

## Prediction of chemical composition of South African *Medicago sativa* L. hay from a near infrared reflectance spectroscopy spectrally structured sample population

G.D.J. Scholtz<sup>1#</sup>, H.J. van der Merwe<sup>1</sup> and T.P. Tylutki<sup>2</sup>

<sup>1</sup>Department of Animal-, Wildlife-, and Grassland Sciences

P.O. Box 399, University of the Free State, Bloemfontein 9300, South Africa

<sup>2</sup>Agricultural Modeling and Training Systems, 418 Davis Rd Cortland, NY 13045, USA

---

### Abstract

The near infrared reflectance spectroscopy (NIRS) to predict chemical and digestibility parameters was investigated. Samples (n = 168) representing the spectral characteristics of the South African *Medicago sativa* L. hay population were chemically analysed for the development of calibration equations. Values for  $r^2$  and ratio of prediction to deviation (RPD) used as estimates of calibration accuracy for these parameters were classified as follows: good for dry matter (DM) ( $r^2 = 0.97$ ; RPD = 4.84), crude protein (CP) ( $r^2 = 0.97$ ; RPD = 4.57), acid detergent fibre (ADF) ( $r^2 = 0.95$ ; RPD = 3.97), neutral detergent fibre (NDF) ( $r^2 = 0.95$ ; RPD = 3.99), lignin ( $r^2 = 0.94$ ; RPD = 3.61), ash ( $r^2 = 0.93$ ; RPD = 3.12) and chloride (Cl) ( $r^2 = 0.95$ ; RPD = 3.74); intermediate for NDF-crude protein (NDF-CP) ( $r^2 = 0.91$ ; RPD = 2.96), sugar ( $r^2 = 0.91$ ; RPD = 2.82), *in vitro* organic matter digestibility at 24 hr (IVOMD24) ( $r^2 = 0.90$ ; RPD = 2.84) and 48 hr (IVOMD48) ( $r^2 = 0.89$ ; RPD = 2.70); and low (RPD < 2.31) for soluble protein (SP), ADF-crude protein (ADF-CP), fat, starch, NDF digestibility (NDFD) and the macro minerals (Ca, P, Mg, P, Na and S). The results recorded in the present study indicated that the NIRS technique is acceptable for DM, CP, ADF, NDF, lignin, ash and Cl analysis and for inclusion in quality models.

---

**Keywords:** Lucerne, fibre, protein, NIRS, calibration

# Corresponding author. E-mail: Scholtzgd@ufs.ac.za

### Introduction

As shown by Scholtz *et al.* (2009) the energy and protein values of lucerne (*Medicago sativa* L.) hay in South Africa could vary considerably and should be determined beforehand in order to accurately compose a balanced diet. Laboratory analysis is laborious, expensive and time consuming so that results often transpire late after consumption. With the recent introduction of near infrared reflectance spectroscopy (NIRS), fast evaluation of feeding quality is now a possibility.

The precision and accuracy of the predictions are critical to the acceptance of NIRS as an analytical tool. In practice, the reproducibility of the NIRS method is usually equal, and often superior, to that of the reference method (Williams, 2007). Furthermore, NIRS has less variance in analyses of the same sample than laboratory analyses (Shenk *et al.*, 1979). Timely decision making on strategic use of nutritional supplements or adjustments in ration formulation to efficiently sustain milk, meat or fibre production is also facilitated by this method. A major advantage of NIRS is its ability to analyse samples rapidly, non-destructively, and its non-polluting technology (Williams, 2001).

Norris *et al.* (1976) are credited with the first application of NIRS to the analysis of forages. Since then numerous workers have explored the use of NIRS for the prediction of both chemical composition and digestibility of lucerne hay (Van Waes *et al.*, 1998). However, limited calibrations have been developed and published for South African grown lucerne hay (Snyman & Joubert, 1992; Scholtz, 2001). There is an urgent need to ascertain whether NIRS could be relied upon to analyse all the lucerne hay feeding parameters used in modern nutritional models. This is especially important in developing a model for lucerne hay quality grading in South Africa

The objective of his study was to identify useful predictive relationships from a pre-selected spectrally structured lucerne hay sample population in South Africa.

## Materials and Methods

Six hundred lucerne (*Medicago sativa* L.) hay samples were collected from several commercial irrigation farms in the main lucerne producing areas in South Africa, which varied in location, soil characteristics (texture, organic matter, N content, pH) and farm management. The samples were collected and prepared as described by Scholtz *et al.* (2009).

Hundred and sixty eight samples that represented the South African Lucerne hay population were selected with a NIR-System Model 5000 scanning monochromator (Foss NIR Systems, Silver Springs, MD, USA). The selected 168 samples were analysed in duplicate by Cumberland Valley Analytical Services, Inc., Maugansville, Maryland, USA, as described by Scholtz *et al.* (2009).

Calibration equations were developed from the 168 samples selected from the population set ( $n = 600$ ), using modified partial least square regression (MPLS) (Shenk & Westerhaus, 1993) with internal cross-validation (NIRS 2, 1995) to encompass all of the chemical and spectral variation (Williams, 1987) and the physico-chemical characteristics that are likely to be found in the population (Williams & Cordiero, 1985).

Parameters in the mathematical processing were sought through trial and error in order to minimize the standard error of cross-validation, giving best results with the mathematical treatment 1, 4, 4, 1, which means 1: number of derivative of spectra, 4: extent of gap over which the derivative was to be calculated, 4: the smoothing of points, 1: second smoothing (almost never used and normally set as 1). Only terms were included in the equation that had an  $F$  statistic of more than 8 to minimize over-fitting of data.

The standard error of cross validation (SECV) was calculated by an internal validation of 33% of all samples randomly taken by the software routine, which was predicted by an equation based on a calibration with the remaining 67% of all samples. Since the cross-validation results are reliable indicators of equation performance (Shenk & Westerhaus, 1994), the lower SECV and standard error of calibration (SEC), together with their coefficient of determination ( $r^2$ ) values, were the criteria used for the final selection of the equations to be tested further.

Another measure for equation selection was the number of  $H$  outliers when the constituents were predicted in the South African lucerne hay population. According to Shenk & Westerhaus (1994) a low number of  $H$  outliers would indicate the span of the equation for predictive purposes. Murray (1993) as cited by Herrero *et al.* (1996) suggested that the number of  $H$  outliers when predictions are carried out, in a routine forage evaluation by NIRS, should not exceed 4%. Accordingly this was the limit used in the present study.

Williams (2007) stated that the  $r^2$  and the ratio of prediction to deviation (RDP) are the most meaningful statistics for appraisal of analytical efficiency by NIRS. Other calibration evaluation statistics used in the present study included the SEC and SECV.

## Results and Discussion

The most commonly used measures of the strength of the relationship between a dependent variable and independent variable are the  $r^2$ . However, the  $r^2$  statistic is population dependent (Weiss, 1993). This means that, when used for comparison of different models that were derived from different populations,  $r^2$  should not be the sole statistic used to evaluate the models (Weiss, 1993). Thus, Williams (2001) proposed the use of  $r^2$  together with RPD as the most meaningful statistics for appraisal of analytical efficiency by NIRS. However, interpretation of RPD in the literature revealed mixed results. Williams (2001) suggested minimum requirements for RPD values of 3.1 – 4.9 for screening and 5.0 – 6.4 for quality assurance. On the other hand Edney *et al.* (1994) considered RPD values of 2.5 – 3.0 adequate for screening, but values of 3.0 – 5.0 are required for quality assurance. Several authors are however of the opinion that RPD values higher than 2.5 in any equation are required for an acceptable predictability (Edney *et al.*, 1994; Mathison *et al.*, 1999).

According to these criteria the RPD obtained in the present study for DM content (4.84), CP (4.57), ADF (3.97), NDF (3.99) and lignin (3.61) could be considered as good and suggested that NIRS calibration models might be used for routine analysis of these parameters. Moderate NIRS calibration models, adequate for screening, were found for NDF-CP (2.96) and sugar (2.82) whereas poor models were reported for SP (2.08), ADF-CP (2.06), fat (1.82) and starch (1.51). The poor calibration performance of ADF-CP ( $SD = 0.49$ ), fat ( $SD = 0.31$ ) and starch ( $SD = 0.55$ ) could possibly be explained by a lack of variation in these parameters for lucerne hay. However, according to Williams & Norris (2001) criterion, RPD values vary

from very poor for starch, fat, ADFCP, SP; to poor for sugar and NDFCP; to fair for lignin, ADF, NDF, CP and DM (Table 1).

**Table 1** Calibration and prediction (cross-validation) statistics of quality parameters of South African lucerne hay

Parameter	SEC	r <sup>2</sup>	SECV	1-VR	SD	RPD
Dry matter (g/kg)	0.17	0.97	0.20	0.96	0.97	4.84
Crude protein (g/kg)	0.47	0.97	0.55	0.95	2.53	4.57
Soluble protein (g/kg)	0.52	0.87	0.68	0.77	1.42	2.08
Acid detergent fibre – crude protein (g/kg)	0.20	0.83	0.24	0.77	0.49	2.06
Neutral detergent fibre – crude protein (g/kg)	0.73	0.91	0.84	0.89	2.50	2.96
Acid detergent fibre (ADF) (g/kg)	0.98	0.95	1.12	0.94	4.44	3.97
Neutral detergent fibre (NDF) (g/kg)	1.38	0.95	1.62	0.94	6.47	3.99
Lignin (g/kg)	0.32	0.94	0.36	0.92	1.29	3.61
Fat (g/kg)	0.13	0.81	0.17	0.71	0.31	1.82
Starch (g/kg)	0.39	0.51	0.44	0.36	0.55	1.25
Sugar (g/kg)	0.45	0.91	0.54	0.87	1.51	2.82
Neutral detergent fibre digestibility at 24 hours (%NDF)	1.75	0.85	1.96	0.82	4.51	2.31
Neutral detergent fibre digestibility at 48 hours (%NDF)	2.29	0.78	2.64	0.72	4.93	1.87
<i>In vitro</i> organic matter digestibility at 24 hours (%)	1.31	0.90	1.47	0.88	4.18	2.84
<i>In vitro</i> organic matter digestibility at 48 hours (%)	1.68	0.89	1.91	0.86	5.16	2.70
Ash (g/kg)	0.83	0.93	1.02	0.90	3.18	3.12
Calcium (g/kg)	0.16	0.74	0.21	0.58	0.32	1.54
Phosphorus (g/kg)	0.03	0.78	0.03	0.66	0.05	1.71
Magnesium (g/kg)	0.14	0.41	0.16	0.23	0.18	1.14
Potassium (g/kg)	0.25	0.87	0.30	0.80	0.68	2.23
Sodium (g/kg)	0.06	0.76	0.07	0.67	0.12	1.73
Sulphur (g/kg)	0.05	0.24	0.06	0.16	0.06	1.08
Chloride (g/kg)	0.07	0.95	0.09	0.93	0.35	3.74

SEC - standard error of calibration;

r<sup>2</sup> - coefficient of determination;

SECV - standard error of cross validation;

1-VR - Cross validation coefficient of determination;

SD – standard deviation;

RDP – ratio of prediction to deviation.

From Table 1 it seems that RPD values calculated for IVOMD24 and IVOMD48 calibrations met the requirements (>2.5) for screening purposes whereas calibration models for NDFD24 and NDFD48 failed. According to Williams (2001) guidelines the RPD values for NDFD24, IVOMD24, and IVOMD48 could be considered as poor and that these NIRS calibration models might only be used for very rough screening.

According to Williams & Norris (2001) it could be expected that statistics assessing the performance of the calibration equations were not as good when spectra were used to predict some nutritional attributes, such as digestibility kinetics. Digestibility is a characteristic, while protein is a chemical entity. Therefore, according to Stuth *et al.* (1999) NIRS prediction of digestibility represents a prediction of a predicted component of feed. According to Andrés *et al.* (2005) parameters measured by biological methods are subjected to higher uncontrolled variability due to multiplicity of sources of experimental and sampling error sources (e.g. basal diet, days, incubation runs, replicates). This can affect the predictive capability of NIR equations that are highly dependent upon the error of the reference method. Thus, the prediction ability of NIRS is expected to decrease with the complexity of parameter such as NDFD.

According to Edney *et al.* (1994) criterion the RPD values calculated for ash and Cl calibrations met the requirements (>3.0) for quality assurance, whereas calibration models for the rest of the minerals could be not recommended for NIRS use (Table 1). The SEC, SECV and r<sup>2</sup> values indicate how well the equations

will perform within the same population (Cozzolino & Moron, 2004). However, with minerals, the SEC and especially  $r^2$  are not good indicators of calibration performance because of the NIRS not directly measuring the element (Stuth *et al.*, 2003; Cozzolino & Moron, 2004).

## Conclusions

The results recorded in the present study indicated that the NIRS technique is acceptable for DM, CP, ADF, NDF, lignin, ash and Cl analysis. These parameters can be used in models for lucerne hay quality grading. Intermediate NIRS calibration models were found for NDF-CP, sugar, IVOMD24 and IVOMD48 that could be considered adequate for screening purposes. Poor calibration results for the remaining parameters, namely SP, ADF-CP, fat, starch, NDFD and the macro minerals (Ca, P, Mg, P, Na and S), with the exception of Cl, indicated that the NIRS technique is less accurate than chemical methods. However, the usefulness of applying NIRS in these cases will largely depend on the accuracy that is required.

## References

- Andres, S., Calleja, A., Lopez, S., Gonzalez, J.S., Rodriguez, P.L. & Giraldez, F.J., 2005. Prediction of gas production kinetic parameters of forages by chemical composition and near infrared reflectance spectroscopy. *Anim. Feed Sci. Technol.* 123, 487-499.
- Barnes, R.J., Dhanoa, M.S. & Lister, S.J., 1989. Standard normal variate transformation and de-trending of near infrared diffuse reflectance spectra. *Applied Spectroscopy* 43, 772-777.
- Cozzolino, D. & Moron, A., 2004. Exploring the use of near infrared reflectance spectroscopy (NIRS) to predict trace minerals in legumes. *Anim Feed Sci. Technol.* 111, 161-173.
- Edney, M.J., Morgan, J.E., Williams, P.C. & Campbell, L.D., 1994. Analysis of feed barley by near infrared reflectance spectroscopy. *J. Near Infrared Spectrosc.* 2, 33-41.
- Herrero, M., Murray, I., Fawcett, R.H. & Dent, J.B., 1996. Prediction of the *in vitro* gas production and chemical composition of kikuyu grass by near-infrared reflectance spectroscopy. *Anim. Feed Sci. Technol.* 60, 51-67.
- Mathison, G.W., Hsu, H., Soofi-Siawash, R., Recinos-Diaz, G., Okine, E.K., Helm, J. & Juskiw, P., 1999. Prediction of composition and ruminal degradability characteristics of barley straw by near infrared reflectance spectroscopy. *Can. J. Anim. Sci.* 79, 519-523.
- NIRS 2, 1995. Routine analysis manual (NIRSystem User's Manual). NIRSystem, Infrasoft International, Port Matilda, USA.
- Norris, K.H., Barnes, R.F., Moore, J.E. & Shenk, J.S., 1976. Predicting forage quality by near infrared reflectance spectroscopy. *J. Anim. Sci.* 43, 889-897.
- Scholtz, G.D.J., 2001. Models for lucerne quality grading. M.Sc. (Agric) dissertation, UFS, Bloemfontein, South Africa.
- Scholtz, G.D.J., Van der Merwe, H.J. & Tylutki, T.P., 2009. The nutritive value of South African *Medicago sativa* L. hay. *S. Afr. J. Anim. Sci.* 39 (WCAP Supplement), 179-182.
- Shenk, J.S. & Westerhaus, M.O., 1993. Analysis of agriculture and food products by near infrared reflectance spectroscopy. Infrasoft International (ISI), Silver Spring, MD, USA.
- Shenk, J.S., Westerhaus, M.O. & Hoover, M.R., 1979. Analysis of forages by infrared reflectance. *J. Dairy Sci.* 62, 807-812.
- Snyman, L.D. & Joubert, H.W., 1992. Prediction of the chemical composition and *in vitro* dry matter digestibility of a number of forages by near infrared reflectance spectroscopy. *S. Afr. J. Anim. Sci.* 23, 20-23.
- Stuth, J.W., Freer, M., Dove, M. & Lyons, H., 1999. Nutritional management for free-ranging livestock. In: *Nutrition of herbivores*. Eds. Jung, H., ASAC. pp. 696-751.
- Van Waes, C., Kirilov, A., Krachounov, I., Petkov, P. & Carlier, L., 1998. Prediction of *in vitro* digestibility and chemical composition of lucerne by NIRS. *Plant Sci. (Sofia)* 35(9), 708-711.
- Weiss, W.P., 1993. Symposium: prevailing concepts in energy utilisation by ruminants: predicting energy values of feeds. *J. Dairy Sci.* 76, 1802-1811.
- Williams, P.C., 1987. Variables affecting near-infrared reflectance spectroscopic analysis. In: *Near-infrared technology in the agricultural and food industries*. Eds Williams, P. & Norris, A., American Association of Cereal Chemists, St. Paul, MN. pp. 143-167.

- Williams, P.C., 2001. Implementation of near infrared technology. In: Near infrared technology in the agricultural and food industries. Eds Williams, P.C. & Norris, K.H., American Association of Cereal Chemist, St. Paul, Minnesota, USA. pp. 145-169.
- Williams, P.C., 2007. Near-infrared technology - Getting the best out of light. Edition 5.1, PDK Projects, Inc., Nanaimo B.C., Canada.
- Williams, P.C. & Cordiero, H.M., 1985. Effect of calibration practice on correction of errors induced in near-infrared protein testing of hard red spring wheat by growing location and season. *J. Agric. Sci., Camb.* 104, 113-123.
- Williams, P.C. & Norris, K.H., 2001. Variables affecting near-infrared spectroscopic analysis. In: Near infrared technology in the agricultural and food industries. Eds. Williams, P.C. & Norris, K.H., American Association of Cereal Chemist, St. Paul, Minnesota, USA. pp. 171-185.