Food Chemistry 123 (2010) 800–805

Contents lists available at ScienceDirect

Food Chemistry

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

Analytical Methods

The development of near-infrared spectroscopy (NIRS) calibration for prediction of ash content in legumes on the basis of two different reference methods

Milica Pojić *, Jasna Mastilović, Dragan Palić, Mladenka Pestorić

University of Novi Sad, Institute for Food Technology, Bulevar cara Lazara 1, 21000 Novi Sad, Serbia

ARTICLE INFO

Article history:
Received 14 July 2009
Received in revised form 20 April 2010
Accepted 1 May 2010

Keywords:
Ash content
NIRS
MPLS
Gravimetric method
Thermogravimetric method

ABSTRACT

The aim of this study was to develop optimal NIRS calibration for ash content prediction in legumes by using the thermogravimetric (TGA) and gravimetric (GA) analytical methods. The calibration was performed on the basis of whole and structured sample sets ($n = 143$ and $n = 99$, respectively). Samples were scanned using a Rapid Content Analyzer in reflectance mode (400–2500 nm). Different mathematical treatments of the spectra preceded modified partial least squares (MPLS) regression analyses. The performance of the models was assessed by cross validation and external validation ($n = 44$). Models developed for the whole sample set on the basis of the TGA and GA methods were characterised by *standard error of calibration* (SEC) ranged from 0.28 to 0.50, *standard error of cross validation* (SECV) ranged from 0.43 to 0.60, *coefficient of determination* (R^2) ranged from 0.97 to 0.89, *explained variance* ($1 - \mathrm{VR}$) ranged from 0.94 to 0.85 and *residual predictive deviation* (RPD) ranged from 4.23 to 2.68, respectively. Models developed for the structured sample set on the basis of the TGA and GA methods were characterised by *standard error of calibration* (SEC) ranged from 0.32 to 0.42, *standard error of cross validation* (SECV) ranged from 0.53 to 0.56, *coefficient of determination* (R^2) ranged from 0.97 to 0.94, *explained variance* ($1 - \mathrm{VR}$) ranged from 0.91 to 0.89 and *residual predictive deviation* (RPD) ranged from 3.52 to 2.98, respectively. The obtained results showed the potential of NIRS method to accurately predict the ash content of legume grass samples that correspond to ash content determined by the TGA and GA methods.

1. Introduction

Ash represents the total amount of the inorganic content of forage and is used to estimate energy and calculate non-fibre carbohydrate content. Measurement of ash may also reveal the contamination which might have occurred initially due to soil collected along with the sample (Hoffman, 2005; Vazquez de Aldana, Garcia-Criado, Garcia-Ciudad, & Perez-Corona, 1996; Windham, Hill, & Stuedemann, 1991). The official and widely-accepted method for determining ash content in feed is the gravimetric method (GA) and this is routinely performed in forage testing laboratories. Research into development of new analytical procedures with more desirable characteristics for routine analysis has been stimulated by the numerous sources of uncertainties associated with measurement of ash. Although these newer methods are characterised by higher sample throughput, easier operation and better precision, the gravimetric method is still widely used as a benchmark method. The thermogravimetric method (TGA) has been exploited for ash determination in different kinds of food (including coffee, milk powders, starches, flours and oil seeds) and has advantages in terms of shorter analysis time, use of smaller sample sizes and no sample pretreatment (Felsner, Cano, Matos, de Almeida-Muradian, & Bruns, 2004; Tomassetti, Campanella, & Aureli, 1989; Windham et al., 1991). Near-infrared spectroscopy (NIRS) is now increasingly being used for the determination of feed quality. NIRS is a secondary analytical technique requiring proper instrument calibration with samples that adequately represent the population range for which it is to be used. NIRS calibration development must be based on widely-accepted analytical procedures, but these are frequently poorly-defined chemically and sometimes do not relate to spectroscopic data (Shenk & Westerhaus, 1994). Theoretically, inorganic substances do not absorb energy in the near infrared (NIR) region but some papers have demonstrated that NIRS can reliably predict the ash content of feed (Deaville & Flinn, 2000; Mentink, Hoffman, & Bauman, 2006; Osborne, 2001; Park, Agnew, Gordon, & Steen, 1998; Pérez-Marín, Garrido-Varo, Guerrero-Ginel, & Gómez-Cabrera, 2004). Since ash content cannot be directly measured by NIRS, it is assumed that it is predicted by correlation with the total amount of organic compounds and water present because of the large number of wavelengths used in the process of calibration development that give significant information (Clarke, Mayland, & Lamb, 1987; Frankhuizen, 2001; Garnsworthy, Wiseman, & Fegeros, 2000; Osborne, 2007). However, a number of

---

* Corresponding author. Tel.: +381 21 4853811; fax: +381 21 450725.
E-mail address: milica.pojic@fins.uns.ac.rs (M. Pojić).

0308-8146/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.foodchem.2010.05.013

literature reports describing the successful measurement of ash content of flour by NIRS led to the method being included amongst the methods approved by the American Association of Cereal Chemists. It has recently been demonstrated that ash calibrations are matrix sensitive and require frequent recalibration (Osborne, 2007; Williams, 2007).

The aim of the present work was to optimise NIRS calibration for ash content determination in legume grass samples using two different reference methods, TGA and GA, on the basis of randomly- and spectrally chosen sample sets. Taking into consideration the previously reported results, the aim of the study and its approach and methodology, the general hypothesis was formulated and examined: Near-infrared spectroscopy, coupled with appropriate TGA and GA calibration models, can be used to predict the ash content of legume grass samples.

2. Materials and methods

2.1. Samples

A total of 143 legume samples were collected in 2008 from two localities in Serbia; specifically 91 samples from Novi Sad and 52 from Kruševac. All were pure cultivars bred and grown in the experimental fields of the Forage Crops Department of the Institute of Field and Vegetable Crops, Novi Sad (located between 19°51′ of the East longitude and 45°20′ of the North latitude at the altitude of 87 m) and the Institute of Forage Crops in Kruševac (located between 21°19′ of the East longitude and 43°35′ of the North latitude at the altitude of 166 m). The Novi Sad site had a slightly calcareous chernozem soil, with that in Kruševac being alluvium of heavy mechanical composition. Average rainfalls for Novi Sad and Kruševac in 2008 were 528 and 647 mm, respectively, and average temperature was 12.7° and 12.9 °C, respectively. The botanical composition comprised 90 samples of alfalfa (*Medicago sativa* L.), 24 samples of red clover (*Trifolium pratense* L.), 12 samples of white clover (*Trifolium repens* L.), 9 samples of Persian clover (*Trifolium resupinatum* L.) and 8 samples of bird's-foot-trefoil (*Lotus corniculatus*). The sampling was conducted in four cuts at four phenological stages (vegetative, 10% bloom, 50% bloom and full bloom); samples were hand cut with scissors at 5 cm height, dried at 60 °C in a forced air oven for 24 h and ground to pass a 1 mm screen, before being used for chemical analyses and NIRS.

2.2. Chemical analysis

Samples were analysed in duplicates and averaged. The GA method was performed according to ISO methodology (ISO, 2002). The TGA method was performed using a Leco thermogravimetric analyzer TGA701 using ceramic crucibles, sample mass of about 1.0 ± 0.1 g, dynamic air atmosphere with high flow rate (10 dm^3/min), ramp rate of 20 °C per minute and in the temperature interval of 25–550 °C, until the constant weight was reached.

2.3. NIRS analysis

NIRS analyses were conducted with a XDS Rapid Content Analyzer (FOSS) over the spectral range 400–2500 nm using a Natural Product Cell (length 214 mm, depth 56 mm, height 44 mm, scanning surface 192 mm^2) to obtain a larger scanning surface. Duplicate scans of ground samples obtained by sample repacking were collected and averaged. Spectra were collected and managed using ISIScan software (Infrasoft International Port Matilda, PA) and calibration models were developed using WinISI software, version 1.50.

To develop NIRS calibration models for prediction of ash content, two calibration sets that encompassed samples that were likely to be encountered in routine analysis were employed: *the whole sample set* contained 143 randomly-chosen samples and the *structured sample set* contained 99 spectrally-chosen samples. This experimental design was chosen because a sample population could contain spectrally similar samples whose characterisation, in terms of the reference analysis, could be expensive and unjustified (Shenk & Westerhaus, 1991).

Principal component analysis (PCA) was performed before the application of modified partial least squares (MPLS) regression to reduce the spectral data and derive the first 10 principal components in order to examine the possible grouping of samples and to detect the spectral outliers as well. Samples with the global Mahalanobis distance (GH) <3.0 from the mean of all spectra were eliminated as spectral outliers ($n = 3$). Also, samples with neighbourhood Mahalanobis distance (NH) smaller than 0.6 between the neighbouring samples were eliminated from the initial sample set in order to form the structured sample set, whilst eliminated samples were used to create the *external validation sample set* ($n = 44$). Prior to modified partial least squares (MPLS) regression, standard normal variate and detrend (SNV–DT) scatter correction was applied. Twenty two mathematical treatments (D,G,S,S), varying in terms of order of the derivative (D), the gap over which the derivative was calculated (G) and the number of data points used to smooth the data (S,S), were tested during development of the NIRS calibrations that resulted in 22 calibration models. Four cross validation groups were used to select the optimal number of partial least squares (PLS) terms. The procedure for outlier elimination was performed in two passes prior to the completion of final calibration model on the basis of t statistics. Application of t statistics indicated five samples with high residuals ($t > 2.5$) which were excluded from the further analysis. The standard error of calibration (SEC), the coefficient of determination in calibration (R_c^2), the standard error of cross validation (SECV), and explained variance (1 − VR) were calculated to evaluate the predictive ability of the models. The residual predictive deviation (RPD) defined as the ratio between the standard deviation of the population (SD) and the standard error of cross validation (SECV) was used to test the accuracy of the calibration models developed. If the RPD is <3, the calibration models are considered as acceptable for analytical purposes (Williams, 2001). Moreover, the standard error of prediction (SEP), bias, slope and the coefficient of determination in validation (R_v^2) were calculated on the basis of external validation to evaluate the performance of models developed for the structured sample set. The significance of differences between the parameters of predictive ability of the models developed for the whole and structured sample sets was determined by *T*-test.

Table 1
Descriptive statistics for calibration and validation sets with regard to ash content (%) determined by the thermogravimetric (TGA) and gravimetric (GA) methods.

Method	*n*	Range	Mean	SD
Whole sample set				
TGA	143	6.60–15.46	9.34	1.82
GA	143	5.69–13.12	8.11	1.61
Structured sample set				
TGA	99	6.66–15.33	9.45	1.87
GA	99	5.39–13.12	8.21	1.67
Validation sample set				
TGA	44	6.60–15.27	9.11	1.72
GA	44	5.65–12.58	7.92	1.57

SD – standard deviation.