincia. Cada muestra se conformó con 10 tallos limpios y despuntados. Los tallos fueron procesados en un desfibrador a martillo obteniéndose caña desfibrada con un porcentaje de células abiertas cercano al 95%. Posteriormente se prensaron en una prensa hidráulica de 60 toneladas que aplica una presión de 250 kg/cm^2 durante 1 minuto para obtener el jugo primario. Las variables de interés se determinaron siguiendo la metodología propuesta por Diez *et al.* (2000).

Determinaciones químicas de referencia

Se determinaron por análisis húmedo las variables utilizadas como referencia para el desarrollo de los modelos de calibración NIRS:

- Brix refractométrico en jugo (BxRef %), se midió con un refractómetro digital marca KEM RA-620 (Chen y Chou, 1993).
- Pol en jugo (%), se determinó utilizando subacetato de plomo como clarificante y midiendo en un polarímetro digital Anton Paar modelo MCP 500 Sucromat (Spencer-Meade, 1967a). El cálculo se realizó mediante la ecuación 1.
- Pol en caña (%), se calculó mediante la ecuación 2, empleando el valor de extracción de la prensa (ecuación 3), la pol del jugo y la pol del bagazo (ecuación 4) esta última obtenida mediante el método de Norris (Spencer-Meade, 1967b), que incluye la determinación de fibra en bagazo mediante una técnica basada en el método de difusión en bolsa (Anon, 2014).
- Fibra en caña se determinó mediante la técnica basada en difusión en bolsa (Anon, 2014) y se calculó con la ecuación 5.
- Cenizas conductimétricas (CC%), se determinaron utilizando un conductímetro Orion 145A+ (ICUMSA, 2019).
- Azúcares recuperables (AR%), se calcularon de manera teórica según Diez et al. (2000).

Obtención de espectros, desarrollo de calibraciones NIRS y validación externa

Se utilizó la base de datos completa de espectros NIRS. El conjunto se depuró previamente para eliminar espectros atípicos o sospechosos con el fin de asegurar la robustez de los modelos de calibración. Al día siguiente de la cosecha, se registraron los espectros de caña desfibrada empleando el sistema CPS conectado por fibra óptica a un espectrómetro Matrix FT-NIR (Bruker Optik GmbH, Ettlingen, Alemania), tal como se muestra en la Figura 1.

El CPS es un sistema automático de arranque y parada con cinta transportadora, que opera en un rango

Ecuación 1

$$Pol\ en\ jugo\ (\%) = \frac{0{,}26 \times Lectura\ de\ pol}{(0{,}004209 \times BxRef + 0{,}9955)}$$

Ecuación 2

$$Pol\ en\ caña\ (\%) = \frac{(Extracción \times Pol\ en\ jugo + (100 - Extracción) \times Pol\ Bagazo)}{100}$$

Ecuación 3

$$Extracción\ (\%) = \frac{Peso\ jugo}{Peso\ total\ de\ caña} \times 100$$

Ecuación 4

$$Pol\ en\ bagazo\ (\%) = (A - B - C) \times L \times 0{,}0026$$

Donde:
A = peso de digestor + bagazo + agua + carbonato de sodio
B = peso del digestor
C = fibra % en bagazo
L = lectura de pol del líquido posterior a la digestión

Ecuación 5

$$Fibra\ en\ caña\ (\%) = \frac{(100 - Extracción) \times Fibra\ Bagazo}{100}$$

espectral de 4000 a 12000 cm^{-1} con una resolución de 32 cm^{-1}. Los datos espectrales y de laboratorio se procesaron con los softwares OPUS/LAB y OPUS/QUANT.

Las calibraciones se desarrollaron mediante regresión por mínimos cuadrados parciales (PLS, siglas del inglés de Partial Least Squares) a partir de la división aleatoria del conjunto total de muestras en dos grupos: calibración y validación externa (Test).

Para cada modelo, se optimizaron el número de términos o factores PLS (Rank), que representan las variables latentes utilizadas por la regresión para relacionar la información espectral con los valores analíticos de referencia, el segmento espectral, correspondiente al intervalo de números de onda seleccionado para la calibración y los pretratamientos matemáticos aplicados a los espectros, destinados a reducir efectos de dispersión, corregir variaciones de línea de base y mejorar la información espectral relevante para la predicción. Estos incluyeron derivadas de Savitzky-Golay, corrección multiplicativa de la dispersión (MSC, siglas del inglés), y normalización estándar (SNV, siglas del inglés).

El desempeño de los modelos se evaluó utilizando el conjunto de validación externa, compuesto por muestras no incluidas en la calibración, mediante la comparación entre los valores predichos por NIR y los resultados analíticos de laboratorio. Los criterios de evaluación incluyeron el error cuadrático medio de la estimación (RMSEE, siglas del inglés), el error cuadrático medio de predicción (RMSEP), la razón de desempeño del modelo o desvío de predicción residual (RPD, siglas del inglés), los coeficientes de determinación de calibración y validación (R^2) y sesgo (Bias).

Se seleccionó como calibración óptima aquella que en la validación presentó el RMSEP más bajo, el RPD más alto, el Bias más cercano a cero, y el mayor valor de R^2.

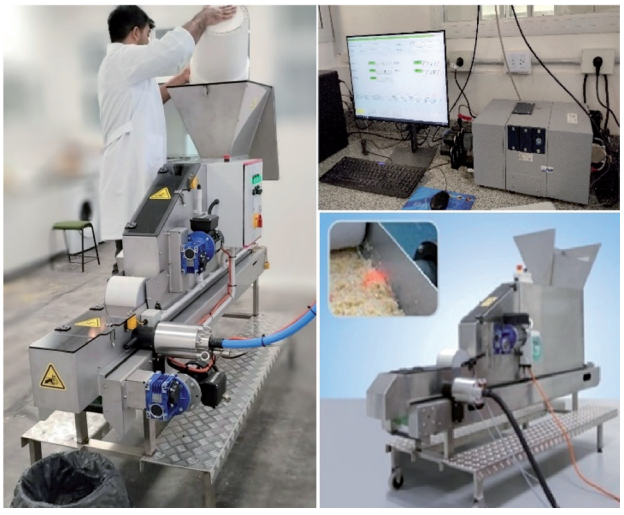


Figura 1. Sistema de presentación de caña (CPS, Cane Presentation System) y espectrómetro Matrix FT-NIR, Bruker.

RESULTADOS

La estadística descriptiva de las variables analizadas se presenta en la Tabla 1.

El diseño experimental con un amplio panel de variedades, zonas y condiciones de cultivo muy contrastantes permitió obtener un rango amplio para las variables analizadas, favoreciendo el desarrollo de modelos de calibración. Los parámetros de calibración obtenidos mediante regresión PLS se presentan en la Tabla 2.

En general, los modelos mostraron un buen nivel de ajuste (el software expresa R^2 en porcentaje, por lo que valores cercanos a 90 corresponden a $R^2 \approx 0{,}90$), con coeficientes de determinación (R^2) entre 82 y 97% según la variable, excepto para CC. Los valores de RMSEE fueron bajos y consistentes con la variabilidad de cada analito, lo que indica que los modelos lograron describir adecuadamente la relación entre las respuestas espectrales y las mediciones de laboratorio dentro del conjunto de calibración. No obstante, este indicador por sí solo no garantiza la validez del modelo, por lo que su interpretación debe complementarse con la validación externa. Los modelos requirieron un Rank de 8 a 10 factores PLS, un número acorde a la complejidad espectral de la matriz vegetal analizada.

En la mayoría de los casos, el pretratamiento matemático más efectivo fue SNV, mientras que la primera derivada resultó particularmente útil para la calibración de cenizas conductimétricas (CC%).

La validación externa permitió evaluar la capacidad predictiva de las calibraciones desarrolladas, tal como se muestra en la Tabla 3.

Los modelos de Brix, Pol en jugo y Pol en caña alcanzaron los valores más altos de desempeño, con R^2 superiores a 94% y valores de RPD mayores a 4, lo que indica una excelente capacidad predictiva y potencial para aplicaciones de rutina. Los de fibra y azúcares recuperables también mostraron buen desempeño, aunque con mayor dispersión, con RPD entre 2,3 y 3,8, valores aceptables para tareas de selección o clasificación. En el caso de cenizas conductimétricas, el modelo presentó la menor capacidad predictiva lo cual se atribuye principalmente a que la predicción por NIRS se basa en correlaciones indirectas, dado que los compuestos inorgánicos no presentan bandas de absorción en la región del infrarrojo cercano. La señal espectral proviene de modificaciones en componentes orgánicos asociados a la acumulación de minerales, por lo que la capacidad predictiva del modelo depende en gran medida de la variabilidad presente en la matriz. En este estudio, el rango reducido de cenizas limitó la fortaleza de dichas correlaciones, lo que explica el menor desempeño observado en la calibración.

En todas las variables, los sesgos obtenidos fueron cercanos a cero, lo que indica que no se observaron errores sistemáticos importantes en el conjunto de validación, y la proporción de outliers se mantuvo dentro de valores razonables para una población diversa.

La Figura 2 muestra, a modo de ejemplo, la relación entre los valores de Brix y Fibra obtenidos por laboratorio y los estimados por NIR (validación externa). La buena distribución de los puntos alrededor de la línea demuestra la consistencia de los modelos.

Tabla 1. Estadística descriptiva de las variables analizadas en 453 muestras de caña de azúcar.

Variable	Mínimo	Máximo	Media	SD
BxRef%	16,17	23,36	19,96	1,37
Pol%jugo	12,29	21,80	18,07	1,65
Pol%caña	10,59	18,16	15,33	1,31
Fibra%caña	8,13	16,42	11,52	1,32
CC%	0,18	1,41	0,80	0,18
AR%	5,97	15,65	12,50	1,54

SD: desvío estándar.

Tabla 2. Parámetros de calibración, obtenidos mediante regresión PLS.

Variable	N	RMSEE	R^2	Rank	Intervalo Espectral (cm^{-1})	Pretratamiento Espectral
BxRef%	291	0,192	97,68	9	9411,4 - 7498,2; 6109,7 - 5446,3	SNV
Pol%jugo	298	0,301	96,13	10	9411,4 - 7498,2; 6109,7 - 5446,4	SNV
Pol%caña	304	0,304	93,61	8	9411,4 - 5446,3	SNV
Fibra%caña	301	0,518	82,12	9	9411,4 - 4597,7	SNV
CC%	300	0,089	69,82	9	9411,4 - 4597,7	1° Derivada
AR%	302	0,408	91,28	10	9411,4 - 7498,2; 6109,7 - 5446,4	SNV

Tabla 3. Parámetros de validación externa.

Variable	N	RMSEP	R^2	RPD	Bias	Outliers
BxRef%	119	0,181	98,56	8,42	-0,03	43
Pol%jugo	108	0,299	97,42	6,31	-0,05	47
Pol%caña	113	0,344	94,36	4,67	-0,15	36
Fibra%caña	108	0,574	81,34	2,36	0,11	44
CC%	112	0,120	58,78	1,56	0,01	41
AR%	109	0,452	92,78	3,83	-0,11	42

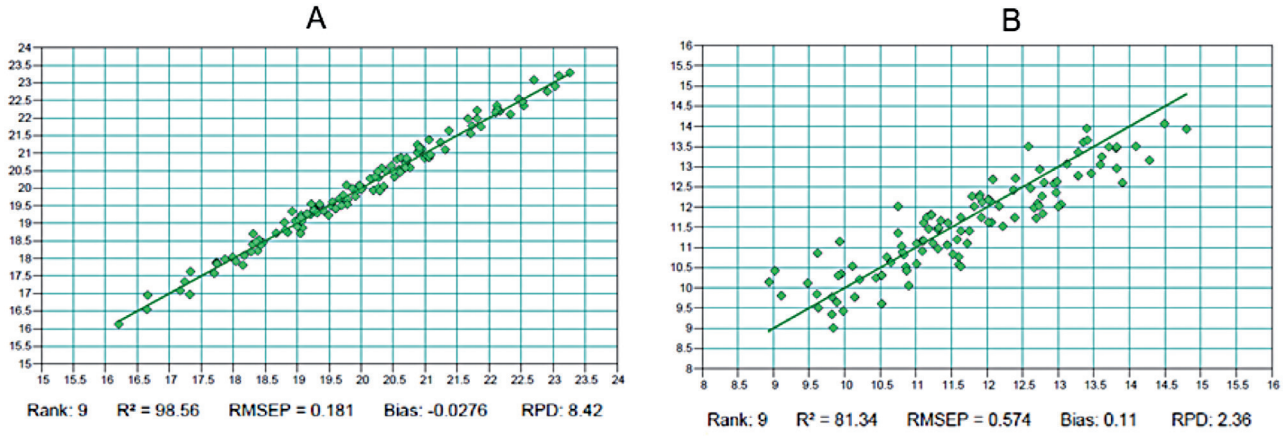


Figura 2. Gráficas de dispersión de valores predichos comparados con valores de laboratorio para Brix (BxRef%, A) y (Fibra en caña%, B) para el conjunto de muestras de validación (Test).

CONCLUSIONES

Los resultados preliminares confirman que el sistema FT-NIR Bruker CPS constituye una herramienta robusta y confiable para la determinación rápida de variables de calidad en caña desfibrada, y la integración de un sistema de presentación de gran volumen de muestra contribuye a mejorar la representatividad y robustez de los modelos. El enfoque demostró ser adecuado para estimar, con distintos niveles de precisión, los principales componentes, mostrando especial fortaleza en aquellos parámetros cuya variabilidad natural y respuesta espectral favorecen la calibración. Incluso en variables de menor respuesta directa al NIRS, como las cenizas conductimétricas, fue posible obtener modelos funcionales basados en relaciones indirectas con señales orgánicas asociadas. En conjunto, estos resultados evidencian que la metodología aplicada constituye una herramienta eficiente para el análisis rápido y no destructivo de un elevado número de muestras en programas de mejoramiento genético y control de calidad industrial.

BIBLIOGRAFÍA CITADA

Anon. 2014. The standard laboratory manual for Australian sugar mills: Volume 2. Sugar Research Australia, Brisbane.

Chen J.C.P. and C. C. Chou. 1993. Cane Sugar Handbook, 12th ed., Wiley &Sons, p. 121.

Diez, O.; S. Zossi, E. Chavanne y Cárdenas, G. 2000. Calidad industrial de las variedades de caña de azúcar de maduración temprana LCP 85-384 y LCP 85-376 en Tucumán. Análisis de sus principales constituyentes físico–químicos. Rev. Ind. y Agric. Tucumán 77 (2): 39-48.

International Commission for Uniform Methods of Sugar Analysis (ICUMSA). 2019. Method GS 1/3/4/7/8-13: The determination of conductivity ash in raw sugar, brown sugar, juice, syrup and molasses – official. Bartens, Berlin, Germany.

Spencer - Meade. 1967a. Manual del Azúcar de Caña, Editorial: Montaner y Simón, p. 457.

Spencer - Meade. 1967b. Manual del Azúcar de Caña, Editorial: Montaner y Simón, p. 670-672.