

Espectroscopía NIRS en el estudio de calidad de caña de azúcar Parte I: Correlaciones para Brix, Pol, Sacarosa, Glucosa y Fructosa

Natalia Sorol*, Silvia Zossi*, Beatriz Juarez*, Paula Diez*, Soledad Medina* y Marcelo Ruiz*

RESUMEN

La determinación de parámetros de calidad en jugos de caña de azúcar tales como brix, pol, sacarosa, glucosa y fructosa, mediante química húmeda en la gran cantidad de muestras diarias que se procesan en el laboratorio, insume tiempo y reactivos. Es por eso que en la Estación Experimental Agroindustrial Obispo Colombres (EEAOC) comenzó a implementarse la Espectroscopía de Infrarrojo Cercano (NIRS) para estudiar la calidad en jugos de caña de azúcar. Se desarrollaron métodos de calibración multivariada (seleccionando distintos algoritmos de pre-procesamiento y rangos de longitudes de onda) y se eligieron modelos definitivos para cada variable bajo estudio, obteniéndose resultados satisfactorios de error estándar de validación cruzada (SECV), error estándar de predicción (SEP) y coeficiente de correlación (RSQ) para brix, pol y sacarosa. Para glucosa y fructosa se obtuvieron RSQ más bajos y errores (SECV y SEP) más altos, posiblemente debido a las bajas concentraciones de estos azúcares en la mayoría de las muestras analizadas. Los resultados obtenidos presentaron concordancia con los valores encontrados en la literatura.

Palabras clave: jugo de caña, NIRS, calidad de caña.

ABSTRACT

NIRS spectroscopy in the study of sugarcane quality Part I: Correlations for Brix, Pol, Sucrose, Glucose and Fructose

Determination of quality parameters in sugar cane juice, such as brix, pol, sucrose, glucose and fructose by wet chemistry in the large number of daily samples that are processed in the laboratory takes time and reagents. Because of that Estación Experimental Agroindustrial Obispo Colombres (EEAOC) began to implement Near Infrared Spectroscopy (NIRS) to study sugarcane quality. Multivariate calibration methods were developed (selecting different pre-processing algorithms and wavelength ranges) for the above variables, obtaining satisfactory standard errors of cross validation (SECV), and of prediction (SEP), and correlation coefficients (RSQ) for brix, pol and sucrose. For glucose and fructose, lower RSQ and larger errors (SECV and SEP) were obtained, possibly because of the low concentrations of these sugars in the majority of the samples. The results agreed with those found in the literature.

Key words: sugarcane juice, NIRS, cane quality.

Fecha de recepción: 01/11/18 - Fecha de aceptación: 05/01/21

INTRODUCCIÓN

La caña de azúcar está compuesta por agua, sólidos solubles (principalmente compuestos azúcares y no azúcares) y sólidos insolubles (cenizas y fibra), variando su composición de acuerdo a la edad, la variedad, el manejo del cultivo y las condiciones ambientales.

Los parámetros principales que se analizan en el jugo de caña para evaluar su calidad son los compuestos azúcares tales como el contenido de pol (sacarosa aparente), brix (sólidos totales disueltos), sacarosa y azúcares reductores (fructosa y glucosa). Los compuestos no azúcares son generalmente factores importantes también, responsables de efectos negativos en la fabricación de azúcar. Ellos son sales de ácidos orgánicos e inorgánicos, ácidos carboxílicos, aminoácidos, proteínas, polisacáridos solubles, almidón, ceras, grasas y otros compuestos minoritarios tales como polifenoles, flavonoides, etc. (Van der Poel *et al.*, 1998)

Considerando que el análisis de los parámetros de calidad de una gran cantidad de muestras por metodologías tradicionales de laboratorio implica gasto de tiempo y dinero, la Espectroscopía de Infrarrojo Cercano (NIRS por sus siglas en inglés) se presenta como una buena alternativa para reducir los tiempos de análisis y minimizar el impacto ambiental, ya que no se utilizan reactivos y a la vez se mantiene la exactitud en los resultados.

Es una técnica que consiste en el análisis de una muestra para determinar la cantidad de luz que absorbe o refleja en la región del infrarrojo cercano, relacionando este valor con el analito que se quiere determinar.

En el campo industrial, las aplicaciones de espectroscopía NIRS son diversas, especialmente en el control de calidad y la evaluación de materiales en la producción de granos, cereales, forraje, fibras, azúcar, sustancias químicas, farmacéuticos y distintos tipos de alimentos.

Entre sus principales desventajas pueden mencionarse que es una técnica comparativa, el alto costo de los equipos, la competencia técnica que debe poseer el operador y el tiempo requerido para desarrollar una base de datos que permita obtener ecuaciones de calibración robustas que originen predicciones confiables (White *et al.*, 2006)

En la región del NOA, el único ingenio que implementó la espectroscopía NIRS en el proceso de fabricación de azúcar fue Ledesma, en la provincia de Jujuy, tanto para el pago de caña de azúcar como para control del proceso. En cambio esta tecnología cuenta actualmente con pocos estudios en Tucumán y su aplicación industrial no está ampliamente difundida ni implementada. Sin embargo existe un gran número de ventajas y aplicaciones que surgen de su utilización, como está demostrado por su uso en industrias azucareras de diversos países.

En la Sección Química de la Estación experimental Agroindustrial Obispo Colombes (EEAOC) comenzó a utilizarse esta tecnología en 2005 y actualmente ha reemplazado algunas metodologías de referencia en ensayos de investigación en caña de azúcar de los diferentes proyectos de esta institución.

Este artículo es un resumen de trabajo de los últimos diez años para desarrollar modelos de calibración apropiados utilizando la espectroscopía NIRS en el estudio de compuestos azúcares y no azúcares que permitan determinar la calidad de la caña de azúcar.

En esta primera parte se presentan los resultados obtenidos para compuestos azúcares.

MATERIALES Y MÉTODOS

Preparación de la muestra

Durante los períodos de zafra desde el año 2011 hasta el año 2017, en el Laboratorio de Investigaciones Azucareras de la EEAOC se analizaron 3813 muestras de jugo primario procedentes de cañas de azúcar cosechadas de forma manual, sin hojas ni despuntes. Para obtener el jugo primario de las distintas variedades estudiadas, los tallos se procesaron en un desfibrador a martillo con un "open cell" cercano al 95% y luego se prensó la caña desfibrada en una prensa hidráulica que aplica una presión de 250 kg/cm² durante un minuto.

Obtención de las variables de interés mediante métodos químicos tradicionales

En el laboratorio, mediante análisis húmedo se determinaron los analitos de interés que luego se utilizaron como metodología de referencia en el desarrollo de modelos de calibración para el equipo NIRS. Las variables determinadas fueron:

- Brix refractométrico: se determinó utilizando un refractómetro digital marca KEM modelo RA-620 (Chen and Chou, 1993)

- Pol % jugo: se llevó a cabo utilizando subacetato de plomo como agente clarificante y realizando las lecturas en un polarímetro digital Anton Paar modelo MCP 500 Sucromat (Spencer - Meade, 1967)

- Sacarosa, glucosa y fructosa: se analizaron mediante Cromatografía Líquida de Alta Performance (HPLC). Se empleó un equipo marca Waters compuesto por un inyector con bomba cuaternaria Alliance modelo e 2695, un detector del índice de refracción Waters 410 y un horno calefactor. Se empleó una columna Sugar Pack de Waters, rellena con resina de intercambio catiónico en forma cálcica. Las condiciones utilizadas fueron las siguientes:

- Fase Móvil: agua + 0,05 g/l de EDTA cálcico
- Temperatura de columna: 85°C
- Temperatura de detector: 45°C
- Caudal: 0,5 ml/min
- Volumen de inyección: 20 µl

Obtención de espectros de infrarrojo cercano

Las muestras en cuestión fueron analizadas en un espectrofotómetro de infrarrojo cercano (NIRS,) marca Foss NIRSystems modelo 6500 para muestras líquidas, con detector de transmitancia de sulfuro de plomo (PbS), lámpara de tungsteno-halógeno y camino óptico de 1 mm para obtener su espectro en un rango de longitudes de onda de 400 a 2500 nanómetros, utilizando para ello el software proporcionado por el fabricante ISIScan 2.21.

Tratamiento de los datos mediante quimiometría

Para construir los modelos de calibración y llevar a cabo las validaciones correspondientes a cada variable se exportaron los espectros al software WinISI III, el cual permite construir una base de datos o modelo de calibración con los espectros de cada muestra y los resultados obtenidos en el laboratorio para cada una de ellas. El primer paso fue separar al azar, mediante el software, la

totalidad de las muestras en dos sets, uno para calibración con 3413 muestras y otro para validación con 400 muestras.

Calibración

Antes de desarrollar los modelos de calibración se realizó el proceso de selección de muestras utilizando un criterio de discriminación proporcionado por el software, basado en el análisis por componentes principales (PCA) (Massart, 1997) y en la distancia de Mahalanobis (Shenk and Westerhaus, 1991; Gemperline *et al.*, 1989) para identificar las muestras que no pertenecen a la población (“outliers”) y seleccionar las muestras más representativas, candidatas para ser incluidas en la calibración.

Antes de adoptar un modelo de calibración, se procedió a elegir diferentes combinaciones posibles entre las opciones disponibles en el software, teniendo en cuenta, para elegir el modelo definitivo, aquella combinación cuyos parámetros estadísticos de error fueran mínimos, es decir, la que presentó menor error estándar de validación cruzada (SECV) y mayor coeficiente de correlación (RSQ). Estas opciones disponibles incluyen:

- Algoritmos de pre-procesamiento de los espectros: se trata de diferentes tratamientos matemáticos cuyo propósito es aumentar la relación señal ruido mediante corrección de aberraciones relacionadas con efectos de la dispersión de la luz (SNV por standard normal variation), deriva de la línea de base (derivadas) y tendencias curvilíneas de los datos como ser detrend.
- Elección de rango de longitudes de onda: puede elegirse distintos segmentos del mismo con la posibilidad de eliminar ciertas regiones si se consideran poco informativas.
- Elección de métodos de calibración multivariada: cuadrados mínimos parciales (PLS), cuadrados mínimos parciales modificados (MPLS) y regresión por componentes principales (PCR).

Validación

Para verificar la verdadera precisión de la calibración NIRS fue necesario evaluar el comportamiento de esta sobre muestras “incógnitas”. Se utilizó el set de vali-

dación compuesto por muestras que no se incluyeron en la calibración, y se analizaron bajo similares condiciones a las que se esperan en el análisis de rutina.

RESULTADOS Y DISCUSIÓN

De las 3813 muestras de jugo de caña se obtuvo, al azar mediante el software, un set para calibración con 3413 muestras y otro para validación con 400 muestras. En la Figura 1 se muestran los espectros de infrarrojo cercano de las muestras utilizadas. Donde: R = Reflectancia.

Calibración

Antes de realizar la calibración, el software seleccionó 3150 muestras, eliminando las “outliers” y redundantes.

Luego de llevar a cabo el desarrollo de los modelos de calibración eligiendo las diferentes combinaciones posibles entre las opciones disponibles en el software, se seleccionó un modelo definitivo para cada variable en estudio cuyas características están descritas en la Tabla 1, en la cual se muestran el algoritmo de pre-procesamiento, el rango de longitudes de onda y el método de calibración multivariada utilizado para cada uno, donde:

- Derivada es la función utilizada para el suavizado del espectro utilizando el método del polinomio Savitzky-Golay (Savitzky and Golay, 1964)
- Gap es el número de puntos de la ventana elegida. Determina cada cuántos nanómetros se hacen las derivadas utilizando el método del polinomio Savitzky-Golay.
- Smooth es el grado del polinomio. Al calcular la derivada mediante el método de polinomio de Savitzky-Golay, el software coloca puntos al lado de los puntos principales con el fin de suavizar la curva (ya que el espectro proporcionado por el equipo contiene mediciones cada dos nanómetros). El número de puntos a colocar estará definido por el grado de polinomio que se seleccione.
- Segmento 1 es el primer segmento de longitud de onda elegido en nanómetros.
- Segmento 2 es el segundo segmento de longitud de onda elegido en nanómetros.
- SNV o standard normal variation es el algoritmo

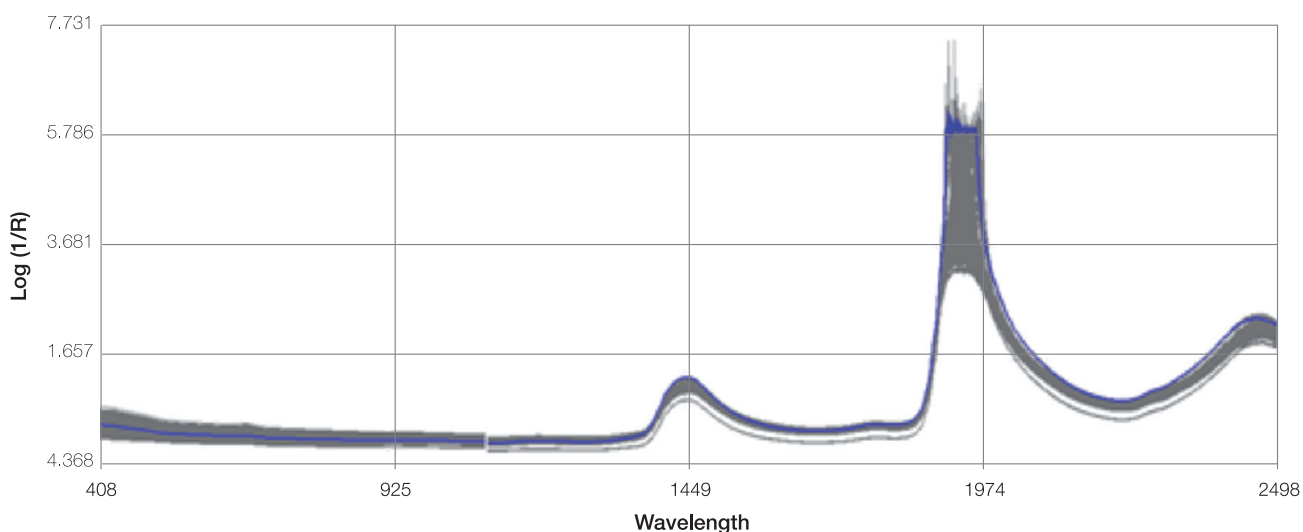


Figura 1. Espectros de infrarrojo cercano (Log (1/R) versus longitud de onda (en nanómetros) de las muestras de jugo de caña en estudio.

Tabla 1. Algoritmos de pre-procesamiento, rango de longitudes de onda y método de calibración multivariada de los modelos definitivos elegidos para cada variable bajo estudio.

Variable	Algoritmo de pre-procesamiento				Rango de Long. de Onda (nm)		Método
	Tratamiento	Derivada	Gap	Smooth	Segmento 1	Segmento 2	
Brix	SNV & Detrend	2	4	4	408-1092,2	1108-2492,2	MPLS
Pol	SNV & Detrend	1	4	4	408-1842,2	2094-2300,2	MPLS
Sacarosa	SNV & Detrend	2	4	4	408-1092,2	1108-2492,2	MPLS
Glucosa	SNV & Detrend	2	4	4	408-1092,4	1108-2492,4	MPLS
Fructosa	SNV & Detrend	1	4	4	1200-1842,2	2094-2300,2	MPLS

de pre-procesamiento que corrige dispersiones de luz.

- Detrend es el algoritmo de pre-procesamiento que corrige tendencias curvilíneas de los datos.

- MPLS es el método de calibración multivariada.

Los parámetros estadísticos del modelo de calibración elegido para cada una de las variables de interés se muestran en la Tabla 2, donde:

N es el número de muestras.

Media es el valor medio; SD, el desvío estándar; y Rango, el intervalo donde están comprendidos los valores obtenidos en el laboratorio por métodos químicos tradicionales.

RSQ es el coeficiente de correlación (R2).

SECV es el error estándar de validación cruzada.

$$SECV = \sqrt{\frac{\sum_i (y_i - Y_i)^2}{n-t-1}}$$

Donde

y_i = valores obtenidos por métodos químicos tradicionales

Y_i = Valores NIRS

n = n° de muestras

t = n° de componentes principales

Otra forma de evaluar las ecuaciones NIRS en relación a su precisión y exactitud es a través de la relación entre el error estándar de la validación cruzada y el desvío estándar del análisis químico (SECV/SD), siendo considerada como una ecuación con alto poder de predicción si la relación es menor a 0,33 (Cozzolino *et al.*, 2003).

En este método de calibración se realiza una validación cruzada que separa las muestras en conjuntos de datos, deja uno de ellos afuera y construye el modelo con los restantes, y luego predice las variables para el conjunto excluido y almacena los desvíos. Esto se realiza hasta que todos los conjuntos de muestras se excluyen una vez. El SECV es la media de todos los desvíos almacenados.

Tabla 2. Parámetros estadísticos de los modelos de calibración.

	N	Media	SD	Rango	RSQ	SECV	SECV/SD
Brix (%)	3110	20,01	1,72	14-25	0,993	0,154	0,09
Pol (%)	2712	17,96	1,95	12-24	0,986	0,243	0,12
Sacarosa(%)	3120	17,65	1,99	12-24	0,971	0,365	0,18
Glucosa(%)	3108	0,15	0,14	0,05-0,60	0,810	0,066	0,47
Fructosa(%)	2772	0,16	0,14	0,05-0,60	0,801	0,066	0,47

La calibración de las variables brix, pol y sacarosa fue adecuada, lo que se refleja en un bajo error estándar de validación cruzada (SECV) y un significativo coeficiente de correlación (RSQ). También puede inferirse que existe una capacidad aceptable de predicción del modelo, dado que la relación entre el error estándar de validación cruzada y el desvío estándar del análisis químico fue menor a 0,33.

En el análisis de glucosa y fructosa se observó un bajo coeficiente de correlación, alto error estándar de validación cruzada y, por consiguiente, un bajo poder de predicción. Esto posiblemente se deba a que la mayoría de las muestras analizadas tenían un alto nivel de madurez, no presentaban deterioro y, como se dijo anteriormente, fueron cosechadas a mano y no tenían hojas ni despuntes. En estas condiciones el jugo de caña presenta niveles de glucosa y fructosa muy bajos (Van der Poel *et al.*, 1998; Chen, 1991).

Al observar un histograma de frecuencia de los valores de laboratorio para ambos parámetros se comprobó que hay una distribución poco homogénea de los datos a lo largo del rango estudiado. No así en el caso de brix, pol y sacarosa.

Validación

Para evaluar los modelos de calibración obtenidos se procedió a realizar su validación, es decir, la comparación de los datos de laboratorio con los obtenidos mediante los modelos NIRS. Se utilizó un set de 400 muestras. En la Tabla 3 se muestran los parámetros estadísticos de la validación, donde:

N es el número de muestras.

RSQ es el coeficiente de correlación.

SEP es el error estándar de predicción, el cual se calcula como el desvío estándar de los residuales (valor laboratorio – valor NIRS)

En la Figura 2 se muestran las gráficas de la validación de los variables estudiadas, expresadas en % (p/p).

Tabla 3. Parámetros estadísticos de la validación de los modelos obtenidos para las variables estudiadas. N = 400.

	RSQ	SEP
Brix	0,994	0,134
Pol	0,985	0,233
Sacarosa	0,972	0,341
Glucosa	0,831	0,060
Fructosa	0,842	0,060

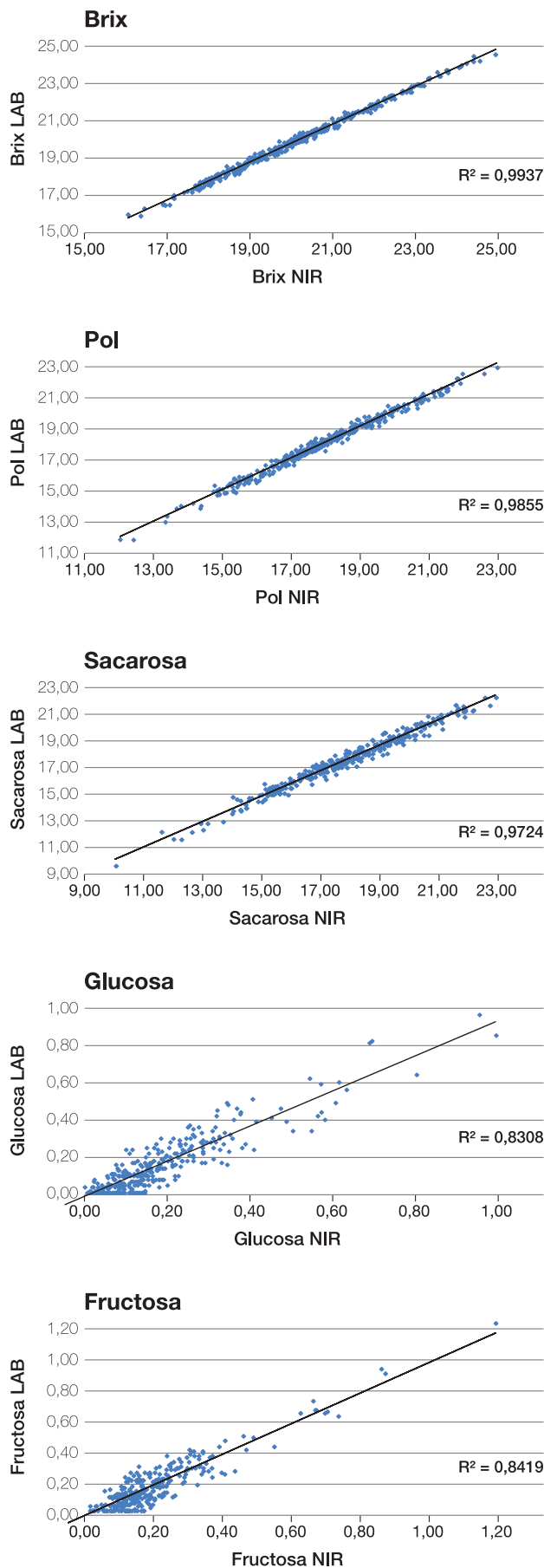


Figura 2. Gráficas de validación de los parámetros estudiados. Valores de NIRS versus valores de laboratorio

Los resultados de las validaciones fueron similares a los de la calibración. De acuerdo a los resultados mostrados en la Tabla 3 y en la Figura 2, los modelos creados para brix, pol y sacarosa fueron satisfactorios.

Nuevamente, al igual que en la calibración, en las gráficas de validación obtenidas para glucosa y fructosa puede observarse la distribución con respecto a la dispersión de todo el grupo de muestras. Se aprecia una mayor cantidad de muestras en la parte más baja del rango y muy pocas en la parte más alta del mismo. Esto se refleja también en los mayores errores y menores coeficientes de correlación obtenidos.

Debe tenerse en cuenta también que estos azúcares reductores son componentes minoritarios del jugo; sin embargo los SEP fueron bajos y podrían llegar a utilizarse los resultados de estos modelos como una medida cualitativa rápida para determinar no solo el grado de madurez, sino también el grado de deterioro de la muestra de jugo de caña.

Los valores determinados en nuestro trabajo presentan resultados concordantes con los obtenidos en otros países. En la Tabla 4 se muestran los parámetros de calibración y validación obtenidos en Colombia para esta metodología (García Asprilla and Ramírez-Navas, 2018) para glucosa y fructosa. En la Tabla 5 se muestran los valores determinados en Colombia para brix, pol, sacarosa, glucosa y fructosa (Larrahondo *et al.*, 2001); y en la Tabla 6, los resultados obtenidos en Brasil para brix y pol (Valde-rama *et al.*, 2007).

Tabla 4. Parámetros de calibración y validación obtenidos en Colombia para glucosa y fructosa (García Asprilla y Ramírez Navas, 2018). N (calibración glucosa) = 141, N (calibración fructosa) = 129, N (validación de glucosa y fructosa) = 66

	Calibración		Validación	
	RSQ	SECV	RSQ	SEP
Glucosa	0,972	0,099	----	0,136
Fructosa	0,996	0,042	----	0,080

En este trabajo realizado en Colombia se utilizaron menor cantidad de muestras que en la EEAOC, 141 y 129 muestras para llevar a cabo la calibración de glucosa y fructosa respectivamente, y 66 muestras para validación. En la calibración, si bien el coeficiente de correlación obtenido para glucosa es más alto que el obtenido en la EEAOC, el valor de SECV es más elevado. En el caso de fructosa, el coeficiente de correlación y el SECV obtenido son mejores que los obtenidos en la EEAOC; sin embargo, en la validación los valores de SEP fueron más elevados que los obtenidos en la EEAOC para ambos azúcares. También a los autores de este trabajo se les presentó el problema de la distribución poco homogénea de los datos a lo largo del rango de estudio, y enfatizan que deben agregar a su modelo muestras con valores más altos para ambas variables.

Los resultados de Larrahondo y colaboradores muestran que los valores de coeficiente de correlación y SECV obtenidos para las variables estudiadas fueron muy similares a las obtenidas en la EEAOC, excepto para pol en la cual obtuvieron un menor SECV.

Tabla 5. Parámetros de calibración obtenidos en Colombia para brix, pol, sacarosa, glucosa y fructosa (Larrahondo *et al.*, 2001).

	Rango	RSQ	SECV
Brix	9,82-22,98	0,99	0,13
Pol	8,21-21,21	0,99	0,14
Sacarosa	12,12-22,02	0,99	0,30
Glucosa	0,11-0,56	0,83	0,07
Fructosa	0,10-0,66	0,80	0,08

Tabla 6. Parámetros de calibración obtenidos en Brasil para brix y pol (Valderrama *et al.*, 2007). N= 1381.

	RSQ	SECV
Brix	0,99	0,30
Pol	0,99	0,44

En las 1381 muestras analizadas en Brasil por Valderrama y colaboradores, los coeficientes de correlación fueron altos y los SECV obtenidos resultaron superiores a los obtenidos en la EEAOC. Esto podría deberse a que utilizaron PLS como método de calibración multivariada, a diferencia de MPLS que se utilizó en la EEAOC.

CONCLUSIONES

Los resultados obtenidos para brix, pol y sacarosa confirman que los modelos de calibración MPLS construidos tuvieron una alta capacidad predictiva. Por el contrario, para glucosa y fructosa no tuvieron un buen desempeño.

Para mejorar los modelos de los azúcares reductores debería procesarse mayor cantidad de muestras con concentraciones que cubran de manera más homogénea el rango a analizar, y así poder obtener una calibración más robusta para ambos azúcares, con valores distribuidos uniformemente en todo el intervalo.

Reemplazar esta metodología por las convencionales de laboratorio brindará grandes ventajas, ya que es una técnica rápida, confiable, no-destructiva y amigable con el medioambiente debido a que no emplea reactivos en su implementación; pero debe recordarse que es una técnica comparativa y que sus errores nunca serán menores que los del método primario o de referencia.

BIBLIOGRAFIA CITADA

- Chen, J. C. P.** 1991. Manual del Azúcar de Caña. Versión autorizada en español de Cane Sugar Handbook. 11th ed., Limusa. pp. 45
- Chen, J. C. P. and C. C. Chou.** 1993. Cane Sugar Handbook, 12th ed., Wiley & Sons, pp. 121.
- Cozzolino, D.; A. Fassio y E. Fernández.** 2003. Uso de la espectroscopía de reflectancia en el infrarrojo cercano para el análisis de calidad de ensilaje de maíz. *Agricultura Técnica*, 63(4): 387-393.
- García-Asprilla, I. and J. Ramírez-Navas.** 2018. Near-infrared spectroscopy: a rapid alternative technique to reducing sugars determination in juice of sugarcane (*Saccharum officinarum* L.). *Journal of Pharmacy & Pharmacognosy Research*, 6 (5): 392-401.
- Gemperline, P.; L. Webber and F. Cox.** 1989. Raw materials testing using soft independent modeling of class analogy analysis of near- infrared reflectance spectra. *Analytical chemistry* 61 (2): 138-144.
- Larrahondo, J.; F. Palau; A. Navarrete and C. Ramírez.** 2001. Applications of near infrared spectroscopy in the sugarcane industry of Colombia. En: *Proc. Int. Soc. Sugarcane Tech.*, 24, Brisbane Australia 2001, 1, pp. 163-165.
- Massart.** 1997. Handbook of chemometrics and qualimetrics. Volume 20A 1st Edition Part A. Elsevier Science. Cap. 17
- Savitzky, A. and M. J. Golay.** 1964. Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* 36 (8): 1627-1639.
- Shenk, J. S. and M. O. Westerhaus.** 1991. Population definition, sample selection, and calibration procedures for near infrared reflectance spectroscopy. *Crop Sci.* 31:469-474.
- Spencer - Meade.** 1967. Manual del Azúcar de Caña, Editorial: Montaner y Simón, p. 457.
- Valderrama, P.; J. Braga and R. Poppi.** 2007. Validation of multivariate calibration models in the determination of sugar cane quality parameters by near infrared spectroscopy. *J. Braz. Chem. Soc.* 18 (2): 256-266.
- White, B. E.; L. S. Polanco and Ch. Verret.** 2006. Utilization of NIR spectroscopy for factory control. En: *Proc. Sugar Processing Research Conference, Frontiers in Sugar Processing.* Aguas de Sao Pedro, Brazil 2006, pp. 230-244.
- Van der Poel P. W.; H. Shiweck and T. Schwartz.** 1998. Sugar Technology. Beet and Cane Sugar Manufacture. Verlag Bartens KG. Berlín.