UDK 636.085

Izvorni znanstveni članak
Original scientific paper

PREDICTION OF DRY MATTER YIELD FROM SEMINATURAL GRASSLAND BY NIR SPECTROSCOPY

Marina Vranić, K. Bošnjak, S. Glavanović, J. Leto, Božica Lukšić

Summary

Near-infrared (NIR) spectroscopy (1100 – 2500 nm) was used to analyze the dry matter (DM) yield from seminatural grassland community Arrhenatheretum medioeuropaeum. Modified partial least square (MPLS), principal component regression (PCR) and partial least square (PLS) techniques were used for spectral data processing. Total of 225 forage samples were used in this investigation.
Standard errors of calibration (SEC) were 0.71, 0.71 and 0.77 for MPLS, PLS and PCR, respectively, while standard errors of cross validation (SECV) 0.80, 0.77 and 0.77 respectively. Comparing NIRS and the chemical procedure the standard error of prediction (SEP) for MPLS, PLS and PCR were 0.748, 0.698 and 0.873 respectively.
The results show the great potential of NIR spectroscopy for DM yield prediction in samples originated from semi natural grassland. Based on SEC, SEP and RSQ (R2) the PLS method is the most relailable to predict DM yield from seminatural grassland samples followed by MPLS method. The PCR method was at least reliable for DM yield prediction in this research..

Key words: dry matter yield, semi natural grassland, NIR spectroscopy, PLS, MPLS, PCR

Introduction

Knowledge of the amount of forage available to grazing or conservation is critical in planning animal production. It is a base for decisions about crop productivity, purchasing or selling forage, fertility and feeding, grazing schemes, and stocking rates. The amount of forage produced per hectare vary significantly from one location to another. These variations are influenced by properties such as soil type, soil moisture concentration, texture, nutrient supply, temperature, forage species and management (Ledgard and Steele, 1992).

A number of methods exist aimed at determining available forage in terms of dry matter (DM) yield. Actual clipping and weighing of forage from a given area is the most accurate method but is time consuming, requiring drying and weighing of clipped forage.

Near-infrared spectroscopy (NIRS) has been widely used in quality control in plant and grain material for different purposes due to its properties such as speed, accuracy, precision and non-destructiveness (Fahey and Hussein, 1999). This technique has been widely used in forage analysis since the 1970s (Norris et al,. 1976). One of the main fields of application of NIRS technology is the quality determination of forages (Shenk and Westerhaus, 1995) as DM and crude protein (CP) concentration (Vranić et al., 2005), chemical composition in semi-natural grassland (Vranić et al., 2010) or DM yield in grass-legume mixtures (Biewer et al. 2009). Red and especially the near infrared region is the most important for DM yield prediction (Biewer et al. 2009).

Prof.dr. Marina Vranić, Prof.dr. Krešimir Bošnjak (kbosnjak@agr.hr), Prof. dr. Josip Leto, University of Zagreb, Faculty of Agriculture, Department of Field Crops, Forage and Grasses, Grassland Research Centre, Svetošimunska cesta 25, 10000 Zagreb, Croatia; Božica Lukšić, student at the University of Zagreb, Faculty of Agriculture, Svetošimunska cesta 25, 10000 Zagreb, Croatia; Siniša Glavanović, Belupo Inc., Research and Development, Danica 5, 48000 Koprivnica, Croatia

The detection of total biomass of grassland ecosystems by spectral reflectance has been described in several studies (Boschetti et al., 2007; Friedl et al., 1994; Ikeda et al., 1999; Numata et al., 2007; Schino et al., 2003). Most studies used complex grassland types and information from satellite-based sensors that were not synchronized with harvesting of the swards. The majority of studies investigated ecosystems with low biomass production or obtained models with only a poor accuracy of prediction.

The primary objective of the present investigation was to evaluate the ability of NIR spectra to predict DM yield from seminatural grassland using three chemometric techniques.

1. Materials and Methods

Experimental design and plant sampling

The study was conducted at the Grassland Research Center owned by the University of Zagreb Faculty of Agriculture located. The Center is located at 638 meters above sea level in the Nature Park Medvednica.

In the ten year period (2003-2014) farm yard manure (FYM) was applied on the seminatural grassland community *Arrhenatheretum medioeuropaeum* to determine the effect of quantity, time and frequency of FYM application on dry matter (DM) yield. The experiment was conducted as a randomized block design with four replications and ten fertilizing treatments incorporation 30 and 50 kg ha^{-1} KSG each fall, each autumn, every third fall, every third autumn and a control treatments studied classical fertilization with mineral fertilizers (NPK 8:26:26) every fall in the amount of 500 kg ha^{-1} and KAN (27% N) in the quantity of 600 kg ha^{-1}.

FYM was manually applied and evenly distributed on plots 6 x 2 m each (12 m^2) 0.5 m spacing between plots.

The experimental plots were mowed with a self-propelled mower at the phase of tasseling the dominant grass to a height of 5 cm. To avoid marginal impact, the strip width of 1.2 m was harvested at the middle of each plot. The study used first spring cut forage samples from 2006, 2007, 2009, 2014 and 2015.

Besides DM, samples originated from 2006, 2007 and 2014 (total of 120 samples), were analysed on chemical composition.

Chemical analysis

Plant mass was weighed and subsamples (about 500 g) dried at a temperature of 60° C in a fan-assisted oven (ELE International) to a constant weight to determine DM concentration and finally DM yield ha^{-1}. Dried forage samples were grinded to pass 1 mm screen.

The ash contents were measured by igniting the samples in a microwave oven (Milestone PIYRO, Italy) at 550 °C for 3 h. The total N concentrations were determined by the Kjeldahl method (AOAC 1990, ID 954.01) using a Gerhardt nitrogen analyser. In addition, N concentration was expressed as crude protein (CP) (total N x 6.25) g kg^{-1} DM.

Neutral detergent fibre (NDF) and acid detergent fibre (ADF) contents were analysed using the procedure described by Van Soest et al. (1991) utilising ANCOM Filter bag technology (USA) with an ANCOM fibre analyser.

Spectral data collection

Total of 225 samples originated from seminatural grassland were used in the calibration development. A calibration set of 170 samples, validation set of 25 samples and a repeatability file consisting 30 samples were collected for this investigation.

Before scanning, the samples were re-dried at 105°C for at least 3 hours. Three independent scans were recorded from each of 225 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). 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.

Analysis of Spectral Data

In performing measurements with the NIRSystem Model 6500 spectrometer, hundreds 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 describe the relationship between spectral data and DM yield, 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 (Geladi and Kowalski, 1986). 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.

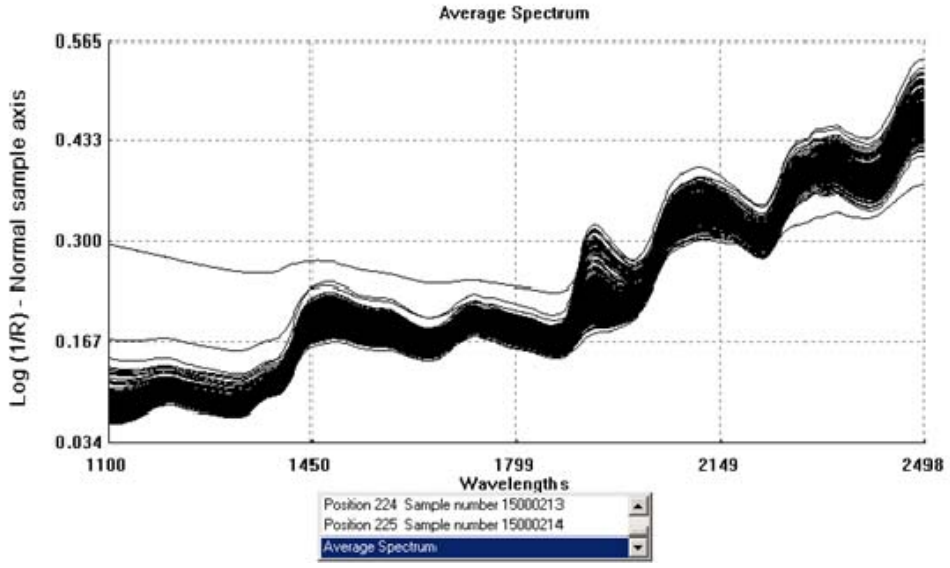
When developing equations, cross-validation is recommended in order to select the optimal number of factors and to avoid overfitting (Davies and Williams, 1996). Validation errors were combined into a standard error of cross-validation (SECV) (Shenk and Westerhaus, 1995). The multivariate regression equations were obtained using PLS, MPLS and PCR, whearas

the standard normal variate and standard multiplicative scatter correction were used for preprocessing of spectral data (Barnes et al., 1989). The statistics used to select the best equations were coefficient of determination (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).

2. Results and discussion

The spectra from samples scanned by reflectance (average of three readings) are presented in Figure 1.

Figure 1. – THE SPECTRA FROM SAMPLES SCANNED BY REFLECTANCE (AVERAGE OF THREE READINGS)



Dry matter (DM) yield (t ha^{-1} $year^{-1}$) and chemical composition of samples used in the experiment are presented in table 1.

Table 1 – THE AVERAGE DRY MATTER YIELD AND THE CHEMICAL COMPOSITION OF SAMPLES USED IN THE EXPERIMENT

Parameter	Minimum	Maximum	Mean	SD	Number of samples
DM yield (t ha^{-1})	0.86	11.12	3.60	1.52	225
CP (g kg^{-1} DM)	106.00	210.64	152.44	33.46	120
OM (g kg^{-1} DM)	882.41	940.33	909.53	12.70	120
NDF (g kg^{-1} DM)	381.63	658.34	498.73	71.92	120
ADF (g kg^{-1} DM)	283.57	458.33	342.47	28.24	120

DM, dry matter; CP, crude protein; OM, organic matter, NDF, neutral detergent fibre; ADF, acid detergent fibre; SD, standard deviation.

There were a great variations in DM yield as well as in chemical composition of samples used in the experiment. Kramberger et al. (2014) presented average annual herbage DM yield which ranged from 4.86 to 9.31 t ha^{-1} in the experiment on the effects of low precipitation periods on the herbage yield of semi-natural grasslands under different cutting regimes.

The requirements for the calibration samples are that the samples are similar enough to each other to represent a single population but diverse enough to allow independent resolution of each constituent of interest (Williams, 2001). Number of samples in the calibration set (N) is influenced by the natural variation in the trait of interest. The narrower the range, the more difficult it is to detect differences. Typically, 80-100 samples are required for developing an initial calibration, with up to multiple-hundreds of samples in a "mature" calibration. A "proof-of-concept" model will utilize 50-60 samples; fully developed prediction models can be built from no fewer than 80-100 samples, but this number can be greater (around 1000 samples) depending upon the error terms associated with each analyte. The final number of samples required is dependent upon the analytical and spectral diversity within the reference samples selected for developing the prediction model (Osborne et al. 1993).

To reduce total error, it is desirable to have multiple replicates of the sample analyzed by the reference method and scanned multiple times with the specific NIR instrument. If the calibration set is being developed for a dried, ground sample NIR instrument, then drying conditions must be standardized. Spectra production is quite sensitive to differences in sample particle size and shape. As a direct result, consistent sample preparation (e.g., grinding) is critical. One of the largest sources of error in NIR predictions between labs that use the same calibration is a result of differences in how the labs prepare the sample (e.g., different type or worn grinders; Williams, 2001).

To develop robust NIRS prediction models and valid results, laboratories must minimize sources of error in the entire process, obtain values for reference samples using analytical methods that have high precision and accuracy, standardize sample preparation and analytical procedures, standardize the NIRS instrument, perform routine instrument maintenance, analyze only samples representative of the original population and obtain routine diagnostics of all associated instruments and undergo yearly prediction model (calibration) updates (Williams, 2001).

Calibration equations

The statistical parameters using three chemometric techniques for DM yield from perennial grass samples are shown in Table 2.

Table 2 – CALIBRATION STATISTICS FOR DRY MATTER YIELD FROM PERENNIAL GRASSLAND

Method applied	SEC	SECV	SEP	RSQ
	t ha^{-1}			
MPLS	0.71	0.80	0.748	0.761
PLS	0.71	0.77	0.698	0.780
PCR	0.77	0.77	0.873	0.688

MPLS, modified partial least square; PCR, principal component regression; PLS, partial least square; SEC, standard error of calibration; RSQ, the coefficient of determination; SECV, standard error of cross validation; SEP, standard error of prediction.

Standard error of calibration (SEC) defines how well the NIRS prediction model predicts the reference values (calibration sample set) that was used to build the model. Low SEC values are desired. To calibrate and validate DM yield, 200 forage samples of semi natural grassland were employed while for testing the model repeatability additional 25 samples of unknown DM yield. 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, SEP and RSQ (R^2) the PLS method is the most relailable to predict DM yield from seminatural grassland samples followed by MPLS method. The equally low SECV was obtained using PLS and PCR chemometric method.

Prediction accurancy of PCR was slightly lower than that of MPLS (Table 2). The common MPLS equation achieved a R^2 of 0.761 with a SEC of 0.71.

The relationship between determined and spectral data show how reliable the chemometric method is. Figures 2, 3 and 4 show the relationship between the data obtained by destructive method data and data calculated from near infrared (NIR) spectral data for DM yield (t ha^{-1}) for the 170 forage samples investigated by MPLS, PLS and PCR method respectively.

Figure 2. – RELATIONSHIP BETWEEN MODIFIED PARTIAL LEAST SQUARE DATA CALCULATED FROM NEAR INFRARED (NIR) SPECTRAL DATA AND THE MEASURED DRY MATTER (DM) YIELD (t ha-1)

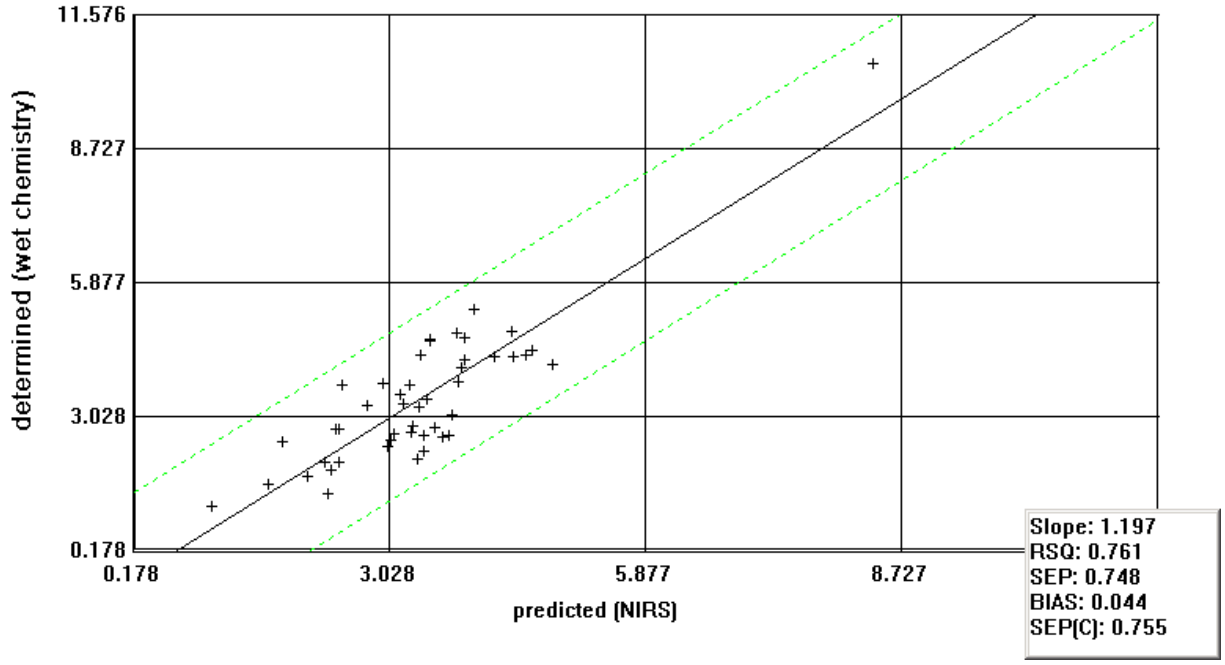


Figure 3. – RELATIONSHIP BETWEEN PARTIAL LEAST SQUARE DATA CALCULATED FROM NEAR INFRARED (NIR) SPECTRAL DATA AND THE MEASURED DRY MATTER (DM) YIELD (t ha-1)

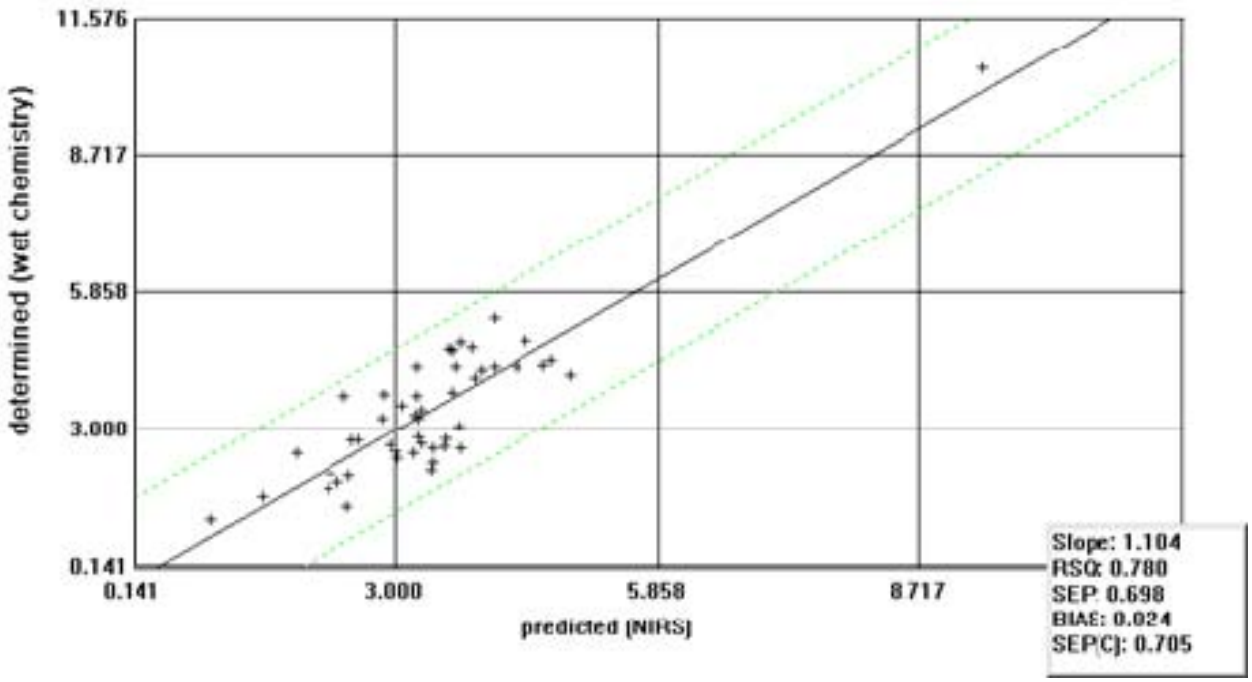
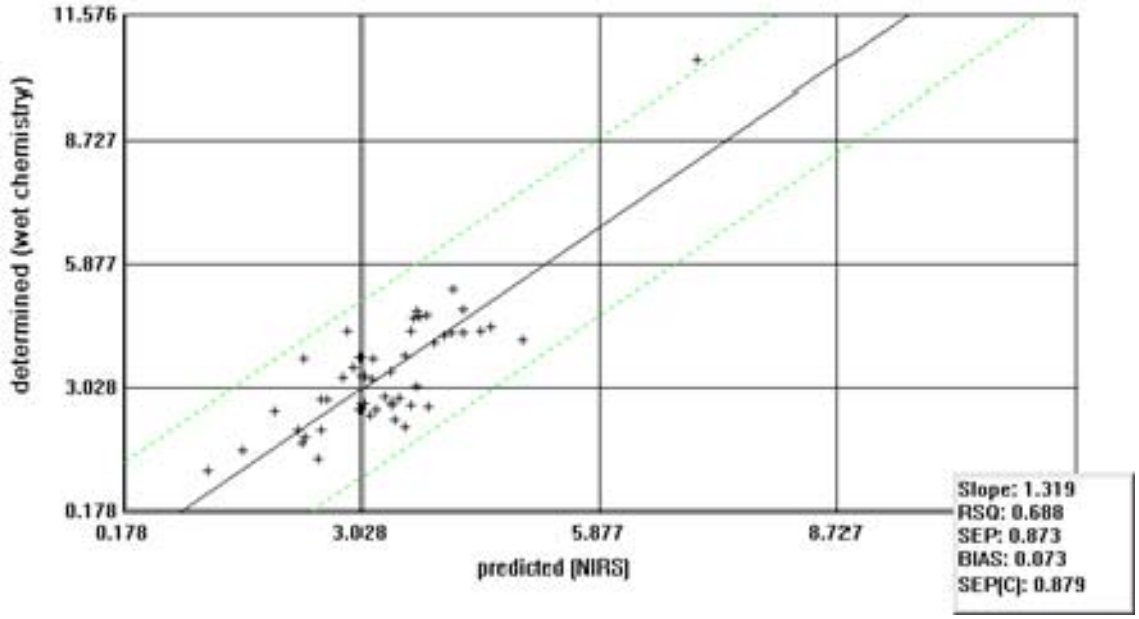


Figure 4. – RELATIONSHIP BETWEEN PARTIAL COMPONENT REGRESSION DATA CALCULATED FROM NEAR INFRARED (NIR) SPECTRAL DATA AND THE MEASURED DRY MATTER (DM) YIELD (t ha^{-1})



Regression coefficient (R^2 or RSQ) is the best fit line when predicted values are plotted against the associated reference values. High R^2 values are desired. An R^2 of 1.0 means 100% of the analyte variance is explained by the prediction equation. In a study assessing the yield of DM from legume - grass swards by field spectroscopy Biewer et al. (2009) have obtained higher accurancy results, and in that R^2 ranged from 0.82 to 0.97. The reasons for obtaining lower R^2 values with plant material might be the lower number of samples used in this study thus lower variability covered in DM yield within the calibration set. Only the DM yield of pure white clover sward that was difficult to determine showing the lower R^2 (0.56) (Biewer et al., 2009) which was under the values obtained in this paper (0.68 – 0.78). The reason might be the structure of white clover swards characterized by a dense layer of horizontally oriented clover leaves (Sanderson et al. 2006). Regularly dispersed leaves of white clover causes a high effective light extinction coefficient (Lantiga et al., 1999) resulting in a strong but undifferentiated pattern of absorption and reflection of light and thus contributing to the weak relationship between DM yield and spectral data.

The heterogenous plant structure caused by the high proportion of weeds and bare soil in the experimental swards and the wide range of the development stages may have confounded the relationship between NIR and DM yield. The plant growth stage influence the reflectance characteristics and model development by NIR. Despite, mowing was done on the nearly the same maturity stage, in 2014 the contents of the SP was lower, and the NDF higher compared to 2004, which could be associated with the term of mowing. However, the content of the SP and the NDF could be affected by the climate, ie. in relation to 2004, but also a multi-year average, 2014 had more rainfall and higher average temperatures. In addition to climate and harvest time, there are

other factors that could affect the quality of forage such as supply of soil nutrients, and the ratio of grass and clover. While developing calibrations, plants at advanced growth stage exhibited spectral attributed that differing strongly from those of less mature swards so the some of them were eliminated as outliers (Biewer et al., 2009)

The NIR is a secondary method based on a regression against a primary (or reference) method. Consequently, the NIR value can never be more accurate than the primary reference method.

Conclusions

The results show the great potential of NIR spectroscopy for DM yield prediction in samples originated from semi natural grassland. Based on SEC, SEP and RSQ (R^2) the PLS method was the most relailable to predict DM yield from seminatural grassland samples followed by MPLS method. The PCR method was at least reliable for DM yield prediction in this research.

REFERENCES

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

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

3. Biewer, S., T. Fricke, M. Wachendorf (2009.): Determination of dry matter yield from legume- grass swards by field spectroscopy. Crop Science, 49: 1927-1936.

4. Boschetti, M., S. Bocchi, P.A. Brivio (2007.): Assessment of pasture production in the Italian Alps using spectrometric and remote sensing. Agriculture, Ecosystems and Environment 118: 267-272.

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

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

7. Friedl, M.A., J. Michaelsen, F.W. Davis, H. Walker, D.S. Schimel (1994.): Estimating grassland biomass and leaf area index using ground and satellite data. International Journal of Remote Sensing 15: 1401-1420.

8. Geladi, P., B.R. Kowalski (1986.): Partial least-squares regression: a tutorial. Analytica Chimica Acta 185: 1-17.

9. Ikeda, H., K. Okamoto, M. Fukuhara (1999.): Estimation of aboveground grassland phytomass with a growth model using Landsat TM and climate data. International Journal of Remote Sensing 20: 2283-2294.

10. Kramberger, B., A. Gselman, M. Podvršnik, M. Lešnik, D. Škorjanc (2014.): Effects of low precipitation periods on the herbage yield of mesic semi-natural grasslands under different cutting regimes. Zemdirbyste-Agriculture, vol. 101, No. 1 (2014), p. 11–18

11. Lantiga, E.A., M. Nassiri, M.J. Kropff (1999.): Modeling and measuring vertical light absorption within grass-clover mixtures. Agric. For. Meteorol. 96: 71-83.

12. Ledgard, S. F. , K.W. Steele (1992.): Biological nitrogen fixation in mixed legume/grass pastures. Plant and Soil 141: 137-153.

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

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

15. Numata, I., D.A. Roberts, O.A. Chadwick, J. Schimel, F.R. Sampaio, F.C. Leonidas, J.V. Soares (2007.): Characterization of pasture biophysical properties and the impact of grazing intensity using remotely sensed data. Remote Sensing of Environment 109: 314-327.

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

17. Sanderson, M.A., K.J. Soder, N. Brezezinski, F. Taube, K. Klement, L.D. Muller, M. Wachendorf (2006.): Sward structure of simple and complex mixtures of temperate forages. Agronomy Journal, 98, 238-244.

18. Schino, G., F. Borfecchia, L. De Cecco, C. Dibari, M. Iannetta, S. Martini, F. Pedrotti (2003.): Satellite estimate of grass biomass in a mountainous range in central Italy. Agroforestry Systems 59: 157-162.

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

20. Van Soest, P.J., J.B. Robertson, B.A. Lewis (1991.): Method for dietary fiber, neutral detergent fiber and nonstarch polysaccharides in relation to animal nutrition. J. Dairy Sci. 74, 3583-3597.

21. Vranić, M., M. Knežević, Z. Seregély, K. Bošnjak, J. Leto, G. Perčulija (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.

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

23. Williams, P. (2001.): Sampling, sample preparation and sample selection. In Handbook of Near-Infrared Analysis edited by Donald A. Burns, Emil W. Ciurczak.

PROCJENA PRINOSA SUHE TVARI KRME POLUPRIRODNOG TRAVNJAKA NIR SPEKTROSKOPIJOM

Sažetak

U ovom istraživanju je korištena NIR spektroskopija (1100 – 2500 nm) za procjenu prinosa suhe tvari (ST) krme poluprirodnog travnjaka, zajednice Arrhenatheretum medioeuropaeum. Za obradu spektralnih podataka je korištena modificirana metoda parcijalnih najmanjih kvadrata (MPLS) (engl. modified partial least squares), regresija glavnih komponenata (PCR) (engl. principal component regression) i metoda parcijalnih najmanjih kvadrata (PLS) (eng. partial least square). Korišteno je ukupno 225 uzoraka voluminozne krme.

Standardne greške kalibracije (SEC) su iznosile 0.71, 0.71 i 0.77 za MPLS, PLS i PCR, respektivno, a standardne greške unakrsne validacije (SECV) 0.80, 0.77 i 0.77 respektivno. U usporedbi NIR spektroskopije i klasičnog utvrđivanja prinosa ST standardna greška procjene (SEP) za MPLS, PLS i PCR je iznosila 0.748, 0.698 i 0.873 respektivno.

Rezultati istraživanja ukazuju na veliki potencijal NIR spektroskopije za procjenu prinosa ST uzoraka poluprirodnog travnjaka. Bazirano na SEC, SEP i R2, PLS metoda se pokazala najpouzdanijom za procjenu prinosa ST, zatim MPLS i naposlijetku, najmanje pouzdana metoda je bila PCR metoda.

Ključne riječi: prinos suhe tvari, poluprirodni travnjak, NIR spektroskopija, PLS, MPLS, PCR.

Primljeno: 17.02.2016.