finalreport

ANIMAL PRODUCTION

Project code: **NBP.359**

Prepared by: **David Coates**

CSIRO Sustainable Ecosystems

Date published: **April 2009**

ISBN: 9781741914207

PUBLISHED BY
Meat & Livestock Australia Limited
Locked Bag 991
NORTH SYDNEY NSW 2059

Monitoring the performance of FNIRS calibration equations for predicting faecal N concentration and faecal $\delta^{13}C$

Meat & Livestock Australia acknowledges the matching funds provided by the Australian government to support the research and development detailed in this publication.

Abstract

The benefits of F.NIRS technology to industry and research depend on accurate predictions. The performance of F.NIRS calibration equations for predicting faecal N concentration and faecal $\delta^{13}C$ was monitored over 3 years using selected faecal samples covering a wide range of pasture types, geographical locations and reference values. The performance of the faecal N calibration was considered to be excellent for each of four validation sample sets with a combined total of 434 samples. The performance of the faecal $\delta^{13}C$ calibration was variable, being most satisfactory for two of the validation sample sets (total of 125 samples) but much less satisfactory for the other two validation sets. The poorer validation statistics were due primarily to a relatively high proportion of faecal samples from pasture types and/or regions not represented in the calibration and these are detailed in the report. Validation sample spectra (434 for faecal N and 393 for faecal $\delta^{13}C$) with matching reference values were added to the existing calibration sets and new calibration equations with enhanced robustness were developed from the expanded sets. The results of this project indicated that similar work focusing on F.NIRS calibration equations for predicting diet quality deserves high priority.

Executive Summary

A range of faecal NIRS calibration equations applicable to grazing cattle was developed in previous projects (CS.253 and NAP3.121). However, no useful NIRS technology is completed simply with the development of calibration equations that possess satisfactory calibration statistics. Rather, calibration equations need to be validated by monitoring predictive performance over time and, where appropriate, the calibration sample sets need to be expanded to improve robustness (i.e. enhance predictive reliability across the entire target population). Since faecal NIRS is a relatively new technology the need to monitor the performance of existing calibration equations is critical to the acceptance and usefulness of the technology. The lack of funding made it impractical to monitor the performance of calibrations covering the full range of attributes but two attributes, faecal N concentration and faecal $\delta^{13}C$, for which monitoring could be conducted at reasonably low cost, were chosen for test casing.

A total of 434 samples for faecal N and 393 samples for faecal $\delta^{13}C$ were selected for validating existing calibration equations and for expanding the calibration sample sets. Validation and expansion was carried out in a step-wise manner using 4 validation sets of samples selected from commercial cattle properties and from a range of grazing experiments in northern Australia. Samples were selected on the basis of:

(i) pasture type, with special emphasis on the inclusion of pasture types either not represented or poorly represented in the pre-project calibration sets
(ii) geographical location, with special emphasis on properties and locations not represented or poorly represented in the pre-project calibration sets, and
(iii) analyte value (with selection based on predictions using existing calibration equations) so that post-project calibration sets would encompass a desirable distribution of samples within the full range of values encountered in the target population. Special emphasis was given to samples with high faecal N concentrations and faecal $\delta^{13}C$ values indicative of high dietary non-grass proportions since these were under-represented in the pre-project calibration sets

Faecal N reference values were determined by "laboratory wet chemistry" at a number of independent analytical laboratories while faecal $\delta^{13}C$ reference values were measured at the CSIRO Plant Industry laboratory in Canberra using mass spectrometry. Faecal spectra were obtained by scanning the dried, milled faecal samples in a NIR monochromator fitted with a spinning sample cup module. ISI software was used to determine validation statistics and to develop new calibration equations during the stepwise process described above.

Faecal $\delta^{13}C$ validations were less satisfactory than those for faecal N with SEP values of 1.39, 0.95, 0.80 and 1.38 and R^2 values of 0.80, 0.67, 0.91 and 0.81 for the 4 validation sets respectively. Clearly, the poorer validation statistics were associated with 2 of the 4 validation sample sets. In the first of these sets, the less satisfactory predictions (i.e. larger prediction errors) were due primarily to samples that had spectra that were significantly different (Mahalanobis distance > 3) from samples in the calibration set. Larger than expected prediction errors in the second set were confined to a number of specific sites and the prediction errors were due mainly to bias rather than random error. At one site the errors were associated with the presence of a pasture legume (*Clitoria ternatea*) not

previously represented in the calibration set. At another 2 sites prediction bias seemed to be associated with locations where cattle grazed buffel grass pastures on properties in southern Queensland despite good representation of buffel pastures in central and northern Queensland. The results suggested the possibility of a latitudinal influence on the predictions. Notwithstanding the poorer than expected validation statistics for two of the validation sets, the stepwise expansion of the calibration set did not have an adverse effect on the calibration statistics (n = 2052, SEC = 0.766, SECV =0.784, R^2 = 0.93). It can be concluded that the structured addition of samples to the calibration set resulted in improved robustness.

The usefulness of any technology that measures something will logically depend on the accuracy with which the measurements are made. This is critical with F.NIRS particularly in so far as the range in values of attributes being measured or estimated is not large. The work in this project was directed towards ensuring and improving predictive accuracy of faecal N and faecal $\delta^{13}C$. Any improvement over existing equations in predictive accuracy must be of benefit to both research and to industry. Thus the beneficiaries of this work can be identified as managers of beef cattle grazing enterprises, consultants, advisors and extension personnel, research scientists who use F.NIRS as a research tool, and those involved in the continued research and development of F.NIRS.

Contents

1 Background

F.NIRS cattle calibration equations for predicting faecal $\delta^{13}C$ (from which the dietary non-grass proportion is calculated) and faecal N concentration were developed in MRC Project CS.253 (Coates 1998) and expanded as part of MLA Project NAP3.121 (Coates 2004). Although the calibration sets of samples were large and diverse and the calibration statistics were good, confidence in the predictive reliability of NIR calibration equations requires a process of continual validation and expansion of the calibration sets to improve robustness and to ensure that predictions remain as accurate as possible. F.NIRS equations to predict dietary attributes (crude protein and digestibility) were also developed in the course of both the above projects and it would have been desirable to monitor the performance of all calibrations by means of validation experiments. However, validation experiments for monitoring the performance of calibration equations that predict dietary attributes are costly because the reference values have to be measured on samples of the diets and this involves pen feeding. On the other hand, monitoring the performance of the calibration equations for predicting faecal $\delta^{13}C$ and faecal N are not costly because both the reference values and the NIR spectra are measured on the faeces. Funding was only available for these less costly validations.

Although measures of dietary non-grass and faecal N concentration do not provide information as critical as that provided by dietary CP and digestibility, they are both useful in association with the predictions of diet quality in establishing a comprehensive picture of the botanical and chemical composition of the diet being selected by free grazing cattle. Such information is useful for both research and commercial application.

Faecal NIRS analysis was commercialised in July 2006 and calibration equations were provided to Symbio Alliance for the purpose of providing the commercial analytical service to industry. The current project was initiated to help ensure that predictive reliability is maintained for the two attributes under consideration.

2 Project Objectives

(i) Conduct validation tests to determine the predictive reliability of F.NIRS calibration equations for predicting faecal $\delta^{13}C$ and faecal N concentration in cattle.

(ii) Improve the robustness of the existing calibration equations by expanding the calibration sets with the validation samples and recalibrating.

3 Methodology

Over the course of three years, July 2005 to June 2008, validation sets were established by selecting approximately 400 validation faecal samples for each analyte (i.e. for $\delta^{13}C$ and N concentration) from those submitted to the CSIRO Davies Laboratory for NIRS analysis. During 2005-06 samples were received from commercial properties as well as from grazing experiments. For the two years 2006-07 and 2007-08, following the commercialisation of F.NIRS, incoming samples were limited to those from grazing experiments. The selection of samples was aimed

primarily at (i) properties and/or pasture types not already represented or not well represented in the pre-project calibration sets, (ii) a wide range in the respective analytes, and (iii) samples that were identified by the ISI software as being spectrally different from samples in the existing calibration sets.

F.NIRS predictions of faecal $\delta^{13}C$ and/or faecal N were made in the usual way using the existing, pre-project calibration equations designated FECN(a).EQA and DELF(a).EQA in this report. The selected samples were then submitted to external laboratories for chemical analysis to obtain reference values for faecal $\delta^{13}C$ and/or faecal N.

Validation statistics were determined by relating predicted values to laboratory reference values using ISI software. In addition the data were assessed to determine whether less accurate predictions were associated with particular pasture types or locations. Validation tests (A, B, C and D) were made with sequential batches of samples over the 3-year period. All samples in validation set C were a subset of samples collected from Pigeon Hole Station in the VRD while samples in sets A, B and D were sourced from a range of locations.

Following validation the samples were added to the existing calibration sets and new calibrations were developed on the expanded sets to improve predictive robustness, (FECN(b).EQA, FECN(c).EQA , FECN(d).EQA and FECN(e).EQA for faecal N and DELF(b).EQA, DELF(c).EQA, DELF(d).EQA and DELF(e).EQA for faecal $\delta^{13}C$.

4 Results and Discussion

4.1 Faecal N validation

Four sequential validations tests (A, B, C and D) were carried out during the course of the project involving 157 samples, 58 samples, 67 samples, and 152 samples in tests A, B, C and D respectively.

The results of validation tests A and B were described in Milestone Reports 1 and 2 and are presented again in Table 1 along with the results from validation tests C and D.

Table 1. Validation statistics for each of the validation sets.

Validation	No. samples	SEP	Bias	Slope	R^2	RPD_{val}
		%N	%N			
A	157	0.18	0.09	1.04	0.92	2.16
B	58	0.14	-0.03	1.16	0.96	2.95
C	67	0.10	0.09	1.02	0.96	4.14
D	152	0.13	-0.07	0.97	0.93	3.15

While the validation statistics were satisfactory for all four tests it can be seen that there was a trend for some improvement as the calibration set was expanded and the calibration equation became more robust.

Calibration statistics for the equation existing prior to the validation exercise and for the equations developed with the addition of each validation set are shown in Table 2.

Table 2. Calibration statistics for equations developed before and after the addition of new samples to the calibration file from the validation sets.

Calibration	Samples in calibration set[1]	Outliers eliminated	SEC %N	SECV %N	R^2	RPD_{cal}
FECN(a).EQA	1012	27	0.075	0.077	0.963	5.05
FECN(b).EQA	1169	35	0.082	0.083	0.961	4.98
FECN(c).EQA	1227	42	0.081	0.082	0.962	5.05
FECN(d).EQA	1294	37	0.082	0.083	0.960	4.99
FECN(e).EQA	1446	43	0.080	0.084	0.961	4.86

[1] includes outliers

There was very little change in the calibration equation statistics during the step-wise expansion of the calibration set with an overall increase in sample numbers of approximately 43%. The substantial difference between the validation SEP values on the one hand (Table 1) and the calibration SEC and SECV values on the other hand (Table 2) requires some comment. The high SEP values were caused by relatively few samples that had large prediction errors and these samples had little influence on the subsequent calibration statistics because the offending samples were mostly eliminated as outliers during calibration. Outliers (defined according to the magnitude of the Mahalanobis distance or h-value, and by the t-statistic calculated from the relationship between reference and predicted values) are identified by the ISI software and eliminated during the calibration procedure. If outliers were not eliminated during calibration the statistics would be poorer. Because no outliers are eliminated in validation tests it is logical to expect validation statistics to be somewhat poorer than calibration statistics.

Of the 43 samples that were eliminated as outliers during calibration of FECN(e).EQA, 22 or 51% came from samples in the validation sets (14 from set A, 3 from set B, and 5 from set D. This compared with only 21 outliers coming from the 1012 samples originally present in the calibration set. No outliers were eliminated from validation set C and, in that case, the validation SEP was more closely aligned with calibration SECV values. The outlier percentages for the different categories were:

original calibration set (n = 1012):	2.08%
validation set A (n = 157):	8.92%
validation set B (n = 58):	5.17%
validation set C (n = 67):	nil
validation set D (n = 152):	3.29%
final calibration set (n = 1446):	2.97%

The high outlier rate for the validation sets was unexpected, especially for validation set A. However, an investigation of the outliers revealed that most were at the high end of faecal N concentration: only 5 outliers in the final calibration set had reference faecal N under 1.5% and 26 of 43 (60%) had reference faecal N over 2%. Moreover, predictions on 75% of the outliers were under-estimates and all predictions were under-estimates on the 22 outliers with reference faecal N over 2.3%. It appears therefore that large prediction errors are more of a problem at high faecal N concentrations and that these large errors are due mainly to predictions being under-estimates. The high outlier rate for samples in validation set A was probably associated with a higher than normal proportion of faecal samples with high faecal N concentration (39 samples with N% over 2 and one third of these were outliers). Although differences between reference and predicted values are designated prediction

errors, it should be noted that faulty reference values can, and do, contribute to these errors. Unfortunately it is difficult, often times impossible, to determine the contribution of laboratory error to the so-called "prediction error" of NIRS values. During the course of building up the calibration set, 5 different chemical laboratories and at least 3 different methods were used to provide reference values for faecal N and this unavoidable situation probably contributed to increased laboratory error.

4.2 Faecal $\delta^{13}C$ validation

Validation sets for predicting faecal $\delta^{13}C$ again included 4 sets with 118, 58, 67 and 150 faecal samples for sets A, B, C and D respectively.

The results of faecal $\delta^{13}C$ validation tests A and B were described in Milestone Reports 1 and 2 and are presented again in Table 3 along with the results from validation tests C and D.

Table 3. Validation statistics for prediction of faecal $\delta^{13}C$

Validation	No. samples	SEP	Bias	Slope	R^2	RPD_{val}
		‰	‰			
A	118	1.39	0.08	0.87	0.80	2.22
B	58	0.95	-0.04	1.01	0.67	3.25
C	67	0.80	0.60	0.87	0.91	3.86
D	150	1.38	-0.90	0.89	0.81	2.18

Validation statistics for faecal $\delta^{13}C$ improved during the first 3 validations but deteriorated substantially in test D. This deterioration was unexpected but a close analysis of the results was enlightening (see below).

Calibration statistics for the equation existing prior to the validation exercise and for the equations developed with the addition of each validation set are shown in Table 4. There was little change in calibration statistics as a result of the step-wise expansion of the calibration.

Table 4. Faecal $\delta^{13}C$ calibration statistics for equations developed before and after the addition of new samples to the calibration file from validation sets A, B, C and D.

Calibration	Samples in calibration set[1]	Outliers eliminated	SEC ‰ $\delta^{13}C$	SECV ‰ $\delta^{13}C$	R^2	RPD_{cal}
DELF(a).EQA	1659	48	0.769	0.788	0.938	3.92
DELF(b).EQA	1777	50	0.793	0.822	0.934	3.75
DELF(c).EQA	1835	49	0.837	0.849	0.925	3.60
DELF(d).EQA	1902	54	0.754	0.777	0.937	3.87
DELF(e).EQA	2052	59	0.766	0.784	0.933	3.77

[1] includes outliers

The substantially higher SEP values for validation sets A and D compared with calibration SEC and SECV values were similar to those for faecal N and can be explained to some extent by the occurrence of outliers in the validation sets.

In validation set A (n = 118) there was a high proportion of samples that differed spectrally from samples in the existing calibration set. There were 11 samples that were single starred (h > 3 but < 4) and 23 samples that were double starred (h >4) and the SEP for those 34 samples was 2.02‰.

This compared with an SEP for the remainder of the samples of only 0.96‰. In validation set B there were no spectral outliers and only one in set C. The absence of spectral outliers was reflected in SEP and RPD_{val} values more consistent with the SEC and RPD_{cal} values of the calibrations. The low R^2 value for validation set B was directly associated with the limited range of reference faecal $\delta^{13}C$ values. Validation set D was somewhat more complex (Table 5 and Figs. 1 and 2). In total there were only 10 samples that were spectral outliers and though the mean error (difference between reference and predicted value) for these 10 samples at 1.52‰ was higher than the mean error of the remaining 140 samples at 1.05‰, there were many samples that were not spectral outliers that contributed to the high SEP and low RPD_{val} statistics. A close inspection of the data indicated that the poor validation statistics were associated with samples from specific locations and/or pasture species (Table 5).

Validation SEP and RPD_{val} statistics for the first 3 entries in Table 5 (also see Fig. 1) were excellent, those for the next 2 entries were satisfactory, while those for the other 5 entries (Fig. 2) were poor. The following points are noteworthy:

(a) The sites and pasture systems where validations statistics were excellent or satisfactory were well represented in the calibration set from which the prediction equation DELF(d).EQA was developed.

(b) There was a deterioration in validation statistics of buffel grass sites from north (Fletcherview) to south (Ticehurst, Sunnyholt, and Banyula –S-E Q buffel sites in Fig. 2). At the southern sites there was a substantial negative bias in faecal $\delta^{13}C$ predictions (i.e. over-prediction of % non-grass), particularly at Sunnyholt and Ticehurst (Fig. 2). The southern sites were not represented in equation DELF(d).EQA. The large prediction bias of -1.91‰ may have been due to the presence of some C_3 grasses in the southern pasture systems or it could be the result of climatically induced changes (particularly temperature) in the chemical composition of buffel grass growing at higher latitudes compared with C_4 grasses in the tropics. There was also a substantial negative bias (-1.20‰) for samples from the CQ buffel grass sites and this was somewhat unexpected because samples from properties in CQ were well represented in DELF(d).EQA.

(c) Predictions on the samples from Brian Pastures showed a mixed result with those for cattle grazing stylo/grass pasture being in accord with expectations (mean error of only 0.20‰ and minimal bias whereas predictions for samples from butterfly pea/grass pasture were subject to large errors (mean 1.51‰), an overall bias (-1.16‰) resulting in over-estimation of non-grass, and a low R^2 value (0.41). These mixed results can be explained in part by samples from cattle grazing butterfly pea pastures not being represented at all in the calibration set compared with samples from stylo pastures that were well represented. However, it should be noted that non-representation in the calibration set does not mean than predictions will necessarily be inaccurate for a given species or location.

(d) Predictions on 4 of the 11 samples from Pigeon Hole were poor. This result was unexpected considering that all 67 samples in validation set C were from Pigeon Hole and the validation statistics on that sample set were excellent. However, all samples in validation set C were collected on the same day while Pigeon Hole samples in validation set D were from a different year.

(e) Prediction errors for Douglas Daly samples in validation set D were large with an overall bias of -2.11‰. This was also unexpected since samples from Douglas Daly were well represented (n = 45) in DELF(d).EQA.

Table 5. Validation statistics for validation set D when categorised for different sites or pasture systems. Predictions were made using calibration DELF(d).EQA and the sites/pasture systems are ranked in ascending order of SEP (and descending order of RPD_{val}).

Site or Pasture system	No. samples	SEP ‰ $\delta^{13}C$	Bias ‰ $\delta^{13}C$	Slope	R^2	RPD_{val}
Salisbury Plains[1]	9	0.46	-0.16	0.84	0.96	6.53
Kidman Springs etc[2]	12	0.61	-0.16	1.11	0.94	4.92
Fletcherview [3]	17	0.65	0.19	0.72	0.71	4.62
Wambiana[4]	24	1.03	-0.59	0.88	0.93	2.92
Belmont etc [5]	11	1.09	-0.95	0.91	0.89	2.76
Brian Past Exp.[6]	20	1.41	-0.78	0.81	0.55	2.17
CQ buffel sites [7]	14	1.49	-1.20	(0.34)	(0.22)	2.02
Pigeon Hole[8]	11	1.95	-1.75	(0.36)	(0.35)	1.54
Southern Buffel [9]	25	2.14	-1.91	1.07	0.88	1.40
Douglas Daly[10]	5	2.64	-2.11	2.87	(0.93)	1.14

[1] Located north of Bowen with native and introduced grasses, often in combination with stylos.
[2] Kidman Springs, Auvergne Station, Manbulloo: located in the Northern Territory west of Katherine; native grass pastures.
[3] Located on basalt soils north of Charters Towers; buffel grass pasture.
[4] Located south of Charters Towers; native pasture in the northern part of the Bothriochloa/Aristida region.
[5] Belmont Research Station (near Rockhampton), Somerville (near Richmond) , Rocky Springs near Mundubbera, Coorabulka near Mt Isa.
[6] Located in the southern speargrass region near Gayndah; pastures of introduced grasses with stylo (*Stylosanthes seabrana* cv. Caatinga) or butterfly pea (*Clitoria ternatea*).
[7] Frankfield north of Clermont, Brigalow Research Station near Theodore, Berrigurra near Comet, and Melrose near Mornish.
[8] Located in the Victoria River District of the NT; native pasture.
[9] Sunnyholt near Injune, Ticehurst near Surat, Banyula near Condamine
[10] South of Darwin in the wet monsoonal area of the NT; sown grass and grass/legume pastures.

While the results for validation test D were somewhat disappointing they did highlight the necessity of continuous monitoring of NIRS calibration equations, especially where the target population is very large and diverse and where there is large between-year variability in pasture conditions. Obviously calibration equation DELF(d).EQA provided accurate predictions for some locations and pasture systems. However, predictions for some other locations and/or pasture systems were less satisfactory and these were mostly associated with poor representation in the calibration set. The Pigeon Hole and Douglas Daly sites were exceptions in that predictive accuracy was poor despite what would normally be classed as adequate representation in the calibration set. There are many factors that can contribute to poor predictive accuracy including sample collection and processing protocols as well as technical factors associated with the target analyte and instrument function, and last, but not least, there is always the possibility of human error leading to faulty reference values.

Instrument function was not likely to have been a problem in this instance because predictions from all sites would have been affected.

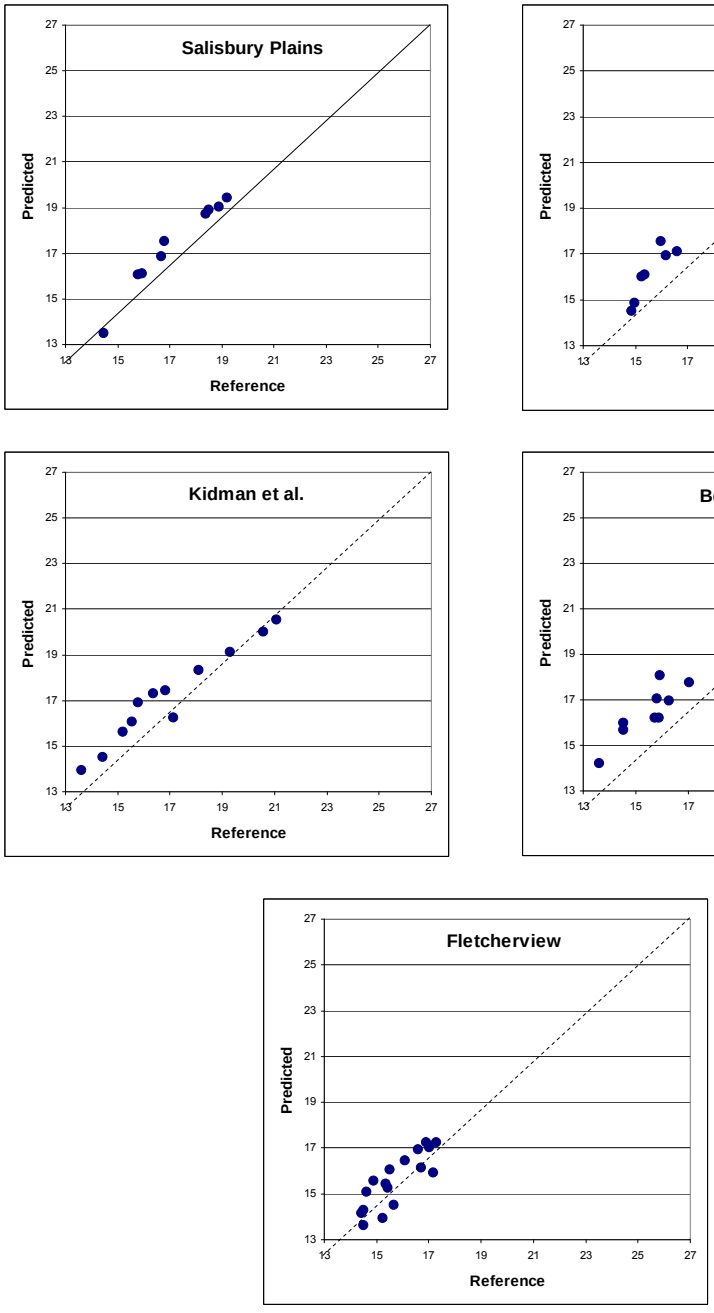

Figure 1. DELF(d).EQA predictions vs reference delta $\delta^{13}C$ for 5 of the sites/pasture systems in validation set D. Predictions at these sites were satisfactory.

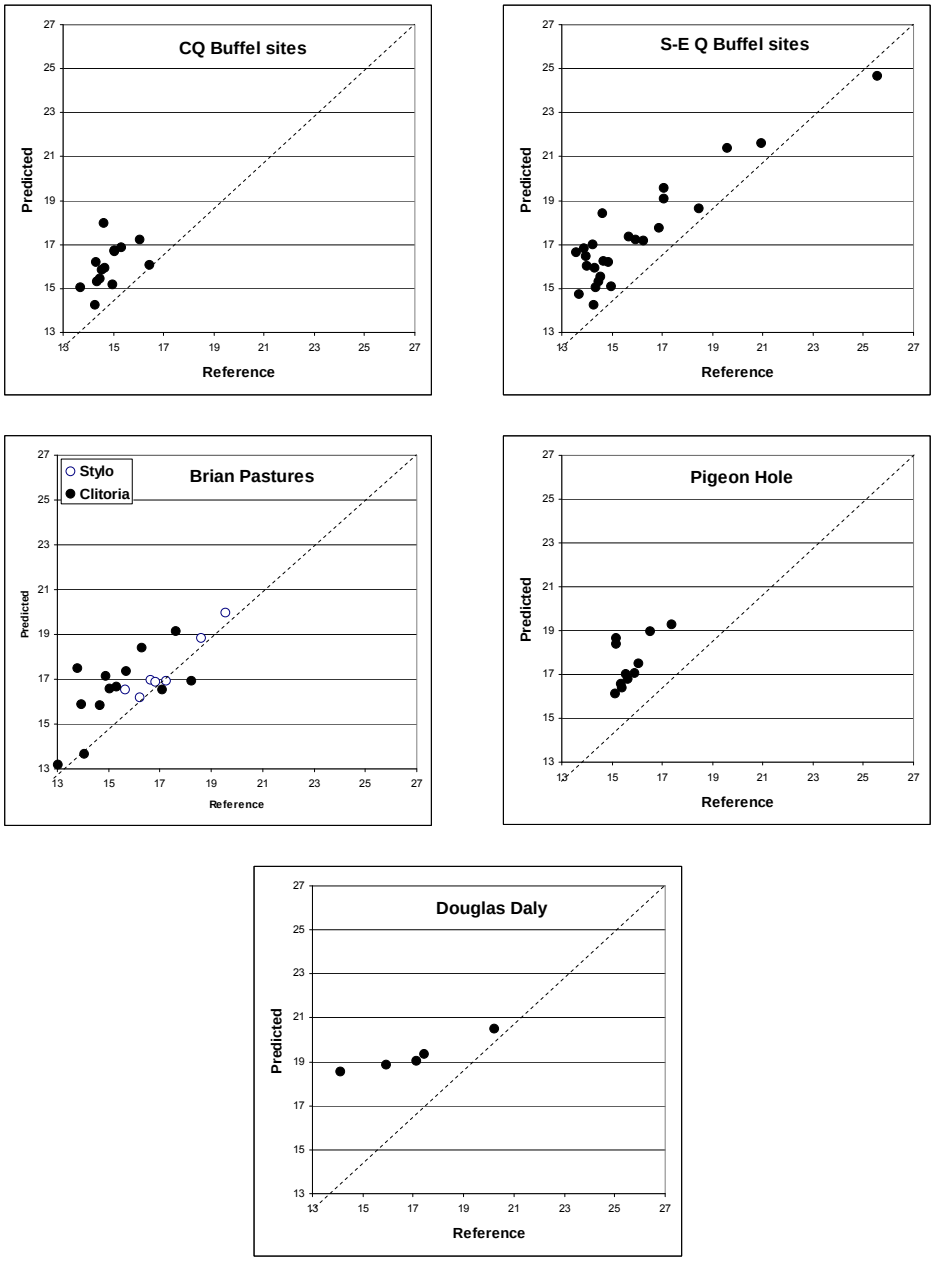

Figure 2. DELF(d).EQA predictions vs reference faecal $\delta^{13}C$ for sites/pasture systems in validation set D where predictions were unsatisfactory.

The presence of C_3 grass or C_4 non-grass in the diet will lead to errors in the prediction of faecal $\delta^{13}C$, the former resulting in a negative bias and the latter to a positive bias. It should be noted, however, that prediction errors in non-grass proportions associated with either C_3 grass or C_4 non-grass in the diet will be substantially less than the associated faecal $\delta^{13}C$ prediction errors (Coates and Dixon 2008). The negative bias observed at the southern buffel grass sites may possibly have been due in part to C_3 grass but this would be most unlikely at the more northern sites. The trend seen in the buffel grass sites proceeding from the northern most site (Fletcherview with positive bias of 0.19‰) to central Queensland sites (negative bias of -1.20‰) and to southern Queensland sites (negative bias of -1.91‰) suggests the possibility of a climatic influence on predictions via an effect on plant fibre composition and this needs further investigation.

Validation sample set D for faecal $\delta^{13}C$ was basically the same as the sample set for faecal N and although the faecal $\delta^{13}C$ validation statistics for samples from butterfly pea pastures at Brian pastures, from CQ and S-E Q buffel pastures, from Pigeon Hole and from Douglas Daley were less than satisfactory, faecal N validation statistics from these same sites were most satisfactory except for the S-E Q buffel pastures where the R^2 was 0.84 (Table 6). The lower correlation between predicted and faecal N reference values at the southern buffel sites was difficult to understand in the light of the high R^2 values at all the other sites. Additionally, the fact that there was a negative bias at all sites (Table 6) suggests the possibility of a bias in the reference values for validation set D.

Table 6. FECN(d).EQA predictions vs reference faecal N concentrations for samples from validation set D according to site/pasture systems.

Site or Pasture system	SEP	R^2	Slope	Bias (%N)
Salisbury Plains	0.07	0.99	1.06	-0.04
Kidman Springs etc	0.11	0.94	0.95	-0.08
Fletcherview	0.07	0.95	1.03	-0.04
Wambiana	0.16	0.97	1.08	-0.14
Belmont etc	0.09	0.95	0.98	-0.07
Brian Past	0.09	0.96	0.96	-0.10
CQ buffel sites	0.09	0.95	1.07	-0.03
Pigeon Hole	0.20	0.94	0.84	-0.15
S-E Q Buffel sites	0.15	0.84	0.93	-0.01
Douglas Daly	0.13	0.94	0.80	-0.12

Overall, it can be reasonably concluded that the pre-project faecal N calibration equation, FECN(a).EQA was probably already quite robust and there is no reason to doubt the robustness and predictive accuracy of the post-project calibration, FECN(e).EQA. Predictive accuracy is likely to be less at high faecal N concentrations but this is of little concern in relation to using faecal N as a nutritional diagnostic.

Project results indicated that the pre-project faecal $\delta^{13}C$ calibration was not as robust as the faecal N calibration with regard to application across all of northern Australia. This is quite understandable due to the nature of the attribute being predicted. In the case of faecal N the attribute is a specific, identifiable chemical component while, in the case of faecal $\delta^{13}C$, the estimate is an index of non-grass proportions (Coates and Dixon 2008) based on the spectral influence of an unknown number of unidentified chemical components. Some weaknesses regarding the robustness of the pre-project

faecal δ^{13}C calibration were exposed in the validation tests and the post-project calibration equation should have improved robustness. Nevertheless, it is to be expected that some weaknesses will remain regarding predictive accuracy in certain circumstances, circumstances relating to plant species, and the interaction of geographic location with seasonal conditions and/or climatic influences. This highlights the need for continuous monitoring of NIRS calibrations and that certain calibrations need more rigorous monitoring than others to ensure a satisfactory level of predictive accuracy. Notwithstanding the problems encountered with the faecal δ^{13}C calibration, the post-project equation should provide dietary non-grass predictions of sufficient accuracy (calculated from predicted faecal δ^{13}C) to be useful in most situations in northern Australia.

5 Success in Achieving Objectives

The objectives of the project were fulfilled. The number of validation samples exceeded the agreed number (300) by 45% and 30% for faecal N and faecal δ13C respectively. The addition of the validation samples to the pre-project calibration files represented 43% and 24% expansions to the faecal N and faecal δ13C calibration sets respectively.

The validation statistics for predicting faecal N concentration indicated that the existing, pre-project calibration equation was accurate and robust. The post-project calibration equation developed on the expanded calibration file will have enhanced robustness. The validation statistics for predicting faecal δ13C indicated lack of robustness leading to prediction biases in relation to specific pasture types and locations. While the expanded, post-project calibration equation will have enhanced robustness and therefore greater predictive reliability, the project was successful in highlighting situations where reduced predictive accuracy may be expected. The project highlighted the critical need for the monitoring and expansion of F.NIRS calibration equations.

6 Impact on Meat and Livestock Industry – now & in five years time

This project was all about ensuring and/or improving the accuracy of F.NIRS predictions. Any incremental improvement in accuracy will have benefits if the attribute being measured or estimated was worth measuring in the first instance. The direct impact of this project on the meat and livestock industry, now and in five years time, will be small. The indirect impact is likely to be more substantial to the extent that the results provide (i) a basis for confidence in the robustness of part of existing F.NIRS technology, *viz*. in the accuracy of faecal N predictions, (ii) a greater understanding of potential areas of weakness in F.NIRS technology, in this case in relation to the prediction of faecal δ^{13}C, and (iii) a clear signal of the critical importance of monitoring the performance of F.NIRS calibration equations and of expanding calibration files to improve robustness and predictive accuracy. Thus, this project can be viewed as a test case giving guidance to future needs. The main outcome is the message that faecal NIRS technology in Australia is a work in progress. Existing calibration equations, especially those for estimating diet quality, need to be monitored and expanded. Currently they should be considered as “immature” and lacking in robustness. The ultimate usefulness of F.NIRS technology, whether for commercial or research purposes, will depend on the priority given to validation and expansion.

7 Conclusions and Recommendations

7.1 Conclusions

(a) The current (post-project) expanded calibration equation for predicting faecal N is both accurate and robust. The need for future validation and expansion can be regarded as minimal. Prediction errors are likely to be greatest at high faecal N concentrations (> 2%) where predictions are likely to under-estimate faecal N concentration. Prediction errors at high faecal N concentrations are not likely to have adverse consequences in relation to decision making.

(b) Validation experiments with regard to the prediction of faecal $\delta^{13}C$ (and the calculation of dietary non-grass proportions) indicated areas of weakness in the pre-project calibration equation. Areas of weakness were apparently associated with pasture type, geographical location, and possibly seasonal conditions. The effect of geographical location may have been associated with climatic effects, especially those associated with latitude such as temperature. Although the expansion of the calibration sample set should have improved robustness of the post-project calibration equation compared with the pre-project equation, areas of weakness are still likely to exist and there is a need for further validation and expansion.

(c) Based on the results of this project it can be concluded that there is a critical need to give priority to the validation and expansion of other F.NIRS calibrations, especially calibration equations for predicting diet quality.

7.2 Recommendations

Ultimately, the benefits and widespread use of F.NIRS technology, whether for research or commercial purposes, will depend of the predictive accuracy of calibration equations and also on the range of attributes that can be accurately predicted. Commercial use of F.NIRS is currently only small-scale but, provided the technology continues to develop, particularly regarding predictive reliability, commercial usage is likely to increase. Use of F.NIRS as a research tool in experiments with grazing cattle in northern Australia is now well entrenched. Indeed, because of the simplicity of faecal sampling and F.NIRS analysis, it can be argued that F.NIRS, as an interpretative tool, should be regarded as essential in any experiment involving grazing beef cattle. With all the above in mind there is a need to continually monitor the performance F.NIRS calibration equations and to expand the calibration sets to enhance predictive reliability. Specific recommendations have been clearly detailed in previous reports, viz. those in relation to NAP3.121 (Coates 2004) and NBP.302 (Dixon 2008).

8 Bibliography

Coates DB (1998) Predicting diet digestibility and crude protein content from the faeces of grazing cattle. Final Report of project CS.253. Meat Research Corporation, Sydney, Australia.

Coates DB (2004) Improving reliability of faecal NIRS calibration equations. Final Report of Project NAP3.121. Meat and Livestock Australia, Sydney.

Coates DB, Dixon RM (2008) Development of near infrared analysis of faeces to estimate non-grass proportions in diets selected by cattle grazing tropical pastures. *Journal of Near Infrared Spectroscopy 16, 471-480.*

Dixon RM (2008) Utilizing faecal NIRS measurements to improve prediction of grower and breeder cattle performance and supplement management, Final Report of Project NBP.302. Meat and Livestock Australia, Sydney.