ORIGINAL SCIENTIFIC PAPER

The potential of NIR spectroscopy for crude protein prediction in spelt (*Triticum spelta* L.)

Marina VRANIĆ, Ana POSPIŠIL, Krešimir BOŠNJAK, Goran PERČULIJA, Hrvoje KUTNJAK, Josip LETO, Damir SOKOLOVIĆ

University of Zagreb, Faculty of Agriculture, Svetošimunska cesta 25, 10000 Zagreb, Croatia, (e-mail: mvranic@agr.hr)

Abstract

Near-infrared (NIR) spectroscopy (1100 – 2500 nm) was used to analyze the crude protein (CP) content of spelt (*Triticum spelta* L.). Modified partial least square (MPLS), principal component regression (PCR) and partial least square (PLS) techniques were used. Standard errors of calibration (SEC) were 5.57, 5.38 and 6.02 for MPLS, PLS and PCR, respectively, while standard errors of cross validation (SECV) were 6.19, 6.34 and 6.47 respectively. Comparing NIRS and the chemical procedure the standard error of prediction (SEP) for MPLS, PLS and PCR were 9.41, 8.87 and 9.50 respectively.

It was concluded that NIR spectroscopy has a great potential for CP prediction in spelt and that the PLS method is superior to the PCR and MPLS methods.

Key words: crude protein, spelt, NIR spectroscopy, PLS, MPLS, PCR

Potencijal NIR spektroskopije za procjenu sadržaja sirovih proteina pira (*Triticum spelta* L.)

Sažetak

Refleksijska spektroskopija u bliskom infracrvenom području (NIR spektroskopija) (1100 – 2500 nm) je korištena za procjenu sadržaja sirovih proteina (SP) uzoraka pira (*Triticum spelta* L.). Za razvoj kalibracije su korištene: metoda parcijalnih najmanjih kvadrata (PLS), modificirana PLS metoda (MPLS) i regresijska metoda glavnih komponenata (PCR). Standardna greška kalibracije (SEC) za MPLS, PLS i PCR metodu je iznosila 5,57; 5,38 i 6,02 respektivno, a standardna greška unakrsne validacije (SECV) 6,19; 6,34 i 6,47 respektivno. U usporedbi NIR i referentne kemijske metode za utvrđivanje sadržaja SP, standardna greška procjene (SEP) za MPLS, PLS i PCR metodu je iznosila 9,41; 8,87 i 9,50 respektivno.

Zaključeno je da NIR spektrsokopija ima veliki potencijal za procjenu sadržaja SP u uzorcima pira, te da je PLS metoda modeliranja ima prednosti u odnosu na PCR i MPLS metodu.

Ključne riječi: sirovi proteini, pir, NIR spektroskopija, PLS, MPLS, PCR

Introduction

Near-infrared spectroscopy (NIRS) has been widely used in quality control in plant and grain material for different purposes due to its properties: speed, accuracy, precision and non-destructiveness (Fahey and Hussein, 1999). One of the main fields of application of NIRS technology is the quality determination of forages (Shenk and Westerhaus, 1995) as fresh grass silage (Vranić et al., 2005) or semi-natural grassland (Vranić et al., 2010). This technique has been widely used in forage analysis since the 1970s (Norris et al,. 1976). The implementation of NIR technology has been growing because of the development of chemometric

procedures used for the calibration of these instruments (Osborne et al., 1993).

Spelt (*Triticum spelta* L.) is close relative of wheat, but the grain resembles barley in appearance. The hulls are not usually removed from the kernels in threshing. Spelt resemble oats in nutritive value. They may be used in the same manner as oats in feeding the various categories of livestock. Crude protein (CP) calibrations of wheat samples at different growth stages in different plant parts have been reported (Noaman and Taylor, 1990) while the prediction ability of NIR for CP concentration in spelt has not been reported.

The aim of the present work is to evaluate the ability of NIR spectra to predict CP concentration in spelt from NIR spectra.

Materials and methods

A calibration set of 36 spelt samples (Pospišil et. al, 2011) and a validation set of 13 samples were collected for this investigation. Dried spelt samples were grinded to pass 1mm screen. Total N concentrations was determined by the Kjeldahl method (AOAC 1990, ID 954.01) using a Gerhardt nitrogen analyzer. Additionally, N concentration was expressed as CP (total N x 6.25) g kg^{-1} DM.

The same samples were than re-dried at 105°C for at least 5 hours. Two independent scans were recorded from each of 49 samples, using a NIRSystem Model 6500 spectrometer (Foss-NIRsystem, Sweden) fitted with a sample transport module and a product sample cup (5 x 6.5 cm). Samples were scanned (32 scans co-aded) using the ISI SCAN Version 1.0 (Infrasoft International, Port Matilda, PA, USA) from 1100 to 2498 nm in reflectance mode (R mode: PbS detector) (Fig.1). Data were collected every 2 nm (700 data points per spectrum). The mean spectral value of each sample was calculated using the WIN ISI III Version 1.5.

Statistical analysis

In performing measurements with the NIRSystem Model 6500 spectrometer, a large number of data are generated for each sample and some data reduction method is needed to facilitate data interpretation. For transforming the measured spectral data into the sample properties producing model that describes the relationship between spectral data and CP, three methods were used: modified partial least square (MPLS), principal component regression (PCR) and partial least square (PLS). PLS regression is similar to PCR but uses both reference data (chemical, physical, etc.) and spectral information to form the factors useful for fitting purposes (Martens and Naes, 2001). MPLS is often more stable and accurate than the standard PLS algorithm. In MPLS, the NIR residuals at each wavelength, obtained after each factor has been calculated, are standardized (dividing by the standard deviations of the residuals at each wavelength) before calculating the next factor.

When developing equations, cross-validation is recommended in order to select the optimal number of factors and to avoid overfitting (Davies and Williams, 1996). Validation errors were combined into a standard error of cross-validation (SECV) (Shenk and Westerhaus, 1995). The multivariate regression equations were obtained using the standard normal variate and standard multiplicative scatter correction methods for scatter correction (Barnes et al., 1989). The statistics used to select the best equations were multiple correlation coefficients (RSQ) and the standard error of cross-validation (SECV). The prediction capacity of the model obtained was evaluated with the standard prediction error encountered in the NIR model (SEP).

Results and discussion

The CP concentration in spelt samples used to perform the investigation is shown in Table 1.

Table 1 Crude protein concentration in the calibration set of spelt samples (g kg^{-1} dry matter)

The name of the sample set	N	Minimum	Maximum	Mean	SD
Calibration and validation set	36	125.97	219.07	172.5	15.51
Selected calibration set	25	133.32	171.43	171.43	16.16

number of samples (N); standard deviation (SD)

The mean, SD and range of values of CP concentration are also included. The concentration of CP in spelt was between 125.97-219.07 g kg^{-1} DM. The randomly selected calibration set consisted of 25 samples that contained 133.32-171.43 g kg^{-1} DM of CP.

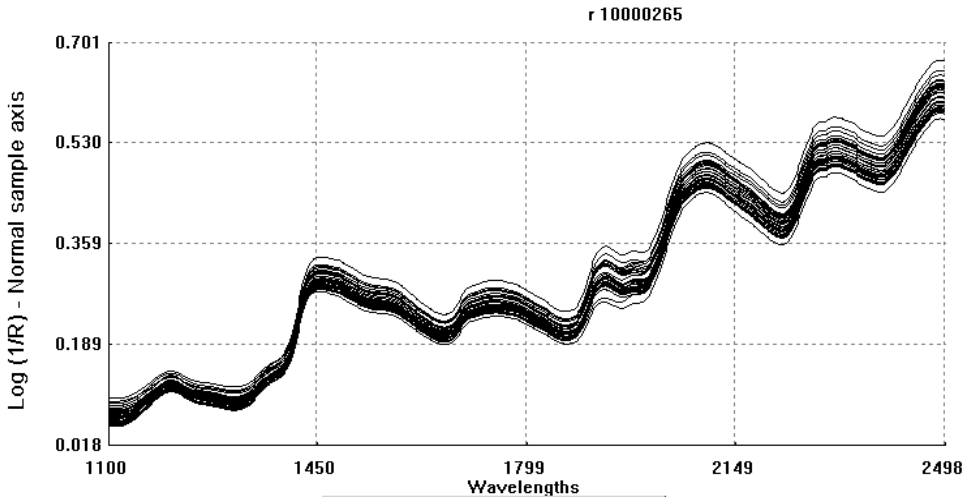

Figure 1 Original spectra of spelt for the calibration and validation set

Calibration equations

The statistical parameters of the calibration equations for CP are shown in Table 2, indicating N, the number of samples used to obtain the calibration equation after removing the samples for spectral (H criterion).

Table 2 Calibration statistics for spelt samples

Method applied	Number of samples	SEC	RSQ	SECV	SEP
PLS	24	5.38	0.87	6.34	8.87
MPLS	24	5.57	0.87	6.19	9.41
PCR	24	6.02	0.85	6.47	9.50

standard error of calibration (SEC); the correlation coefficient (RSQ);
standard error of cross validation (SECV); standard error of prediction (SEP)

To calibrate and validate CP, 36 samples of spelt were employed while for testing model repeatability additional 13 samples of unknown CP concentration. Calibration was carried out using NIRS technology and a remote reflectance fibre-optic probe applied directly on the sample with no previous treatment or manipulation.

Furthermore, the risk of mistake with the equations under practical conditions was very low or almost nil when using the standardised H statistic (Mahalanobis distance) during routine analysis of unknown samples (Martens and Naes, 2001). After the number of principal components had been calculated, detection of anomalous spectra was accomplished using the Mahalanobis distance (H statistic), establishing H>3 as the limit value for spectral reasons; 1 sample was removed.

Based on SEC and SEP the PLS method is the most reliable to predict CP in spelt while the lowest SECV is obtained using MPLS method.

Figure 2 shows the relationship between chemical (LAB) data and data calculated from near infrared (NIR) spectral data for crude protein (g kg^{-1} DM) for the 25 spelt samples investigated.

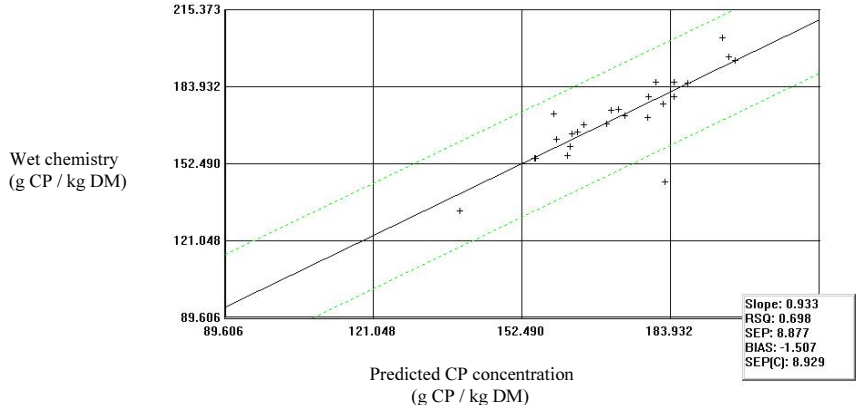

Figure 2
The predicted crude protein (CP) concentration and wet chemistry results for spelt samples obtained by PLS method

The present results corresponds with Noaman et al. (1988) who reported a good ability of the NIR for prediction of CP in wheat along the different growth stages. Lower RSQ in the present work for MPLS, PLS and PLR method (0.87; 0.87 and 0.85 respectively) than previously reported for predicting the CP in feed for cow (De Boever et al.,1994) might be result of lower number of samples used for calibration development.

Conclusions

The results show the great potential of NIR spectroscopy for CP prediction in spelt while the PLS method was the most reliable over the PCR and MLR methods applied.

Literature

AOAC (1990). Official methods of analysis of the association of Official Agricultural Chemists. 15th edn. (AOAC: Arlington, VA, USA).

Barnes R.J., Dhanoa M.S., Lister S.J. (1989). Applied Spectroscopy 43:772–777.

Davies A.M.C., Williams P. (1996). Near infrared spectroscopy: the future waves. Proceedings of the Seventh International Conference on Near Infrared Spectroscopy. NIR Publications, Chichester, West Sussex, UK

De Boever J.L., Cottyn B.G., Vanacker J.M., Boucque Ch.V. (1994). The use of nirs

to predict the chemical composition and the energy value of compound feeds for cattle. Animal Feed Science and Technology 51:243- 253.

Fahey G.C., Hussein H. S. (1999). Forty years of forage quality research: accomplishments and impact from an animal nutrition perspective. Crop Science 39:4–12.

Martens H., Naes T. (2001). Multivariate calibration. Wiley, Chichester, UK

Noaman M.M., Taylor G.A., McGuire C.F. (1988). Estimation of protein concentration in three plant parts at four growth stages in wheat by near infrared reflectance. Canadian Journal of Plant Science 68:217–220.

Noaman M.M., Taylor G.A. (1990). Vegetative protein and its relation to grain protein in high and low grain protein winter wheats. Euphytica 48:1–8.

Norris K.H., Barnes R. F., Moore J.E., Shenk J.S. (1976). Predicting forage quality by infrared reflectance spectroscopy. Journal of Animal Science 43:889–897.

Osborne B.G., Fearn T., Thindle P.H. (1993). Practical NIR Spectroscopy with Applications in Food and Beverage Analysis. Harlow, UK: Longman.

Pospišil A., Pospišil M., Svečnjak Z., Matotan S. (2011). Influence of crop management upon the agronomic traits of spelt (Triticum spelta L.). Plant Soil Environment 57:435-440.

Shenk J.S., Westerhaus M.O. (1995). Analysis of agriculture and food products by near infrared reflectance spectroscopy. Monograph NIRSystems Inc, 12101 Tech Road, Silver Spring MD 20904 USA

Vranić M., Knežević M., Seregély Z., Bošnjak K., Leto J., Perčulija G. (2005) Prediction of Dry Matter and Crude Protein Content in Fresh Grass Silage by Near Infrared Spectroscopy. Progress in Agricultural Engineering Sciences 1:57-75.

Vranić M., Knežević M., Leto J., Bošnjak K., Perčulija G., Kutnjak H., Knezić L., Teskera M. (2010). Procjena udjela sirovih proteina poluprirodnog travnjaka NIR spektroskopijom. 45. hrvatski i 5. međunarodni simpozij agronoma. S. Marić, Z. Lončarić (ur.) 950-954.

sa2012_0526