Use of NIRS, PLS and OPS Variable Selection to Predict Five Quality Parameters of Sweet Sorghum Juice used to Produce Bioethanol

Marcelo Sena
Cristiane Guimarães
Camila Assis
Maria Lúcia Simeone

Track
8. Agriculture and food

First generation biofuels represent a well established sustainable energy source, reducing pollutants emission in the atmosphere. A culture highly feasible to be used for the bioethanol generation in Brazil is the sweet sorghum, which is particularly promising as complementary for the diversification of sugarcane croplands due to its shorter harvest period. Sorghum breed programs have sought to develop new cultivars of higher bioethanol yield, generating a large number of juice samples that must be analysed and characterized before the fermentation. Thus, rapid and non-destructive analytical methods are required. In this work, multivariate calibration models were developed for determining five quality parameters of sorghum juice (Brix, saccharimeter reading/SR, reducing sugars/RS, Pol, apparent purity/Q) based on PLS and NIRS. All of these models were optimized by variable selection with Ordered Predictors Selection (OPS). OPS selects the most predictive variables by a systematic investigation of the PLS informative vectors (correlation coefficients, VIPscores, NAS, etc.) in a cross-validation process, leading to a great reduction in the number of variables.

500 juice samples were obtained from 275 recombinant inbred lines of sweet sorghum derived from two contrasting lines in relation to the quality and quantity of sugars. The five quality parameters were determined by appropriate reference methods. Juice samples were previously filtered in cotton and their spectra were recorded on a Büchi NIRFlex N-500 FT-NIR spectrometer, equipped with a transreflectance accessory, from 10000 to 4000 cm⁻¹, with 4 cm⁻¹ steps. Data were handled with Matlab, PLS_Toolbox and a homemade routine for OPS.

The analytical ranges for the parameters were 5.5-18.1°B (brix), 1.1-53.2°Z (SR), 1.2-5.2%
w/v (RS), 0.3-13.0% w/v (pol) and 9.8-83.0% w/v (Q). Samples were split in 333 for
calibration and 167 for validation using the Kennard-Stone algorithm. The best models were
obtained with the preprocessings first derivative, Savitzky-Golay smoothing and mean
center. All the PLS models were built with 6 latent variables. Subsequently, the models were
optimized by variable selection with OPS, reducing the number of wavenumbers used from
1501 (full spectra) to 50-130, depending on each parameter. All the RMSEC and RMSEP
values for the OPS-PLS models were decreased. RMSEP were decreased from 0.4 to 0.3°B
(brix), 3.2 to 2.1°Z (SR), 0.4 to 0.3% w/v (RS), 0.8 to 0.6% w/v (pol) and 5.8 to 5.3% w/v (Q).
Correlation coefficients (r) for the reference versus predicted values were between 0.896
and 0.992. The randomness of the residuals was checked by appropriate statistical tests,
assuring the linearity of the methods. The developed methods were validated by estimating
figures of merit, such as trueness, precision, analytical sensitivity and bias. RPD (residual
prediction deviation) was estimated between 2.4 and 6.2, attesting the good prediction
ability of the models. The use of OPS allowed developing simpler, more interpretable and
predictive multivariate calibration models. The NIRS methods are rapid, non-destructive and
of low cost, being appropriate for replacing the more laborious reference methods, what is
stressed considering the large number of samples analyzed.

Acknowledgements

CAPES for a fellowship (C.A.), Reinaldo F. Teófilo (UFV, Brazil) for the OPS routine, and
FAPEMIG for the financial support.
15.30 OP14 Simulation method design to link the spectral properties of dense microbial cultures to cell physiology
Dario Benitez, Sarah Belle, Jean Michel Roger - France
Institut IMS TAP

15.50 OP15 Precipitation of soil bulk density for a set of Scottish afforested soils: a comparison of approaches
Jean Robert, Emma Pierson Fernández, J. Potts, A. Lily
The James Hutton Institute - Scotland

ORAL SESSION 4 - AGRICULTURE AND FOOD [Tuesday 20th]
Chair: Dolores Penez-Marin

8:40 Session to honour

8:40-8:50 Session to honour

8:50-9:00 Use of NIRS, PLS and QPSK variable selection to predict the quality parameters of sweet sorghum juice used to produce bioethanol
Marcelo M. Song, Claudine C. Cunha, Camila Assis, Maria Lucia F. Simeone
Universidade Federal de Minas Gerais - CES - Chemistry - Brazil

9:00-9:10 Development of NIRS method for energetic value of high fiber cane clones
Fabio Garcia, Camila Roussel, Milene Melo, Regis Ciacco, Christine-Poeder
CRAA UFRJ, Quatro, 7 chemin de FRAT, 07143 Saint-Pierre, La Réunion - France

9:10-9:20 The impossible made possible: providing a detailed topological map of chemical classification parameters for pork carcasses with on-line slaughter line instrumentation
Klaus Sarnitz, Karen Klungstedt
University of Copenhagen - Department of Food Science - Spectroscopy & Chromometrics - Denmark

ORAL SESSION 5 - AGRICULTURE AND FOOD [Tuesday 20th]
Chair: Maria Lucia F. Simeone

10:20 OP16 Evolution of Vis/NIR bulk optical properties of apple skin and flesh during fruit maturation
Robine Van Nieten, Bert Aerisso, Rodrigo Martí, Ann Schenk, Bart Nicola, Wouter Saeys
KU Leuven Department of Biosystems, Melle, Leuven - Belgium

10:30 OP18 Identification and quantification of bovine meat adulteration by using Vis-NIR spectrometry
Cristina Carletti, Jose Manuel Artes, Darcie Bailey Engelhard, Cristina Rampazzo
University of Milan-CservicolFd Int. Dept of Food Science - Department of Food, Environmental and Nutritional - Italy