



Multi-product calibration models using NIR spectroscopy



M.K.D. Rambo^a, M.M.C. Ferreira^b, E.P. Amorim^c

^a Department of Chemistry, University of Tocantins—UFT, CEP 77838-824 Araguaína, Brazil

^b Institute of Chemistry, University of Campinas—UNICAMP, CEP 13083-970 Campinas, Brazil

^c Embrapa Cassava & Fruits, 44380-000 Cruz das Almas, Brazil

ARTICLE INFO

Article history:

Received 30 September 2015

Received in revised form 11 December 2015

Accepted 16 December 2015

Available online 30 December 2015

Keywords:

PCA

PLS

Chemical composition

Banana

Coffee

Coconut

ABSTRACT

The physical–chemical composition of multiple biomasses can be predicted from one single calibration model instead of compositional prediction conducted by individual models. In this work, multi-product models, involving banana, coffee and coconut samples were built by partial least square regression (PLS) for ten different chemical constituents (total lignin, klason lignin, acid insoluble lignin, acid soluble lignin, extractives, moisture, ash, glucose, xylose and total sugars). The developed PLS models show satisfactory results, with relative error (RE%) less than 20.00, except for ash and xylose models; ratio performance deviation (RPD) values above 4.4 and range error ratio (RER) values above 4.00. This means that all models are qualified for screening calibration. Principal component analysis (PCA) was useful to demonstrate the possibility and the rationale for combining three biomass residues into one calibration model. The results have shown the potential of NIR in combination with chemometrics to quantify the chemical composition of feedstocks.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Near infrared spectroscopy (NIR) has received considerable attention in the last years, as a tool for rapid, non-destructive, non-expensive (1–5% of the wet chemistry procedure cost), of simple application and that allows simultaneous assessment of multiple parameters of biomass composition [1,2]. The combination of NIR with chemometric tools allowed the development of multivariate calibration models for the rapid analysis of the chemical composition of feedstocks [3–7].

To ensure reliable prediction using the correlation of NIR spectra with the reference data from biomass composition, the NIR methods must be calibrated to an accurate primary reference analytical method. For this initial calibration, advanced multivariate models are developed, and although the process cost is slightly increased (30% of the wet chemistry procedure), they are still lower than the wet analysis [1]. Besides, another question raised when building calibration models, is the necessity to have a large variability of the calibration population and of the chemical characteristic of the samples [8].

Most frequently this variability is reached by sampling over different times and locations, what increases the process costs. To avoid such additional costs, some authors have used different botanical fractions from biomass to increase the variability in calibration models [8–10]. One promising alternative for increasing sample variability would be to use

various feedstocks. However, literature [3,8] is scarce on the use of multi-biomass calibration models in which one single model combining different biomasses is developed.

According to Liu et al. [8] the main difficulty in building such models is associated to the dissimilarity among biomasses (different NIR spectra). It is not practical to develop a NIR calibration model with species showing large dissimilarity. So, to ensure a good prediction and reliable result, principal component analysis (PCA) was performed [11] to justify the development of a single calibration model containing three different biomasses. Besides, the usual statistical parameters (calibration and validation plots, calibration and validation errors, among others) were used to ensure the confidence of the models.

This study has shown that it is feasible that the arduous and costly process of sample collection over different times and from different locations was effectively replaced in a simple manner to use different types of biomass wastes to build single multivariate predictive models to analyze multiple constituents. Three quite distinct feedstocks (coffee, banana and coconut) and also different botanic fractions of each plant were considered. So, from the 10 different parameters (total lignin, klason lignin, acid insoluble lignin, acid soluble lignin, extractives, moisture, ash, glucose, xylose and total sugars) of physical–chemical composition analyzed, one model was built for each constituent, but useful for three singular feedstocks.

It proves that the NIR associated to multivariate analysis can be used for screening calibration and quality control to estimate physical–chemical content in biomass residues.

E-mail address: magalarambo@mail.uft.edu.br (M.K.D. Rambo).

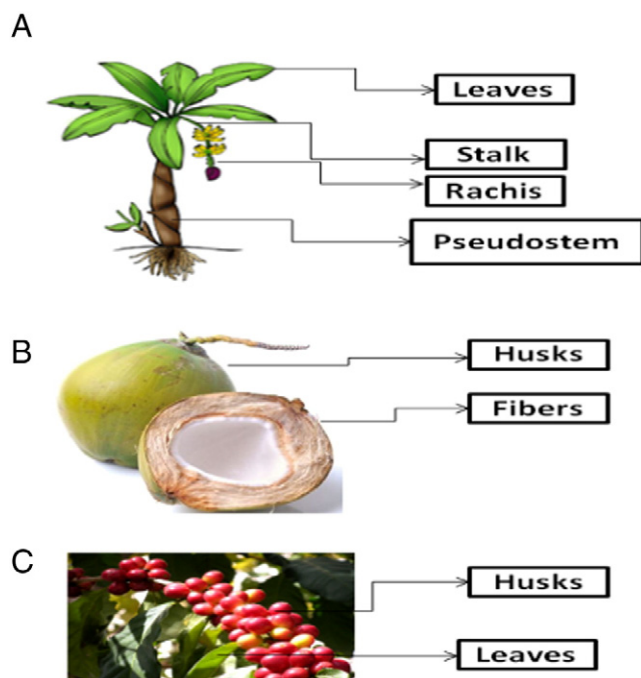


Fig. 1. Botanical fractions sampling of (A) banana (B) coconut and (C) coffee.

2. Material and methods

2.1. Sample collection

A total of 104, 101 and 28 samples of banana, coffee and coconut residues of different botanical parts were collected as illustrated in Fig. 1.

Also, among the different fractions, samples from different locations, soils, cultivars, species and harvest time were sampled to ensure the variability.

Of the 233 samples collected, not all were subjected to the wet analysis steps. All the 233 samples were analyzed for moisture, extractive and ash. The analyses of soluble and insoluble lignin contents were carried out for 137 samples, and for sugars only 94 samples were analyzed.

2.2. Physical–chemical analysis

All the samples were dried, mill and then sieved to a homogeneous particle size of 180–850 μm . The biomass analyses (all in duplicate) of extractives, lignins and sugars were carried out using standard National Renewable Energy Laboratory (NREL) methods [12,13]. For extractives (NREL/TP-510-42619, 2008), the accelerated solvent extraction with 95% ethanol in a Dionex ASE 200 system (Thermo Fisher Scientific, Waltham, MA, USA), was used. Acid hydrolysis (NREL/TP-510-42618, 2011) on the extracted samples was carried out with sulfuric acid 72% in a water bath in the first step, followed by hydrolysis for 1 h at 120 $^{\circ}\text{C}$ (in autoclave) and an acid concentration of 4%. In the hydrolysis step the lignin (soluble and insoluble) and sugar contents were determined. The acid soluble lignin (ASL) content was determined by UV-spectroscopy in a Shimadzu UV-1700 spectrometer (Shimadzu, Kyoto, Japan), at wavelength of 205 nm. Insoluble lignins (klason lignin (KL) and acid insoluble residue (AIR) were determined by gravimetry, and sugars were determined by high pH anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD) for the monosaccharides.

The moisture (105 $^{\circ}\text{C}$) and ash (600 $^{\circ}\text{C}$) analyses were carried out using ASTM 3173-87 [14] and ASTM D 3174-04 methods [15], respectively.

2.3. Multivariate calibration models

The Vis-NIR spectra (400–2500 nm) were collected using a FOSS XDS instrument (FOSS, Hillerød, Denmark). Each spectrum was generated by averaging 32 scans, with 0.5 nm of increment. Two spectra were collected for each sample and the average spectrum was used for data analysis.

Initially, all the 233 raw spectra were submitted to PCA with varimax rotation to reveal the data structure and identify similarity/dissimilarity among the three feedstocks.

Partial least squares regression (PLS1) was used to obtain the multivariate calibration models using the Unscrambler 10.2 (Camo Software, Oslo, Norway). The data set was randomly split into two subsets: the calibration set consisting of 75% of the samples and the external validation set with the remaining 25% of samples. The external validation set may be used to determine the number of latent variables (LV), and is often cited as the most realistic estimate, particularly of the prediction errors. However, it requires a large amount of samples [16,17], such as in the present study. These models were developed with the spectra transformed by taking the Savitzky–Golay second (2D) derivative using a second-order polynomial, with a window of 15 and 25 points [18]. For the extractive model, the best results were obtained by combining the standard normal variate (SNV) with first (1D) derivative transformations using a second-order polynomial, with a window of 2 points [19].

For each model, the coefficient of determination (R^2_{cal} and R^2_{val}), the root mean square error of calibration (RMSEC), the root mean square error of prediction (RMSEP), the standard error of calibration (SEC), the standard error of prediction (SEP) and the numbers of outliers and LV, were obtained. The error vector, \mathbf{e} , which is the difference between

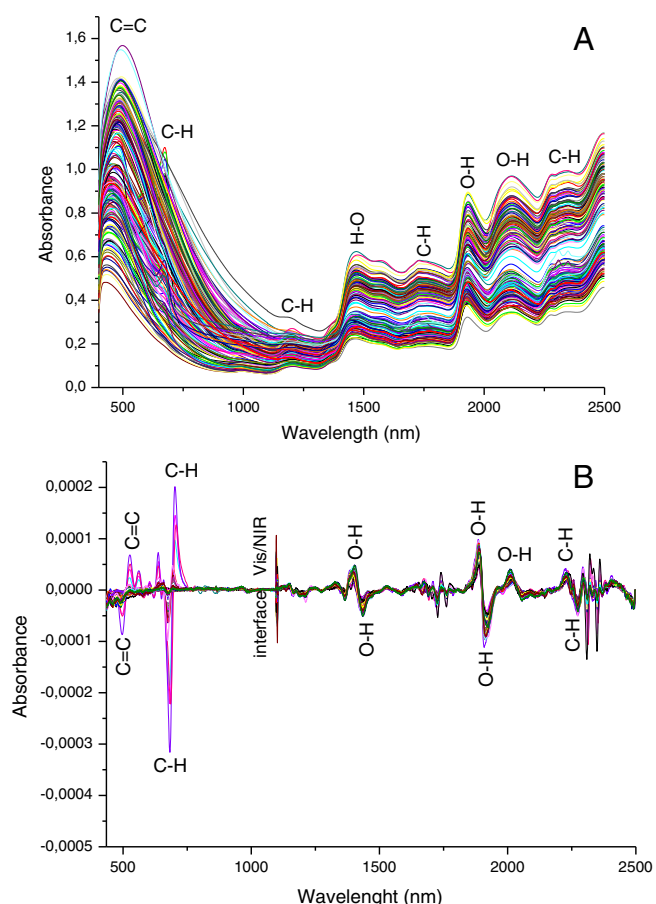


Fig. 2. Vis-NIR raw spectra (A) and Vis-NIR second derivative spectra.

Download English Version:

<https://daneshyari.com/en/article/1179357>

Download Persian Version:

<https://daneshyari.com/article/1179357>

[Daneshyari.com](https://daneshyari.com)