Bioresource Technology 102 (2011) 9542–9549

Contents lists available at SciVerse ScienceDirect

Bioresource Technology

journal homepage: www.elsevier.com/locate/biortech

Direct prediction of bioethanol yield in sugar beet pulp using Near Infrared Spectroscopy

C. Magaña [a], N. Núñez-Sánchez [b,*], V.M. Fernández-Cabanás [b], P. García [a], A. Serrano [b], D. Pérez-Marín [b], J.M. Pemán [a], E. Alcalde [a]

[a] Syngenta Seeds S.A., Barcelona 08006, Spain
[b] NIR Soluciones, S.L., Cordoba 14006, Spain

ARTICLE INFO

Article history:
Received 29 April 2011
Received in revised form 8 July 2011
Accepted 15 July 2011
Available online 4 August 2011

Keywords:
Sugar industry
Biofuel crops
First generation bioethanol
Sugar beet
Near Infrared Spectroscopy

ABSTRACT

Sugar beets are a raw material for the production of sugar and ethanol. The decision on which end product to pursue could be facilitated by fast and reliable means of predicting the potential ethanol yield from the beets. A Near Infrared (NIR) Spectroscopy-based approach was tested for the direct prediction of the potential bioethanol production from sugar beets. A modified partial least squares (MPLS) regression model was applied to 125 samples, ranging from 21.9 to 31.0 g L^{-1} of bioethanol in sugar beet brei. The samples were analyzed in reflectance mode in a Direct Contact Food Analyser (DCFA) FOSS-NIRSystems 6500 monochromator, with standard error of cross validation (SECV), standard error of prediction (SEP), coefficient of determination (r^2) and coefficient of variation (CV) of 0.51, 0.49, 0.91 and 1.9 g L^{-1}, respectively. The NIR technique allowed direct prediction of the ethanol yield from sugar beet brei (i.e. the product obtained after sawing beets with a proper machine) in less than 3 min.

1. Introduction

The 2009 EU Renewable Energy Directive sets a goal of generating 20% of its energy from renewable sources and of increasing share of biofuels to 10% of the transport fuel consumption by 2020 (EU Directive 2009/28/EC, 2009). Ethanol is a renewable fuel source that can be produced by fermentation of sugar from sugar and starch-containing crops (first generation feedstock), or from lignocellulosic biomass (second generation feedstock) (Białas et al., 2010). The utilization of second generation feedstock is still challenging (Nghiem et al., 2010) whereas ethanol production from first generation crops such as sugar beets is relatively simple.

Sugar beet (*Beta vulgaris* L.) is a biennial plant which stores sucrose as primary carbon and energy resources and accounts for 30% of the world's sugar production. During its first growing season, it produces a large (1–2 kg) storage root whose dry mass contains 15–20% sucrose by weight. Sugar beets have the potential to produce 30–40 tons of roots per hectare under non-irrigated conditions and 50–70 tons per hectare with irrigation.

Sugar beet breeding programs have traditionally been focused on developing varieties combining outstanding agronomic performance and improved quality characteristics for human consumption. To meet requirements by the ethanol industry, new breeding criteria such as outstanding potential for ethanol production, need to be taken into consideration to develop varieties better adapted to ethanol production. Such breeding programs would benefit from a fast, inexpensive and non-destructive method of determining the bioethanol yield of beet samples. Similarly, the sugar and ethanol industry would be able to utilize such a tool to quickly determine if batches of sugar beets would be economically more suitable for sugar or ethanol production based on their content of quality parameters (sucrose, total sugars or bioethanol yields, among others).

Near Infrared Spectroscopy (NIRS) has the potential for quantitative and qualitative prediction of the main parameters such as protein, fat, moisture, fiber, ash, starch or sugar content of raw materials related with the quality of the agricultural products (Williams and Norris, 2001). NIR has been applied to the estimation of the production of bioethanol from different crops based on the determination of fermentable products such as starch in cereals (Sohn et al., 2007) or sugars in sugar beet (Roggo et al., 2004). Liebmann et al. (2009) developed an NIR application to analyze glucose and ethanol during bioethanol production for process control purposes, and later on obtained NIR calibrations for the determination of moisture, protein and starch in feedstock (cereals), and glucose, ethanol, glycerol, lactic acid, acetic acid, maltose, fructose and arabinose in the fermentation media (Liebmann et al., 2010). Pohl and Senn (2011) related fermentable substances and ethanol yield in cereals (wheat, rye and triticale), and concluded that near infrared reflectance spectroscopy provides a cost and timesaving tool for the evaluation of cereal grains for fuel ethanol production. However, the study was developed with a reduced number of

* Corresponding author. Tel.: +34 957 21 1004; fax: +34 957 21 8436.
E-mail address: nnunez@nirsoluciones.com (N. Núñez-Sánchez).

0960-8524/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.biortech.2011.07.045

samples and for industrial application further work is necessary in order to cover a broad range of growth locations, varieties and years of cultivation and get a more reliable model.

Since currently no fast analytical methodology to determine the bioethanol production potential of sugar beets or sugar beet pulp exists, the aim of the current study was to develop a NIR spectroscopy-based method that could be employed by industry in the selection of feedstock for use in bioethanol and sugar production.

2. Methods

2.1. Sugar beets

Sugar beets (*B. vulgaris* L., Syngenta Seeds, S.A.) were collected during the 2006/07, 2007/08 and 2008/09 seasons (50, 24 and 50 samples respectively). Field trials were conducted by the Research Association for the Improvement of Sugar Beet Crop of Spain (AIMCRA) in the north-central Spain (Castilla y León), where sugar beet is sown in spring and harvested in autumn, and in southern Spain (Andalucía) where sugar beet is sown in autumn and harvested in summer.

At harvest, all beets from every plot were washed, weighed and passed through a machine with multiple saws to obtain a fine brei (macerated root material), following the standard method used in sugar factories. Different subsamples of this brei were used to measure sucrose, total sugars, α-amino nitrogen and potassium and two subsamples (100 g) were immediately frozen and kept at −20 °C until fermentation and spectral analysis for NIRS calibration, respectively.

2.2. Bioethanol determination

Bioethanol yield was determined for each subsample at the CIEMAT (Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas).

Fermentation was carried out in triplicate in 250-mL Erlenmeyer flasks at 32 °C and 50 rpm until sucrose, glucose or fructose could no longer be detected in the medium. Sugars content was monitored by high performance liquid chromatography (HPLC) in a Waters 2695 HPLC System with a refractive index detector. An AMINEX HPX-87P carbohydrate analysis column (Bio-Rad, Hercules, CA) operating at 85 °C with ultrapure water as a mobile phase (0.6 mL/min) was used. Fermentations were performed with substrate concentration of 30% of total mass (v/v). The medium consisted of 1 g L^{-1} yeast extract, 7.5 g L^{-1} $(NH_4)_2SO_4$, 3.5 g L^{-1} KH_2PO_4 and 0.7 g L^{-1} $MgSO_4$, and was autoclaved for 20 min at 121 °C before inoculation with 2 g L^{-1} of *Saccharomyces cerevisiae* (grown in a media containing 50 g L^{-1} sucrose and the above described nutrients for 12 h at 32 °C and 150 rpm). Ethanol in the supernatant was determined by HPLC using the same method previously described for sugars content. Three replicates per sample were analyzed and the average ($n = 9$) was used in the development of NIRS prediction model.

2.3. NIR data collection

A FOSS-NIRSystems 6500 monochromator (Silver Spring, Maryland, USA) equipped with a Direct Contact Food Analyzer Module (DCFA) with a 42 × 27 mm window and provided with six cuvettes and a gold reflector with path length of 0.1 mm (Sampling Kit NR-6543) was used to measure reflectance spectra from 400 to 2498 nm, every 2 nm.

Spectral absorbance values were recorded as $\log 1/R$, where R is the sample reflectance. Circular quartz cuvettes (48 mm Ø) were used for reflectance analysis. As the glass window of the DCFA module is rectangular and the cuvettes are circular, a centering device with the right diameter size was used so that the cuvette was centered on. After being thawed, approximately 40 g of sugar beet brei were used to fill about half of the cuvette. During the sample analysis, the internal white ceramic used as reference material in the NIR instrument was scanned 16 times and afterwards the beet sample was scanned 32 times, obtaining a final spectrum for each cuvette and sample. The number of replicates necessary to obtain a representative spectrum of a given sample was also optimized. For that purpose, the 50 samples belonging to the 2007 campaign were analyzed by filling four cuvettes per sample. The mean spectra of the four spectra and only the two first spectra obtained with each one of the 50 samples were used to develop initial prediction models of bioethanol yield. After optimization of the number of replicates per sample, two cuvettes per sample were filled and the average spectrum was used in the data analyses. Spectral data were collected by using WINISI II package version 1.50 (Infrasoft International, Port Matilda, PA, USA).

2.4. Data processing and calibration development

The WinISI II software package was also used for the chemometric management of data. Prior to NIRS calibration development, the CENTER algorithm was applied to determine the structure and spectral variability of population (Shenk and Westerhaus, 1991a,b). This algorithm consisted of a principal component analysis (PCA) of the samples and was used to reduce the dimensionality of the data matrix and to retain the maximal amount of possible variability in the spectral data. The scores are the values of the samples represented in the new n-dimensional space, and the loadings are the coefficients of the combined m original variables, i.e. absorbance values ($\mu \mathrm{Log}\, 1\ R^{-1}$) from 400 to 2500 nm. It is common in NIRS approaches to apply pre-treatments designed to correct for scatter effects usually seen in absorbance data as well as derivatives prior to the PCA or calibrations. For structuring the population, an initial PCA was performed by using Standard Normal Variate and Detrending (SNV + DT) methods (Barnes et al., 1989), together with Norris–Williams spectral mathematical derivation treatment "1,5,5,1". The CENTER algorithm also calculates the GH value (Mahalanobis global distance to the center of the population) of each sample, sorts them according to it and considers as outliers those samples with GH value 3.0 or above, which are eliminated (Shenk and Westerhaus, 1991a). The population was divided into calibration and validation sets in order to perform an external validation. For that purpose, one of each six sorted samples from the resulting set of the CENTER algorithm was selected to build up the validation set, and the rest remained as the calibration set (Flores et al., 2009). In this way, the samples of the validation set are consistently distributed in the population.

Prediction equations were obtained using the modified partial least squares (MPLS) method of regression owing to MPLS is often more stable and accurate than the standard PLS algorithm. Cross validation was used to select the optimal number of factors and to avoid overfitting (Shenk and Westerhaus, 1995a). Cross validation was performed by splitting the calibration set into four groups according to Shenk and Westerhaus (1995b, 1996). All multivariate regression equations were obtained using the Standard Normal Variate and Detrending (SNV + DT) (Barnes et al., 1989) and Multiplicative Scatter Correction (MSC) (Geladi et al., 1985) algorithms for scatter correction. Both algorithms removed artifacts or imperfections (e.g. undesirable scatter effect) from the data matrix prior to data modeling (Rinnan et al., 2009). In addition, Norris–Williams spectral derivatives were used to remove additive and multiplicative effects in the spectra (Norris and Williams, 1984) after a smoothing of the raw spectra and application of a first or second