

## In-field hyperspectral proximal sensing for estimating quality parameters of mixed pasture

R. R. Pullanagari · I. J. Yule · M. P. Tuohy · M. J. Hedley ·  
R. A. Dynes · W. M. King

Published online: 13 November 2011  
© Springer Science+Business Media, LLC 2011

**Abstract** A study was conducted to explore the potential use of a hand-held (proximal) hyperspectral sensor equipped with a canopy pasture probe to assess a number of pasture quality parameters: crude protein (CP), acid detergent fibre (ADF), neutral detergent fibre (NDF), ash, dietary cation–anion difference (DCAD), lignin, lipid, metabolisable energy (ME) and organic matter digestibility (OMD) during the autumn season 2009. Partial least squares regression was used to develop a relationship between each of these pasture quality parameters and spectral reflectance acquired in the 500–2 400 nm range. Overall, satisfactory results were produced with high coefficients of determination ( $R^2$ ), Nash–Sutcliffe efficiency (NSE) and ratio prediction to deviation (RPD). High accuracy (low root mean square error-RMSE values) for pasture quality parameters such as CP, ADF, NDF, ash, DCAD, lignin, ME and OMD was achieved; although lipid was poorly predicted. These results suggest that in situ canopy reflectance can be used to predict the pasture quality in a timely fashion so as to assist farmers in their decision making.

**Keywords** Hyperspectral proximal sensing · Pasture quality · Partial least squares regression (PLSR)

---

R. R. Pullanagari · I. J. Yule (✉) · M. P. Tuohy  
New Zealand Centre for Precision Agriculture (NZCPA), Institute of Natural Resources, Massey University, Private Bag 11-222, Palmerston North, New Zealand  
e-mail: I.J.Yule@massey.ac.nz

M. J. Hedley  
Institute of Natural Resources, Massey University, Private Bag 11-222, Palmerston North, New Zealand

R. A. Dynes  
AgResearch, Lincoln Research Centre, Christchurch 8140, New Zealand

W. M. King  
AgResearch, Ruakura Research Centre, Hamilton 3240, New Zealand

## Introduction

Grazed pasture systems support New Zealand's major export earnings from milk, meat and wool production. The meat and milk industries require high quality pastures to maintain productivity and profitability. In addition, pasture quality is a critical factor in determining animal performance, stocking rates and methane emissions (FAO 2010). The key components of pasture quality are CP, fibre, minerals, ash, organic matter digestibility (OMD), sugars and metabolisable energy. Pasture quality is highly variable, and its feeding value depends to a large extent upon the species composition of the pasture, its maturity, stage of growth as well as topography and climatic factors (Holmes et al. 2007).

Intensification of NZ farming systems and consumer demand for changing product specifications contribute to an increasing need for objective measurement and control of pasture quality. The key components of feed quality estimates have typically been measured using conventional methods of wet chemistry according to the Association of Official Analytical Chemists (AOAC 2005). These procedures are time consuming and expensive. The need for faster and cost effective analysis options has led to the wide-spread use of laboratory-based near infrared spectroscopy (NIRS) for estimating foliar chemistry without any chemical treatments and analysis. Lab-NIRS has been widely accepted as a common method to estimate chemical components in materials such as forage (Marten et al. 1983), maize (Volkers et al. 2003), cereals (Stubbs et al. 2009), tuber crop flowers (Lebot et al. 2009), meat (Prieto et al. 2006) and soil (Kusumo et al. 2008). Analysing forage quality using NIRS was initiated by the United States Department of Agriculture (USDA) because it allowed more rapid processing of laboratory samples, multiple analyses with one operation (Marten et al. 1985). However emerging demand is for 'real-time' analysis which overcomes the issues of spatial and temporal variability in pasture quality.

Remote sensing technologies, particularly hyperspectral remote sensing, have enabled field study of vegetative biochemical features at a canopy level and can also record spatial differences (Zarco-Tejada 2000). This reduces the tedious process of intensive sampling and lab analysis. In the early 1990s NASA initiated a programme called Accelerated Canopy Chemistry Programme (ACCP) (NASA 1994). ACCP examined the relationships between spectral data acquired from High Resolution Imaging Spectrometer (HIRIS) and measured foliar chemistry. The nitrogen and lignin concentrations in forest canopy were predicted successfully using this method with  $R^2$  values of 0.87 and 0.77, respectively (Martin and Aber 1997). Despite the successful application of remote sensing in field crops, grassland provides a more diverse set of challenges when adopting these technologies. In general, pastures have greater diversity as spatial and temporal heterogeneity result from a number of confounding factors, including: diverse species, morphology and interactions between the grazing animals, the natural environmental conditions and management practices. To accomplish this task, in such a complex environment, Schellberg et al. (2008) has recommended use of a high resolution spectral sensor, where high spectral and spatial resolution proximal sensors could provide reasonable information with high precision. Recently portable or field spectroradiometers (hyperspectral sensors) have been developed with similar features for research studies in various industries. Research by Sanches (2009), Mutanga (2004) Biewer et al. (2009b) and Pullanagari et al. (2011) found significant relationships between nitrogen concentration and in situ green vegetation. Although there were problems of water interference with biochemical concentrations (Mutanga 2004), soil background and canopy structure, satisfactory results with high accuracy were obtained.

This paper investigates the ability of a proximal hyperspectral sensor to estimate pasture quality parameters (CP, ADF, NDF, ME, ASH, Lignin, Lipid, OMD and DCAD) on

commercial NZ pastures. It can process large numbers of in situ samples cost effectively compared with wet chemistry, creating an opportunity to improve pasture management. Proximal hyperspectral sensors encompass spatial and temporal variations and near real time data is produced to aid effective farm decision processes to be implemented. The objective of this study was to evaluate the spectral differences related to in situ pasture quality, as well as developing and validating relationships between acquired reflectance data and pasture quality parameters.

## Materials and methods

### Study area

The study was conducted on four commercial farms across New Zealand, a total of 320 locations or subplots (Table 1) were sampled. The farms and paddocks within farms varied in: ratio of green: dead and vegetative: non-vegetative plant material, geographical location, botanical combinations of pasture, soil type, climate and livestock enterprise (dairy or sheep and beef). All pastures were based on perennial ryegrass (*Lolium perenne* L.) and white clover (*Trifolium repens* L.). A range of less dominate grasses and a small portion of weeds such as: buttercup (*Ranunculus* spp.), catsear (*Hypochaeris radicata*), chickweed (*Stellaria media*), docks (*Rumex* spp.), Californian thistle (*Cirsium arvense*) and yarrow (*Achillea millefolium*), were also evident.

### Spectral measurements

Canopy spectral measurements were taken during the autumn season of 2009, on 23rd–26th March at Massey University Dairy Farm, Aorangi (sheep and beef), Palmerston North; from 22nd–23rd April at Lincoln University Dairy Farm and from 27th to 28th May at Ruakura Dairy Farm, Hamilton (Table 1). The spectral measurements were acquired in situ using an ASD FieldSpec<sup>®</sup> Pro FR spectroradiometer (Analytical Spectral Devices Inc., Boulder, CO, USA).

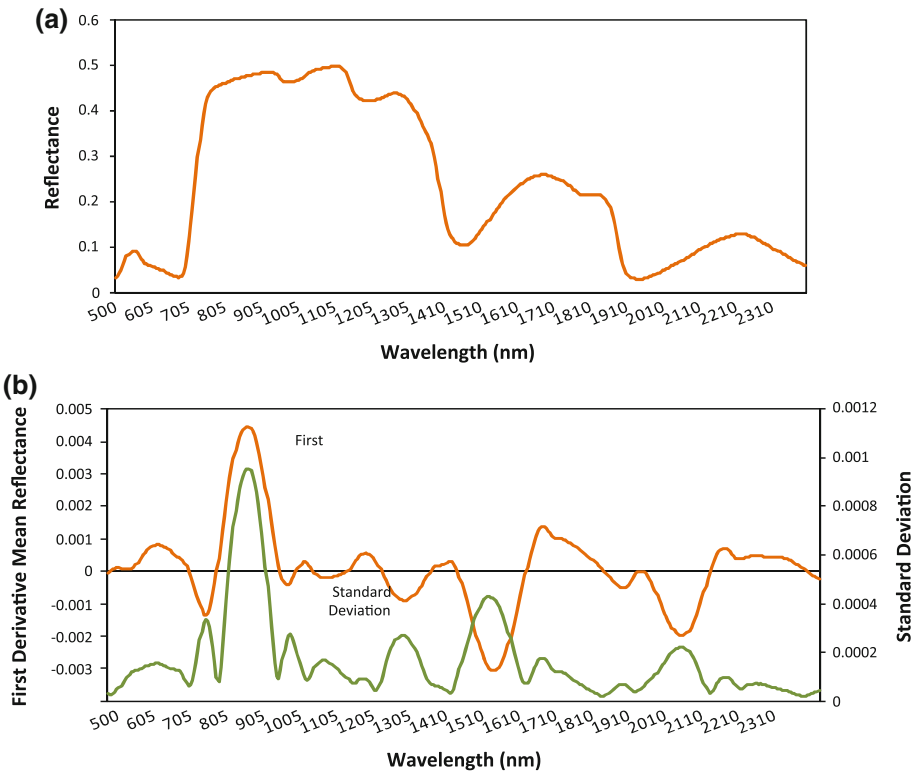
The ground field-of-view was approximately 25° and covered a sample area of 0.25 m<sup>2</sup>. The spectral range was 350–2 500 nm, with 1.4 nm resolutions in the 350–1 000 and 2 nm in the 1 000–2 500 nm. This was re-sampled as 1 nm resolution spectral data (from

**Table 1** The experimental site locations

Site no.	Farm name	Location	Latitude (N)	Longitude (E)	Number of samples	Number of samples considered in analysis
1	Massey University Dairy Farms	Palmerston North	−40.3785	175.6029	80	30
2	Aorangi Research Station, AgResearch	Palmerston North	−40.2523	175.5901	40	18
3	Lincoln University Dairy farm	Lincoln	−43.6399	172.4413	100	92
4	Ruakura Dairy farm, AgResearch	Hamilton	−37.7777	175.3125	100	74

350–2 500 nm) by using ASD, RS<sup>3</sup>™ software. To ensure consistent illumination, the canopy pasture probe (CAPP)-top grip, developed by Sanches (2009), was used. This consisted of an inverted black bin coupled with a 50 Watt tungsten-quartz-halogen bulb as the light source. The CAPP-top grip allowed acquisition of consistent reflectance spectra using an artificial light source under variable natural lighting (e.g. cloudy) and weather (e.g. windy) conditions. At each sub-plot, ten spectral measurements were acquired and subsequently averaged, using View Spec Pro<sup>®</sup> software, to a single reflectance spectrum.

A total of 214 spectral measurements (Table 1) were included in the data analysis. Samples were excluded due to soil contamination the sample submitted for NIR analysis and missing samples. The radiance was converted into reflectance so as to optimise the reflectance by using scans from a reference panel. In this case, a matt white ceramic tile was used as a reference which has been proven as a reasonable and reliable reflectance standard (Sanches et al. 2009). Figure 1 illustrates the population mean reflectance (Fig. 1a), and first derivative mean reflectance of 214 pasture measurements at wavebands ranging from 500–2 400 nm, (Fig. 1b). The first derivative reflectance illustrates most of the variation is in the visible to near infrared with wavelengths of 550–1 000 nm followed by 1 460–1 800 and 2 000–2 300 nm. This indicates the importance of the visible-near infrared region for the study of green vegetation, and is consistent with previous research (Biewer et al. 2009a).



**Fig. 1** a Mean reflectance, b mean and standard deviation of first derivative reflectance of acquired pasture samples ( $n = 214$ )

## Sampling

Subplots in each paddock were selected randomly. At each subplot, a wooden framed quadrant with inside edge dimensions of  $0.5 \times 0.5$  m, was positioned on the pasture so as to obtain spectral signatures. At each site, after spectral measurements, the herbage samples were cut to ground level with an electric shearing hand piece. The dominant species and ancillary data (e.g. weeds, treading damage, and species present) were visually assessed and recorded separately.

## Chemical Analysis

The clipped pasture samples were collected in polythene plastic bags and transported to a laboratory immediately for processing. The samples were oven-dried at  $60^\circ\text{C}$  for 24 h and ground to pass a 1 mm sieve. The CP, acid detergent fibre (ADF), neutral detergent fibre (NDF), ash, DCAD, lignin, lipid, metabolisable energy (ME) and OMD contents were estimated using near-infrared reflectance spectroscopy—NIRS at FeedTECH (Corson et al. 1999) laboratory based at AgResearch in Palmerston North, New Zealand.

## Data processing and statistical analysis

### Data manipulations

The objective of the research was to build relationships between acquired spectral signatures and pasture quality parameters which are important for livestock management.

The spectral data were manipulated to remove spectral abnormalities which occur randomly across the spectrum and to improve absorption features. These abnormalities might be due to internal (detectors and electronic circuits and baseline fluctuations) (Ozaki et al. 2005) and external factors (light leak and humidity). View Spec Pro<sup>®</sup> (ASD Inc.) programme was used to process the raw spectra and eliminated unusual spectra (other than typical green vegetation spectrum) which might reduce the calibration accuracy. In acquired reflectance, spectral data were removed at the two edges of spectra, 350–500 and 2 400–2 500 nm, due to natural light leak into the CAPP. The collected contiguous hyperspectral data (1 900/1 nm wavebands) were reduced to 380/5 nm wavebands using Microsoft Visual Basic<sup>®</sup> software. Data transformation, pre-processing and pre-treatment were followed, (Viscarra Rossel 2008) to enhance spectral properties. Initially, transformation, converting spectral data in R units to  $\log(1/R)$  units, was used to reduce non-linearity. The Savitzky–Golay filter, (Savitzky and Golay 1964) pre-processing smoothing procedure, was then used to improve signal-to-noise ratio. The Savitzky–Golay filter had a window size of 30 and polynomial order of 4. The next step of the pre-processing procedure was to compute the first derivative (Tsai and Philpot 1998) of reflectance differentiation to enhance the absorption spectral features and to minimise background noise. The first derivative (FD) transformation of the reflectance spectrum calculated the slope values from the reflectance which can be derived from the following Eq. 1 (Mutanga et al. 2005):

$$R'_{\lambda(i)} = \frac{(R_{\lambda(j+1)} - R_{\lambda(j)})}{\Delta\lambda} \quad (1)$$

where  $R'$  is the first derivative reflectance at a wavelength  $i$  midpoint between wavebands  $j$  and  $j + 1$ .  $R_{\lambda(j+1)}$  is the reflectance at the  $j + 1$  waveband,  $R_{\lambda(j)}$  is the reflectance at the

$j$  waveband and  $\Delta_j$  is the difference in wavelengths, between  $j$  and  $j + 1$ . As a final step of data manipulation, mean centering of pre-treatment was assigned which may minimise multicollinearity (one variable correlated with other variables) (Aiken and West 1991). It also increased precision and stability of estimates by reducing the standard error and producing least squares of estimates.

### Data Analysis

After data manipulation, multivariate statistics were used by adopting the PARLES software, developed by Viscarra Rossel (2008) to develop relationships between processed spectral data and measured variables of interest. Among multivariate statistics, partial least squares regression (PLSR) is a prominent modelling method which effectively deals with numerous, multicollinear variables and also when the number of explanatory (number of wavelengths) variables is greater than the number of observations (Wold et al. 2001). Before the application of PLSR, principal component analysis was adopted to identify the spectral outliers but in this case none were found.

The calibration model was developed using the PLS technique, then the regression model was developed to predict the unknown quality estimates known as validation or test set. The regression model (2) was represented by (Kawamura et al. 2008):

$$\mathbf{Y} = \beta\mathbf{X} + \varepsilon \quad (2)$$

where  $\mathbf{Y}$  was the dependent variable (pasture parameters),  $\mathbf{X}$  was the independent variable (spectral reflectance),  $\beta$  was the coefficient, and  $\varepsilon$  was the residual. The calibration model was built with the minimum number of components required to minimise the RMSE full cross-validation (leave-one-out method). This predicts each sample, with a PLSR model constructed using the remaining samples ( $n - 1$ ) and is a method of estimating the accuracy of the calibration model internally (Kusumo et al. 2009). The validation errors were then combined into statistical measurements to test the performance of the calibration model. For external validation, the total dataset (214) was divided into 1:1 ratio as calibration (107) and validation (107) sets. The calibration model was used to predict the unknown samples or validation set, thereby estimating the practical accuracy of the developed model. The whole dataset was divided by ranking the samples from smallest to largest. Even and odd number samples were recognised as calibration and validation sets respectively.

### Quantifying Model Accuracy

The accuracy of the calibration and validation models was evaluated by statistical measurements;  $R^2$  (coefficient of determination), RMSE (root mean square error), RMSE% (root mean square error percentage), bias, RPD (ratio prediction to deviation) and Nash–Sutcliffe efficiency (NSE). In general,  $R^2$  indicates the degree of collinearity between predicted and measured data and describes the percentage of variation of the  $\mathbf{X}$  variable in the  $\mathbf{Y}$  variable. Although  $R^2$  has been widely used for model evaluation, this statistic is oversensitive to outliers and insensitive to additive and proportional differences between model predictions and measured data (Legates and Jr McCabe 1999).

The difference of standard deviation between the measured and the predicted values of functional properties of pasture was measured as RMSE (root mean square error). The RMSECV and RMSEP, RMSE measure of cross-validated calibration and validation sets respectively, were calculated according to Eq. 3. The RMSE is an absolute measure of fit

whereas  $R^2$  is a relative measure of fit. Lower values of RMSE indicated a better fit. RMSE is a measure of how accurately the model predicts the response, and is the most important criterion for fit if the main purpose of the model is prediction.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{y} - y)^2}{n}} \quad (3)$$

where  $\hat{y}$  indicates predicted value and  $y$  was the measured laboratory value,  $\bar{y}$  was the mean of measured values and  $n$  was the number of samples. Although, RMSE is more sensitive to outliers and therefore, RMSE% was also calculated using Eq. 4.

$$RMSE\% = 100 \frac{RMSE}{\bar{y}} \quad (4)$$

The bias, mean difference between the reference data and NIRS-predicted data indicated the systematic error in the model and was computed according to Eq. 5.

$$Bias = \frac{1}{n} \sum_{i=1}^n (\hat{y} - y) \quad (5)$$

The RPD is the ratio of the standard deviation of laboratory values of pasture characteristics to the RMSE. This calculation was made to show how much more accurate (measured by the standard error) a prediction from the model was than simply quoting the overall mean (Kusumo et al. 2008). The RPD values were calculated from Eq. 6.

$$RPD = \frac{SD(y)}{\sqrt{\frac{\sum_{i=1}^n (\hat{y} - y)^2}{n}}} \quad (6)$$

In addition, the Nash–Sutcliffe efficiency (NSE) statistic also called the model efficiency was used to examine the relative magnitude of the residues compared to the variance in measured data and was calculated from Eq. 7 (Nash and Sutcliffe 1970). The values of NSE ranged between  $-\alpha$  to 1.0. The acceptable NSE values were from 0 to 1, whereas the negative values ( $<0$ ) were deemed unacceptable which indicated poor model performance (Moriassi et al. 2007) and (Miehle et al. 2006). However, it is strongly sensitive to the variation within the data (Schut et al. 2006).

$$NSE = 1 - \left( \frac{\sum_{i=1}^n (\hat{y} - y)^2}{\sum_{i=1}^n (\hat{y} - \bar{y})^2} \right) \quad (7)$$

Accurate and precise prediction was shown by high  $R^2$ , NSE and RPD, and low RMSE and RMSE %.

After PLSR modelling, the magnitude of each waveband ( $x$ ) in modelling of  $y$  computed using PLS-regression weighted coefficients and represented by variable importance for the projection (VIP) (Wold et al. 2001) and calculated by Eq. 8. A larger score indicates the waveband had greater importance in building a model that predicts  $y$ , while the waveband having a lower score ( $<1$ ) had less importance in developing a model.

$$VIP_k(a) = K \sum_a w^2 a k \left( \frac{SSY_a}{SSY_t} \right) \quad (8)$$

The Eq. 8, where  $VIP_k(a)$  is the importance of the  $k$ th predictor (wavelength) variable based on a model with  $a$  factors (PLS-components),  $W_{ak}$  represents PLS-weights of  $k$ th variable in a  $a$ th PLS-factor,  $SSY_a$  is the explained sum of squares of  $Y$  by a PLSR model with  $a$  factors,  $SSY_t$  is the total sum of squares of  $Y$  explained in all  $a$  factors of a PLS model.

## Results

### Summary statistics of NIRS data

The dataset ( $n = 214$ ) comprised of 107 assays for the calibration set and 107 assays for the validation set. The descriptive statistics of pasture quality estimates were analysed by bench-top laboratory NIRS with respective calibration set and validation sets and are presented in Table 2.

In this experiment a wide range of quality estimates were found. The variation was mainly due to biotic (botanical composition and weeds) and abiotic (slope, soil, altitude and climate) factors (Mutanga and Skidmore 2003) and also samples acquired from different locations and different growth stages. To build a robust calibration model Marten et al. (1985) recommended a wide range of datasets.

To summarise the descriptive statistics illustrated in Table 2, the lignin had the widest variation expressed as a CV(%) with a range of values from 1.30 to 5.14% followed by lipid (0.97–4.24% DM), crude protein (8.08–28.41% DM), NDF (28.29–67.32% DM), DCAD (305.30–874.40 mEq  $\text{kg}^{-1}$  DM), ADF (19.40–38.19% DM), OMD (48.77–92.23% DM) ash (7.30–13.58% DM) and ME (8.39–13.16 MJ  $\text{kg}^{-1}$  DM) in calibration and validation sets. Overall, lignin had the greatest CV of 30.61 and 30.7% while ME had a low CV (10.34 and 10.26% in the calibration and validation datasets respectively).

### Correlation among the pasture quality parameters

A strong linear intercorrelation existed, listed in Table 3, between various measured quality parameters. The majority of the parameters have intercorrelation at a significance level of  $p < 0.001$ . The CP had shown a strong positive correlation ( $R^2$ ), significance at level of  $p < 0.001$ , with ash (0.73), DCAD (0.81), lipid (0.62), ME (0.88), OMD (0.88) while having a strong negative correlation ( $p < 0.001$ ) with ADF (−0.84), NDF (−0.80) and lignin (−0.75). ADF had shown significant correlation with all pasture quality parameters with a range of  $R^2$  (−0.94 to 0.94) values. Similarly, NDF had significant correlation with all quality constituents except ash. Ash had stronger correlation with CP ( $R^2$  0.73) while other components correlated with moderate  $R^2$  values. Lignin and lipid had significant correlations with all quality parameters except with ash.

### Principal component analysis

After mathematical transformations of the reflectance spectra, a principal component analysis (PCA) was conducted to visualise the spectral variance and detect any influence of each object's spectral data within the whole dataset (Esbensen et al. 2009). In the process of PCA decomposition, the available spectral data were transformed into non-linear principal components or latent variables. This described the majority of variation present in



**Table 2** Descriptive statistics of the pasture quality parameters of calibration ( $n = 107$ ) and validation sets ( $n = 107$ ) measured by NIRS

Pasture quality parameters	Calibration set ( $n = 107$ )					Validation set ( $n = 107$ )				
	Mean	Minimum	Maximum	SD	CV (%)	Mean	Minimum	Maximum	SD	CV (%)
	CP	20.49	8.08	28.21	4.88	23.82	20.58	9.38	28.41	4.83
ADF	26.18	19.40	37.31	4.91	18.78	26.27	19.99	38.19	4.96	18.91
NDF	43.95	28.29	67.32	9.16	20.85	44.14	28.66	67.32	9.19	20.84
Ash	10.59	7.35	13.40	1.32	12.49	10.58	7.30	13.58	1.35	12.75
DCAD	632.7	305.30	872.30	123.22	19.48	635.50	311.40	874.40	121.9	19.18
Lignin	2.65	1.30	5.10	0.81	30.61	2.66	1.31	5.14	0.82	30.78
Lipid	2.73	0.97	4.24	0.76	27.98	2.71	0.44	4.18	0.78	28.78
ME	11.24	8.39	12.78	1.16	10.34	11.26	8.51	13.16	1.15	10.26
OMD	77.82	48.77	89.91	10.07	12.94	78.06	53.64	92.23	9.84	12.61

CP crude protein (% DM), ADF acid detergent fibre (% DM), NDF neutral detergent fibre (% DM), DCAD dietary cation-anion difference (mEq kg<sup>-1</sup> DM), ME metabolisable energy (MJ kg<sup>-1</sup> DM), OMD organic matter digestibility (% DM), SD standard deviation, CV coefficient of variation

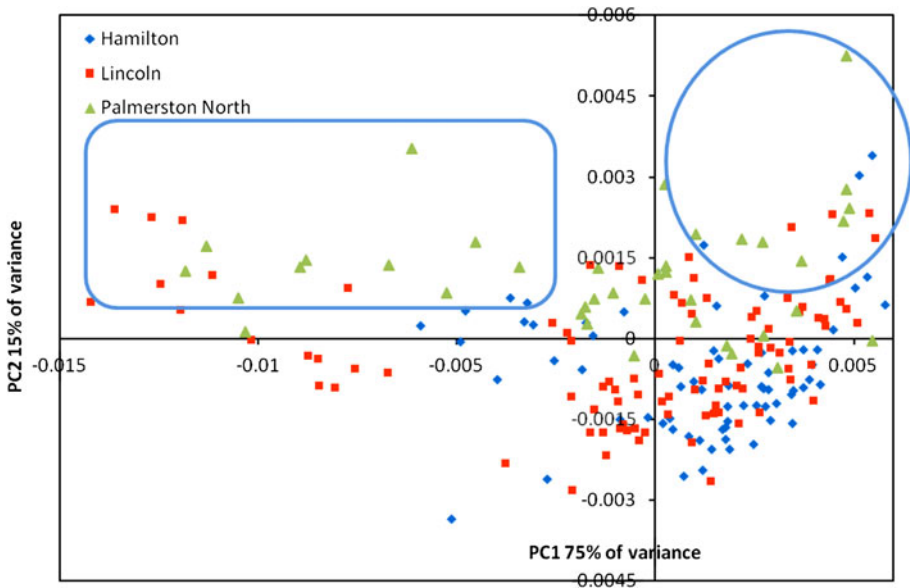
**Table 3** Intercorrelation coefficients of measured pasture quality parameters

Pasture quality Parameters	CP	ADF	NDF	Ash	DCAD	Lignin	Lipid	ME	OMD
CP	1.000								
ADF	-0.838**	1.00							
NDF	-0.804**	0.941**	1.00						
Ash	0.731**	-0.439*	-0.414	1.00					
DCAD	0.809**	-0.767**	-0.776**	0.549**	1.00				
Lignin	-0.749**	0.745**	0.743**	-0.442**	-0.742**	1.00			
Lipid	0.617**	-0.731**	-0.683**	0.396**	0.659**	-0.614**	1.00		
ME	0.877**	-0.912**	-0.855**	0.606**	0.820**	-0.786**	0.667**	1.00	
OMD	0.880**	-0.940**	-0.878**	0.603**	0.830**	-0.778**	0.690**	0.984**	1.00

CP crude protein, ADF acid detergent fibre, NDF neutral detergent fibre, DCAD, dietary cation-anion difference, ME metabolisable energy, OMD organic matter digestibility

\*\* Significant:  $p < 0.001$

\*Significant:  $p < 0.05$



**Fig. 2** Score plot of first and second principal components from the PCA of reflectance spectra

the spectra. Furthermore, the score plot spectral distinctions were useful for identifying outliers and to discriminate the spectral differences as clusters which were used for model development.

The resulting biplot (Fig. 2) illustrated 90% of variance in total spectra: PC1 (principal component 1) accounts for 75% and PC2 (principal component 2) accounts for 15% of the variance. Geographical location had little impact on spectral variation. The location-specific spectral discrimination was not strong, therefore the score values of four quadrants in the score plot (Fig. 2) were evaluated with recorded ancillary data and visual images.

#### PLSR models for calibration and validation datasets

From the calibration and validation datasets the quality estimates of CP, ADF, NDF, ash, DCAD, lignin, ME and OMD were significantly predicted using the spectral data (Table 4). The results of the prediction models and regression equations were described in Table 4. High levels of coefficient of determination ( $R^2$ ) values ranging from 0.71–0.83 were accompanied by low RMSE values, indicating high accuracy. From this, it can be deduced that the performance of the PLSR models were consistent among the calibration and validation datasets with a slight variation of  $R^2$  values (0–5.2%) and RMSE values (0–13%). Although  $R^2$  is a commonly used calibration statistic, it is not the best measure of the merit of a calibration model because it depends on range (Davies and Fearn 2006). To offset this problem, RPD was calculated for the above parameters and found to range in value between 1.88 to 2.46. Under laboratory conditions, the desired level of prediction accuracy is:  $R^2 > 0.8$  and  $RPD > 2$  (Kusumo 2009). For field measurements, however, lower RPD values are acceptable according to Biewer et al. (2009b). The NSE values for the above pasture quality parameters ranged between 0.63 and 0.83, which indicated competent performance of the models.

**Table 4** PLSR results between first derivative reflectance and pasture quality concentrations for cross-validated calibration and validation datasets

Pasture quality parameters	Cross-validated calibration set						Validation set						
	$R^2$	RMSECV	RMSECV%	Bias	RPD	NSE	Regression equations	$R^2$	RMSEP	RMSEP%	Bias	RPD	NSE
CP	0.82	2.08	10.10	0.07	2.34	0.81	$y = 0.849x + 3.18$	0.78	2.33	11.32	-0.12	2.04	0.75
ADF	0.81	2.13	8.163	-0.02	2.28	0.80	$y = 0.841x + 4.12$	0.82	2.23	8.54	0.52	2.11	0.77
NDF	0.77	4.22	9.703	-0.07	2.08	0.76	$y = 0.755x + 10.7$	0.75	4.63	10.56	-0.15	1.95	0.72
Ash	0.65	0.74	6.931	0.00	1.68	0.63	$y = 0.725x + 2.94$	0.65	0.74	6.95	0.01	1.70	0.64
DCAD	0.77	55.92	8.891	0.31	2.10	0.77	$y = 0.814x + 117$	0.73	63.58	9.85	0.38	1.90	0.72
Lignin	0.71	0.41	15.60	0.00	1.87	0.71	$y = 0.739x + 0.695$	0.71	0.42	16.31	0.06	1.84	0.69
Lipid	0.56	0.47	16.82	0.01	1.51	0.55	$y = 0.604x + 1.11$	0.55	0.48	17.08	-0.01	1.44	0.50
ME	0.83	0.46	4.082	0.01	2.46	0.83	$y = 0.859x + 1.59$	0.83	0.46	4.074	-0.05	2.48	0.83
OMD	0.83	4.13	5.30	0.01	2.43	0.83	$y = 0.869x + 10.3$	0.83	4.02	5.14	0.22	2.46	0.83

CP crude protein, ADF acid detergent fibre, NDF neutral detergent fibre, DCAD dietary cation–anion difference, ME metabolisable energy, OMD organic matter digestibility,  $R^2$  coefficient of determination, RMSECV root mean square error of cross validation, RMSECV% root mean square error of cross validation percentage, RMSEP root mean square error of prediction, RMSEP% root mean square error of prediction percentage, RPD ratio prediction to deviation, NSE nash–sutcliffe efficiency

$y$  = measured parameter,  $x$  = predicted parameter using reflectance

In contrast, the remaining quality parameter, lipid was not predicted well by the spectra with results of low  $R^2$  values: 0.52, and 0.18 respectively. Although lipid had a wider range (CV, 24–28%) in the dataset, the precision and accuracy were low.

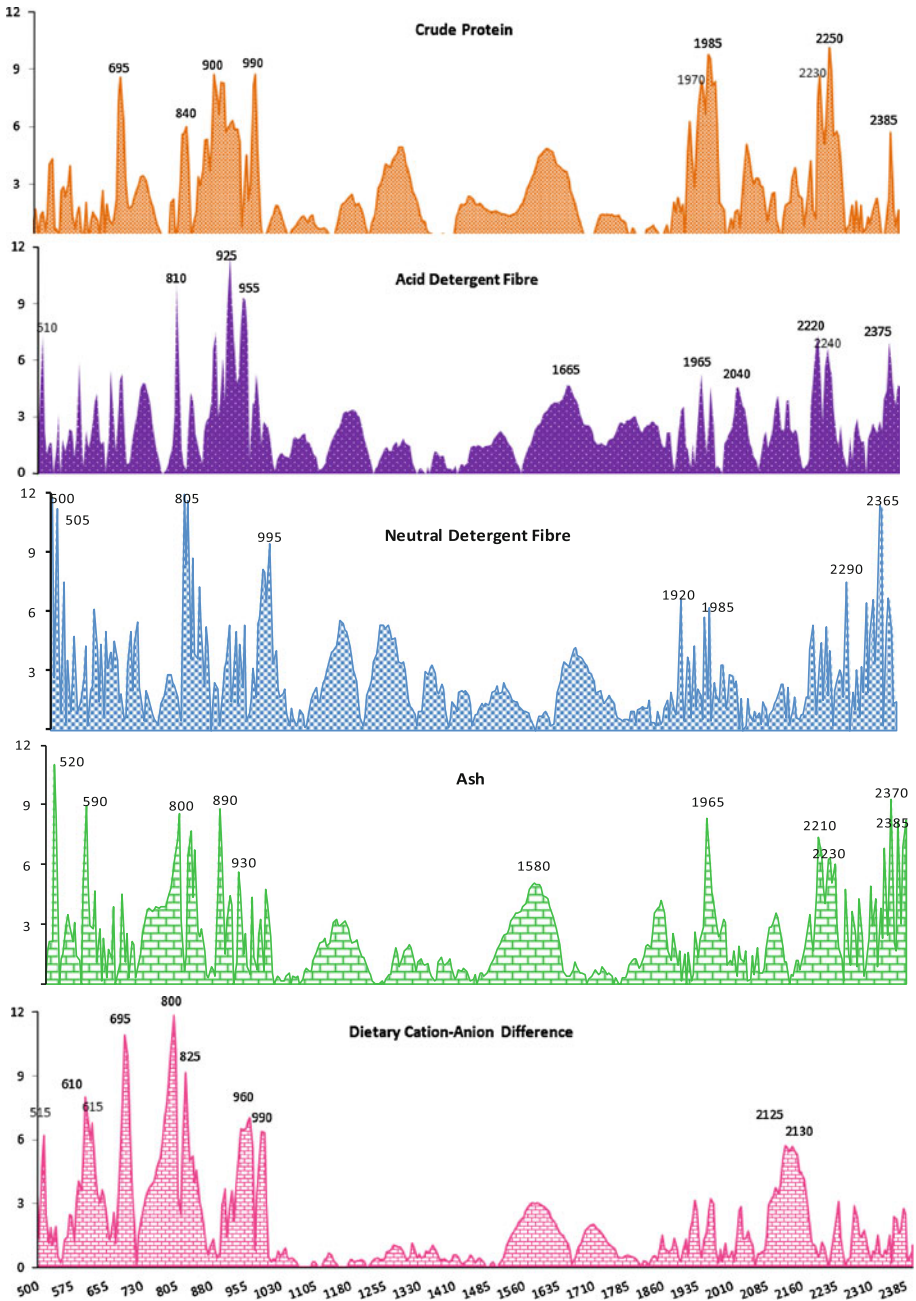
### Important wavebands explaining the variance of pasture quality components

The contribution of each waveband can be visualised by computing the variable importance in projection (VIP), which is illustrated in Fig. 3. As expected, the majority of the important first derivative reflectance wavebands with high VIP scores occurred in the visible region (500–750 nm) (Fig. 3). This was attributed to absorbance of visible radiance by green vegetation. In addition to this, the near infrared region (800–1 000 nm) and shortwave infrared region from 1 900–2 400 nm had shown importance in corresponding to each pasture quality parameter.

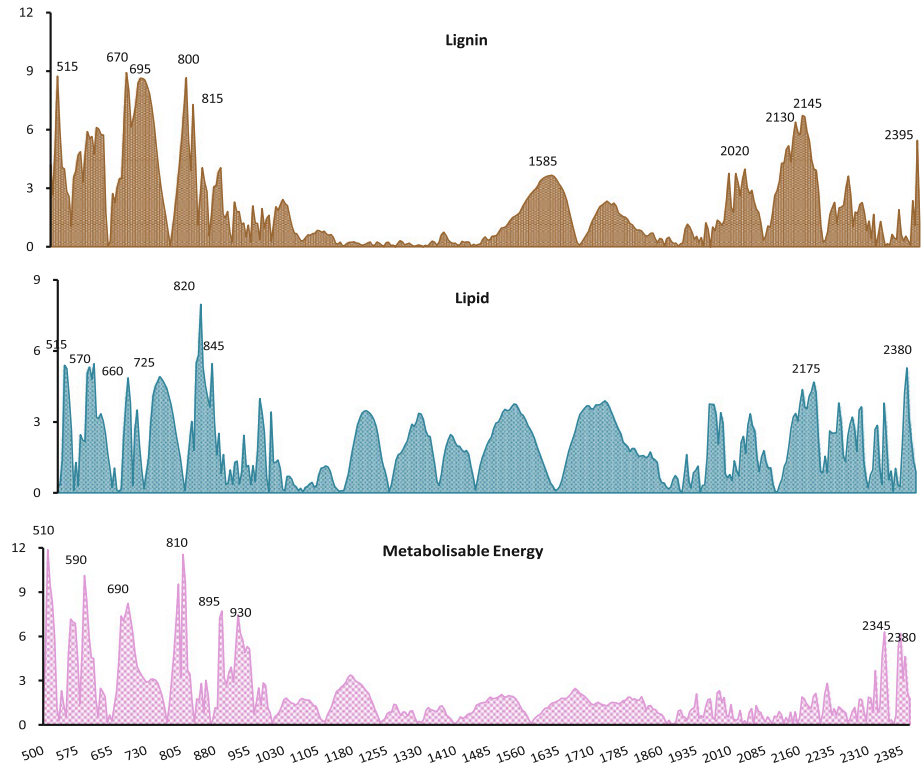
## Discussion

The majority of forage quality prediction experiments using hyperspectral sensors, so far, are confined to reasonably well controlled experimental sites and have produced satisfactory results (Albayrak 2008; Mutanga et al. 2005; Beeri et al. 2007). In order to evaluate the practical application of these sensors and to develop suitable methodologies for the analysis of data, this study was conducted under commercial field conditions. Using single wavebands or broad band indices to explain the variation in foliar chemistry is limited (Zhao et al. 2005). Single band values are not directly related to any plant chemical constituent due to overlapping of chemical absorption features. Therefore, the use of contiguous spectral wavebands, with a full-spectrum approach, was investigated to determine if the relationships between reflectance measurements and in situ pasture quality could be improved. The models were developed using acquired and processed spectral data using 380 wavebands with 5 nm resolution and data produced from lab-NIRS measurement of pasture quality. Although many mathematical transformations are available to develop a “best functional relationship” between measured pasture quality parameters and in situ reflectance measurements, the first derivative (FD) of Log (1/R) was found to be useful in prediction with an improved statistical accuracy compared to reflectance alone. This study has shown that using PLSR analysis resulted in predictive models with high  $R^2$ , RPD and NSE, and lower RMSE and RMSE % values. Biewer et al. (2009b) stressed the importance of using full spectral data using a modified partial least squares regression (MPLSR) algorithm for estimating quality parameters of CP, ME, ash and ADF in highly variable mixed swards with high  $R^2$  and RPD values rather than two-wavebands ratios (low  $R^2$  values). A similar pattern was also observed with the results of (Mutanga et al. 2005). Zhao et al. (2007) also reported the significant performance, with improved  $R^2$  values, of PLSR models used to predict forage quality parameters, compared with simple reflectance ratio and multiple regression (MAXR) with 10-waveband models.

Principal component analysis (PCA) showed the reasons for major spectral discrimination in the score plot may be due to the presence of dead material (highlighted by the rectangle, Fig. 2), variance in botanical composition and pasture colour (highlighted by the circle, Fig. 2). On the top-left corner of the score plot, the samples in the marked region contained some dead material, and the top-right quarter of the score plot samples had light green coloured pasture. These observations were consistent with Sanches (2009). In



**Fig. 3** Variable importance in projection (VIP) plot showing the importance of each waveband in developing a model of pasture quality attributes across the electromagnetic spectrum; X-axis represents wavelength (nm) and Y-axis represents VIP-scores



**Fig. 3** continued

addition, Biewer et al. (2009b) has highlighted the spectral reflectance values obtained from dried grass swards and regarded these as outliers to improve the model accuracy.

Principal component regression (PCR) was applied to investigate the predictive ability of pasture quality parameters, but weaker relationships were obtained, with  $R^2$  values of 0.15–0.45 (data not shown). However, PCA is useful for recognising major sources of variance (fraction of green and dead vegetation) rather than small variances (chemical concentrations) in the vegetation spectral data. This implied that the majority of spectral variance might be influenced by confounding factors such as canopy structure, chemical interactions with other factors, soil background and botanical composition. To offset this problem, PLSR analysis was performed where the spectral PLS-components were more strongly directed towards parameters of interest by providing these parameters with extra weight (Esbensen et al. 2009).

In this study, the pasture quality estimates of CP, ADF, NDF, ash, DCAD, lignin, ME and OMD were predicted with high accuracy, a wide range of chemical constituents of pasture samples caused by natural heterogeneity in permanent pastures (Schellberg et al. 2008) may support improved accuracy. Added to this, fertiliser application rates, varieties and measurement times also contributed to create a large variance (Nguyen et al. 2006). For developing a best fit model a wide range of data within the dataset is essential (Williams and Norris 1987). However, lipid was not predicted well, which might be due to the lower fraction present in the sample. Despite satisfactory results in prediction of biochemical concentrations in green vegetation there is no consistency of statistical

accuracy in various experimental studies (Mutanga et al. 2005; Schut et al. 2006; Kawamura et al. 2008). It could possibly depend on the range of samples used in datasets and the influence of confounding factors (biotic and abiotic). Moreover, there is high level of intercorrelation (Table 3) between quality parameters, which will assist with more precise predictions. However, there are still opportunities to improve the accuracy of models which might be influenced by various interference factors such as: different botanical and floristic composition, weeds (Schut et al. 2006), growth stages, soil background effects (Kokaly 2001) and canopy structure.

Developing calibration equations using the data estimated by laboratory-NIRS has some limitations since there were significant errors associated with prediction and in addition to this, standard error (SE) varies with each chemical compound. For example, the SE of CP of hay was higher than ADF, might be due to an absence of precise methods to analyse detergent fibres rather than CP (Marten et al. 1985). In addition, Biewer et al. (2009b) has explained the relative importance of using wet chemistry values as a reference for reducing the prediction errors as seen in NIRS analysed samples. Considering this statement, this study has used laboratory-NIRS to predict chemical composition of dried samples of each sward. Therefore the accuracy of the measurements might be slightly lower compared to standard procedure (wet chemistry). This suggests that the wet chemistry might improve the model accuracy.

The predictive contribution of each waveband can be visualised by computing the VIP, an output of PARLES (Viscarra Rossel 2008) and shown in Fig. 3. However, as expected, the majority of the important first derivative reflectance wavebands occurred in the visible region (400–750 nm), near infrared region (800–950 nm) and in the shortwave infrared region (1 950–2 350 nm). This can be attributed to absorbance of visible radiance by chlorophyll, which is abundant in green vegetation. Past studies have shown that there was a strong relationship between chlorophyll concentration and nitrogen content in plants due to the presence of N–H bonds (Curran 1989). The leaf organic materials such as: lignin, protein, starch, cellulose, hemicellulose and sugar have common fundamental molecular bonds such as O–H and C–H. The vibrational and bond stretching absorbance's associated with these bonds lie across the spectral region of shortwave infrared from 1.720 to 2.350  $\mu\text{m}$  (Kokaly and Clark 1999). The wavelength at 2.078  $\mu\text{m}$  is responsible for O–H stretch/O–H deformation bond, which are the prominent bonds in starch or sugar (Curran 1989) and water. Absorptions around 1 960, 1 980, 2 100, 2 240 and 2 340 nm are responsible for O–H, N–H, O=H and O–H combinations, C–H (aromatic), C–H and O–H combination bonds respectively (Curran 1989) which are the common bonds in pasture quality parameters.

Protein is the major nitrogen containing biochemical component in plants. For CP, the peaks with higher VIP values surrounded the wavebands from 695–990 nm, and from 1 950–2 400 nm. Absorption in the spectral region from 2 100–2 200 nm has also been attributed to N–H bonds in proteins (Martin and Aber 1997). The absorptions at 2.054 and 2.172  $\mu\text{m}$  is directly related to the presence of C–N and N–H bonds in proteins (Kokaly 2001). This implies that the visible and near infrared wavebands are reliable when estimating biochemical concentrations in pasture.

Accurate and real-time estimation of pasture quality enables farmers to adopt precise management practices. Such practices include fertilizer application which can be applied in response to pasture quality status thereby, fertilizer use has been optimised. Moreover, field or spatial variability maps can be obtained when the sensor integrated with global positioning system (GPS) which allow for the identification chronically poor and high productive areas. These variability maps allow farm managers to maintain pasture evenly



across the field using site-specific practices. Murray and Yule (2007) clearly indicated the economic benefit and increased fertilizer use efficiency by adopting the variable rate application technology as site-specific practice.

Regular monitoring of forage nutrient status provides an opportunity to schedule rotations in an efficient way in order to meet the requirements of stock while maintaining threshold levels in the field, thereby supplements can be provided when there is an inadequate level of nutrient from pasture. Based on the information of available nutrients at the paddock level, stocking rate would be allocated as to meet animal needs. Finally, the successful adoption of precise management practices on grasslands leads to economic and environmental benefits and better utilization of pastures and ensures animal health and performance.

## Conclusion

This paper explains the potential use of in-field hyperspectral proximal sensing to estimate mixed pasture quality using a PLSR algorithm. Satisfactory results were obtained that reflect the strong relationship between spectral measurements and pasture quality parameters. The PLSR models predicted measured attributes with reasonable precision (high  $R^2$ , NSE and RPD values) and accuracy (low RMSE and RMSE % values) compared to other models.

The PLSR algorithm performed better in estimating pasture quality attributes such as CP, ADF, NDF, ash, DCAD, lignin, ME and OMD, while, the estimates of lipid was predicted with lower precision for various reasons. The information produced using in-field hyperspectral proximal sensing of pasture would help pastoral farmers and graziers to improve their productivity, on-farm performance and build business resilience, by enabling them to make more accurate and timely decisions. These are, but not limited to: manipulation of stocking rates, grazing intervals, optimising timing of grazing individual paddocks, benchmarking each paddock within their farm to optimise and tailor capital inputs of fertiliser, plan from which paddocks conserved feed is to be made, and gauge what the quality of the pasture is before harvesting. To extend the results of this study towards a practical outcome for farmers, it is recommended that further research be carried out to investigate the spectral changes in permanent pastures throughout the year and across the seasons with the view of evaluating the need for seasonal calibration of NIRS to pasture quality.

**Acknowledgments** The authors are grateful for the technical support received from: Michael Killick, Massey University, Palmerston North, New Zealand; Grant Rennie, AgResearch, Ruakura, Hamilton, New Zealand; Brian DeVantier, AgResearch Grasslands, Palmerston North, New Zealand; Staff from AgResearch, Lincoln, New Zealand and comments on the manuscript from Carolyn Hedley and Cathie Goulter.

## References

- Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park: Sage Publications, Inc.
- Albayrak, S. (2008). Use of reflectance measurements for the detection of N, P, K, ADF and NDF contents in sainfoin pasture. *Sensors*, 8(11), 7275–7286. doi:10.3390/s8117275.
- AOAC. (2005). *Official methods of analysis of AOAC international* (18th ed., Vol. 1). Gaithersburg: Association of Official Analytical Chemists Inc., AOAC International.

- Beerli, O., Phillips, R., Hendrickson, J., Frank, A., & Kronberg, S. (2007). Estimating forage quantity and quality using aerial hyperspectral imagery for northern mixed-grass prairie. *Remote Sensing of Environment*, *110*(2), 216–225.
- Biewer, S., Erasmii, S., Fricke, T., & Wachendorf, M. (2009a). Prediction of yield and the contribution of legumes in legume–grass mixtures using field spectrometry. *Precision Agriculture*, *10*(2), 128–144. doi:10.1007/s11119-008-9078-9.
- Biewer, S., Fricke, T., & Wachendorf, M. (2009b). Development of canopy reflectance models to predict forage quality of legume–grass mixtures. *Crop Science*, *49*(5), 1917–1926. doi:10.2135/cropsci.2008.11.0653.
- Corson, D. C., Waghorn, G. C., Ulyatt, M. J., & Lee, J. (1999). NIRS: Forage analysis and livestock feeding. *Proceedings of the New Zealand Grassland Association*, *61*, 127–132.
- Curran, P. J. (1989). Remote sensing of foliar chemistry. *Remote Sensing of Environment*, *30*(3), 271–278. doi:10.1016/0034-4257(89)90069-2.
- Davies, A. M. C., & Fearn, T. (2006). Back to basics: Calibration statistics. *Spectroscopy Europe*, *18*(2), 31–32.
- Esbensen, K. H., Guyot, D., Westad, F., & Houmoller, L. P. (2009). *Multivariate data analysis—in practice: An introduction to multivariate data analysis and experimental design* (5th ed.). Oslo, Norway: CAMO.
- FAO. (2010). *Greenhouse gas emissions from the dairy sector: A life cycle assessment*. Rome, Italy: Animal Production and Health Division.
- Holmes, C. W., Wilson, G. F., Mackenzie, D. D. S., Flux, D. S., Brookes, I. M., & Davey, A. W. F. (2007). *Milk production from pasture*. Palmerston North, New Zealand: Massey University.
- Kawamura, K., Watanabe, N., Sakanoue, S., & Inoue, Y. (2008). Estimating forage biomass and quality in a mixed sown pasture based on partial least squares regression with waveband selection. *Grassland Science*, *54*(3), 131–145.
- Kokaly, R. F. (2001). Investigating a physical basis for spectroscopic estimates of leaf nitrogen concentration. *Remote Sensing of Environment*, *75*(2), 153–161. doi:10.1016/S0034-4257(00)00163-2.
- Kokaly, R. F., & Clark, R. N. (1999). Spectroscopic determination of leaf biochemistry using band-depth analysis of absorption features and stepwise multiple linear regression. *Remote Sensing of Environment*, *67*(3), 267–287. doi:10.1016/S0034-4257(98)00084-4.
- Kusumo, B. H. (2009). *Development of field techniques to predict soil carbon, soil nitrogen and root density from soil spectral reflectance*. PhD thesis unpubl, Massey University, Palmerston North.
- Kusumo, B. H., Hedley, M. J., Hedley, C. B., Arnold, G. C., & Tuohy, M. P. (2009). Predicting pasture root density from soil spectral reflectance: Field measurement. *European Journal of Soil Science*, *61*(1), 1–13. doi:10.1111/j.1365-2389.2009.01199.x.
- Kusumo, B. H., Hedley, C., Hedley, M., Hueni, A., Tuohy, M., & Arnold, G. (2008). The use of diffuse reflectance spectroscopy for in situ carbon and nitrogen analysis of pastoral soils. *Australian Journal of Soil Research*, *46*(6–7), 623–635. doi:10.1071/SR08118.
- Lebot, V., Champagne, A., Malapa, R., & Shiley, D. (2009). NIR determination of major constituents in tropical root and tuber crop flours. *Journal of Agricultural and Food Chemistry*, *57*(22), 10539–10547. doi:10.1021/jf902675n.
- Legates, D., & Jr McCabe, G. (1999). Evaluating the use of “goodness-of-fit” measures in hydrologic and hydroclimatic model validation. *Water Resources Research*, *35*(1), 233–241. doi:10.1029/1998WR900018.
- Marten, G. C., Halgerson, J. L., & Cherney, J. H. (1983). Quality prediction of small grain forages by near infrared reflectance spectroscopy. *Crop Science*, *23*(1), 94–96.
- Marten, G. C., Shenk, J. S., & Barton, F. E. I. (1985). Near infrared reflectance spectroscopy (NIRS): Analysis of forage quality. In *Agriculture handbook: No. 643* (p. 96): United States Dept. of Agriculture, Agricultural Research Service.
- Martin, M. E., & Aber, J. D. (1997). High spectral resolution remote sensing of forest canopy lignin, nitrogen, and ecosystem processes. *Ecological Applications*, *7*(2), 431–443.
- Miehle, P., Livesley, S., Li, C., Feikema, P., Adams, M., & Arndt, S. (2006). Quantifying uncertainty from large-scale model predictions of forest carbon dynamics. *Global Change Biology*, *12*(8), 1421–1434. doi:10.1111/j.1365-2486.2006.01176.x.
- Moriassi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, *50*(3), 885–900.
- Murray, R. I., & Yule, I. J. (2007). Developing variable rate application technology: economic impact for farm owners and topdressing operators. *New Zealand Journal of Agricultural Research*, *50*(1), 65–72.
- Mutanga, O. (2004). *Hyperspectral remote sensing of tropical grass quality and quantity*. PhD thesis, Wageningen University, Wageningen, The Netherlands.

- Mutanga, O., & Skidmore, A. (2003). Continuum-removed absorption features estimate tropical savanna grass quality in situ. In *Proceedings of the 3rd EARSeL workshop on imaging spectroscopy* (Vol. 3, pp. 543–558). Herrsching, Germany.
- Mutanga, O., Skidmore, A. K., Kumar, L., & Ferwerda, J. (2005). Estimating tropical pasture quality at canopy level using band depth analysis with continuum removal in the visible domain. *International Journal of Remote Sensing*, 26(6), 1093–1108.
- NASA (1994). *Accelerated canopy chemistry program*. Washington, DC.
- Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I—A discussion of principles. *Journal of Hydrology*, 10(3), 282–290. doi:10.1016/0022-1694(70)90255-6.
- Nguyen, H., Kim, J., Nguyen, A., Nguyen, L., Shin, J., & Lee, B.-W. (2006). Using canopy reflectance and partial least squares regression to calculate within-field statistical variation in crop growth and nitrogen status of rice. *Precision Agriculture*, 7(4), 249–264. doi:10.1007/s11119-006-9010-0.
- Ozaki, Y., McClure, W., & Christy, A. (2005). Spectral analysis. In: Y. Ozaki, W. McClure, & A. Christy (Eds.), *Near infrared spectroscopy in food science and technology*. New Jersey: Wiley-Interscience, Wiley: Hoboken.
- Prieto, N., Andrés, S., Giráldez, F. J., Mantecón, A. R., & Lavín, P. (2006). Potential use of near infrared reflectance spectroscopy (NIRS) for the estimation of chemical composition of oxen meat samples. *Meat Science*, 74(3), 487–496. doi:10.1016/j.meatsci.2006.04.030.
- Pullanagari, R. R., Yule, I., King, W., Dalley, D., & Dynes, R. (2011). The use of optical sensors to estimate pasture quality. *International Journal on Smart Sensing and Intelligent Systems*, 4(1), 125–137.
- Sanches, I. D. (2009). *Hyperspectral proximal sensing of the botanical composition and nutrient content of New Zealand pastures*. PhD thesis unpubl. Massey University, Palmerston North, New Zealand.
- Sanches, I. D., Tuohy, M. P., Hedley, M. J., & Bretherton, M. R. (2009). Large, durable and low-cost reflectance standard for field remote sensing applications. *International Journal of Remote Sensing*, 30(9), 2309–2319. doi:10.1080/01431160802549377.
- Savitzky, A., & Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, 36(8), 1627–1639. doi:10.1021/ac60214a047.
- Schellberg, J., Hill, M. J., Gerhards, R., Rothmund, M., & Braun, M. (2008). Precision agriculture on grassland: Applications, perspectives and constraints. *European Journal of Agronomy*, 29(2–3), 59–71. doi:10.1016/j.eja.2008.05.005.
- Schut, A. G. T., van der Heijden, G. W. A. M., Hoving, I., Stienezen, M. W. J., van Evert, F. K., & Meuleman, J. (2006). Imaging spectroscopy for on-farm measurement of grassland yield and quality. *Agronomy Journal*, 98(5), 1318–1325. doi:10.2134/agronj2005.0225.
- Stubbs, T. L., Kennedy, A. C., & Fortuna, A.-M. (2009). Using NIRS to predict fiber and nutrient content of dryland cereal cultivars. *Journal of Agricultural and Food Chemistry*, 58(1), 398–403. doi:10.1021/jf9025844.
- Tsai, F., & Philpot, W. (1998). Derivative analysis of hyperspectral data. *Remote Sensing of Environment*, 66(1), 41–51. doi:10.1016/S0034-4257(98)00032-7.
- Viscarra Rossel, R. A. (2008). Parles: Software for chemometric analysis of spectroscopic data. *Chemometrics and Intelligent Laboratory Systems*, 90(1), 72–83. doi:10.1016/j.chemolab.2007.06.006.
- Volkers, K. C., Wachendorf, M., Loges, R., Jovanovic, N. J., & Taube, F. (2003). Prediction of the quality of forage maize by near-infrared reflectance spectroscopy. *Animal Feed Science and Technology*, 109(1–4), 183–194. doi:10.1016/s0377-8401(03)00173-1.
- Williams, P., & Norris, K. (1987). *Near-infrared technology in the agricultural and food industries*. American Association of Cereal Chemists, Inc., MN 55121, USA.
- Wold, S., Sjöström, M., & Eriksson, L. (2001). Pls-regression: A basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, 58(2), 109–130. doi:10.1016/S0169-7439(01)00155-1.
- Zarco-Tejada, P. (2000). *Hyperspectral remote sensing of closed forest canopies: Estimation of chlorophyll fluorescence and pigment content*. PhD thesis. York University, Toronto, Ontario, Canada.
- Zhao, D. H., Li, J. L., & Qi, J. G. (2005). Identification of red and NIR spectral regions and vegetative indices for discrimination of cotton nitrogen stress and growth stage. *Computers and Electronics in Agriculture*, 48(2), 155–169.
- Zhao, D., Starks, P. J., Brown, M. A., Phillips, W. A., & Coleman, S. W. (2007). Assessment of forage biomass and quality parameters of bermudagrass using proximal sensing of pasture canopy reflectance. *Grassland Science*, 53(1), 39–49.