Assessing the accuracy of current near infra-red reflectance spectroscopy analysis for fresh grass-clover mixture silages and development of new equations for this purpose

Article

Accepted Version

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1016/j.anifeedsci.2018.03.009

Publisher: Elsevier

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

1 **Assessing the accuracy of current near infra-red reflectance spectroscopy**

2 **analysis for fresh grass-clover mixture silages and development of new**

3 **equations for this purpose**

4 A. L. Thomson[1], D. J. Humphries[1], C. Rymer[1,] J. Archer[2], N. Grant[2], and C. K.

5 Reynolds[1]

6

7 [1] *Centre for Dairy Research, School of Agriculture, Policy and Development,*

8 *University of Reading, Whiteknights Road, Reading, RG6 6AR, UK*

9 [2] *Agri-Food and Biosciences Institute, Large Park, Hillsborough, Co Down, BT26*

10 *6DR, Northern Ireland*

11

12 Corresponding author: C. K. Reynolds. Email: c.k.reynolds@reading.ac.uk

13

14

15

Abstract

The purpose of this study was to ascertain whether Near Infra-Red Reflectance Spectroscopy (NIRS) prediction equations calibrated on grass silage samples, could accurately predict the chemical composition of mixed grass-clover silage samples, and furthermore, to develop and calibrate new grass-clover equations should the grass-based equations be insufficiently accurate for these silages. A set of 94 silage samples from mixed grass-clover swards (clover concentration (CC) ranging from 4 to 1000 g/kg as fed; determined manually) were analysed for chemical composition using reference laboratory techniques, *in vivo* digestible organic matter in the dry matter (DOMD, in sheep), and *in situ* degradability of dry matter and crude protein (in cows). The same samples were scanned fresh (undried and unmilled, as is standard practice for silage analysis within UK laboratories) using NIRS (at AFBI, Northern Ireland) and grass-based prediction equations applied. Predicted and observed results were compared. Of 15 chemical components that were tested for prediction accuracy, only volatile-corrected dry matter and nitrogen were well predicted (RPD values of 4.9 and 2.4 respectively, with low root mean square errors of prediction (RMSEP)). Neutral detergent fibre and DOMD showed low RPD values, however the predicted and observed datasets had no significant bias between them and were therefore also considered as fit for purpose. Variables with significant bias between predicted and observed datasets that were not considered suitably accurate included crude protein, acid detergent fibre, microbial dry matter yield and the effective degradability of protein. For many components, bias could be attributed at least in part to CC and changes in the fractionation of nutrients present. For some variables such as crude protein, grass-based equations were sufficiently accurate at low CCs but became inaccurate as CC increased, as expected. In response to inadequate prediction

accuracy of certain nutrients, new grass-clover equations were calibrated using the obtained spectra. These were validated and results indicated that the grass-clover-based equations outperformed their grass-based counterparts. The adoption of new grass-clover equations, or alternatively, with further development, the use of a CC correction factor to the existing grass-based equations, is recommended for commercial laboratories offering undried and unmilled silage analysis on samples containing clover.

Abbreviations: ADF, acid detergent fibre; aNDF, neutral detergent fibre; CC, clover concentration; CP, crude protein; VCODM, volatile corrected oven dry matter; EDN, effective degradable nitrogen; EDDM, effective degradability of dry matter; EE, ether extract; FiM, Feed into Milk; OM, organic matter; LA, lactic acid; MDM, microbial dry matter; N, nitrogen; NH_3-N, ammonia nitrogen; NIRS, near infrared reflectance spectroscopy; NMSC, normal multiplicative scatter correction; r^2, coefficient of determination of cross validation; RMSEP, root mean standard error of prediction; RPD, ratio of standard deviation of the measured population to the standard error of prediction; SEC, standard error of calibration; SECV, standard error of cross validation; SEP, standard error of prediction; SNVD, standard normal variate de-trending; TMR, total mixed ration; TVC, total volatile content; TVFA, total volatile fatty acids; WMSC, weighted multiplicative scatter correction; WSC, water soluble carbohydrate.

1. Introduction

66   Near Infra-Red Reflectance Spectroscopy (**NIRS**) is a relatively rapid and inexpensive

67   technique, routinely used to provide nutritional analysis of silage and other livestock

68   feeds in the dairy and beef industries. However, obtaining accurate results requires

69   robust prediction equations. This is particularly relevant to the UK where most silages

70   are analysed 'fresh' (i.e. undried and unmilled) for rapid through-put in comparison to

71   Europe where analysis of dried, ground samples is more common. Dry analysis

72   requires lengthy sample preparation but has the benefit of increased precision of NIRS

73   prediction, partly explained by the increased homogeneity of ground samples as well

74   as the stability of the feedstuff after the removal of water (Sorensen, 2004). Currently

75   UK laboratories do not offer NIRS equations for grass-legume mixtures, instead, a

76   prediction equation with a monoculture grass-based calibration is used for a number

77   of different grass and legume-based forages.

78   This study focusses on NIRS analysis for grass-clover silages, since clover is

79   thought to be present within grass swards on 70% of UK dairy farms, and therefore is

80   likely to be the most widely-grown forage legume in the UK (DEFRA, 2015).

81   Furthermore, clover-containing forages are thought to be a promising feed to increase

82   sustainability on farms due to reduced inorganic fertilser required for growth in

83   comparison to ryegrasses (Elgersma *et al.*, 2000), while maintaining high yields of milk

84   or meat due to a fast rate of passage promoting intake (Dewhurst *et al.*, 2009; Copani

85   *et al.*, 2016). A preliminary study has shown that the current NIRS analysis available

86   for use on grass silages in the UK has poor prediction accuracy of crude protein, pH

87   and lactic acid when used on mixtures containing both clover and grass (Davies *et al.*,

88   2012). However, Davies *et al.* (2012) did not evaluate the degradability of dry matter

89   (**DM**), nitrogen (**N**), or the apparent total tract digestibility of organic matter (**OM**; from

90   which metabolisable energy (**ME**) is calculated) for prediction accuracy, despite these

91   nutrient fractions being very important for diet formulation when balancing the ratio of

92   metabolisable protein to metabolisable energy supply. Imbalances in the degradable

93   protein to fermentable energy ratio will result in poor N use efficiency. Creating

94   calibration equations for grass-clover silages poses a challenge because these silages

95   are a mixture of two (or more) forage species, meaning that any resulting equation

96   must be able to deal with a broad spectrum of sample composition. To date, the

97   majority of forage-based NIRS calibrations have focussed on predicting the nutritional

98   composition of just one species, and moreover, in a few instances where mixtures

99   were analysed using NIRS, typically the focus of the study was on predicting botanical

100   composition rather than chemical composition (Wachendorf *et al.*, 1999; Cougnon *et*

101   *al.*, 2014; Karayilanli *et al.*, 2016).

102   The objective of this study was primarily to assess the adequacy of a grass

103   silage-based prediction equation, commonly used in the UK for predicting chemical

104   composition, when it was applied to grass silage samples that contained clover in

105   varying concentrations. Subsequently, a secondary objective was to investigate

106   whether using grass-clover based prediction equations could improve accuracy of

107   predicted chemical composition.

108

109   **2. Material and methods**

110   *2.1 Experimental design*

111   In total, 94 grass-clover silages were sourced from commercial farms and transported

112   to the Centre for Dairy Research (**CEDAR**), (Arborfield, Reading, UK) for processing.

113   Samples were acquired from a diverse range of UK farms to ensure maximum

114   variation within the sample set, in line with the findings of Cougnon *et al.* (2014) for

115   sourcing robust calibration data. Silage was collected over three consecutive years

116 (2012/13, 2013/14, and 2014/15). The clover content range of greatest importance

117 was deemed to be 300 - 600 g/kg DM as a more even distribution of grass-clover

118 within a ley has been shown to create the most advantageous conditions for growth

119 and promote symbiotic N fixation (Nyfeler *et al.,* 2011; Luescher *et al.,* 2014); although

120 samples containing < 300 and > 600 g/kg DM clover were also included to provide

121 sufficient range for statistical analysis and equation evaluation.

122

123 *2.2 The silage sample set*

124 *2.2.1 Sample description.* The set of 94 silage samples consisted of 58 bales and 36

125 samples from clamps which were collected from 50 different locations distributed

126 across the UK. Of the samples where the clover variety was known (n=65) 66 % were

127 red clover, 20 % were white clover and 14 % were a mixture of both. Different cuts

128 were also represented within the set with 36 first, 20 second, 16 third and 4 fourth cut

129 silages (harvest number not reported for 22 samples). The mean CC within the set

130 was 310 g/kg DM (Table 1). The sample containing the least clover contained 4 g/kg

131 DM clover and two samples contained 1000 g/kg DM clover, however all samples

132 originated from swards that were grass-clover mixtures. Twenty-three of the 94

133 samples contained < 70 g/kg DM CC and were conosidered a 'minimal' clover group

134 for which we hypothesised prediction accuracy would be similar to that of a pure grass

135 silage. The measured concentration of weed species within samples (any species

136 other than grass or clover) ranged from 0 - 380 g/kg DM with a mean of 50 g/kg DM.

137 *2.2.2. Sample processing.* Samples sourced were either unchopped bales or chopped

138 clamped material. If in the form of an unchopped bale, it was mixed and chopped in a

139 feeder wagon (Hi-Spec Mix Max, Hi-Spec Engineering, Co. Carlow, Ireland) for 45

140 minutes to minimize variability in chop length. Clamp silages that were already

141 chopped, were mixed in a DataRanger diet mixer which did not contain knives

142 (American Calan, Northwood, NH, USA). The DM content of the silage was estimated

143 from the loss in weight of a subsample after it has been repeatedly placed in a

144 microwave oven (Belling 384TC, 850 Watts) until a constant weight was achieved.

145 From this determination, the amount of silage (fresh weight) required for feeding an

146 individual sheep for 63 days was calculated. This amount was then weighed into

147 polythene bags with one days' feed per bag, the air was removed under vacuum, and

148 the bags were sealed and stored frozen (-20°C) until required. Frozen subsamples of

149 each silage were stored separately for future analysis of chemical and botanical

150 composition.

151

152 *2.3 Nutritional analysis*

153 *2.3.1 NIRS analysis* A 2 kg frozen subsample of each silage was sent to the Agri-Food

154 and Biosciences Institute (**AFBI**; Hillsborough, Northern Ireland) where the reference

155 chemical composition of the silages was determined using UKAS accredited methods

156 and NIRS spectra were obtained. Before scanning, all samples were further chopped

157 by hand to approximately 2.5 cm lengths and then thoroughly mixed. Two separate

158 packages were prepared by wrapping approximately 100 g of fresh sample in non-

159 PVC cling film (Park *et al.*, 1999). These packages were then placed in a rectangular

160 coarse transport cell and scanned through a Foss NIRSystems 6500 instrument (Foss,

161 Hillerød, Denmark). The optical values for each scan were recorded as Log

162 1/Reflectance over the range 400-2498 nm at 2 nm gaps using the ISI v3.10 (Infrasoft

163 International, Port Matilda, PA, USA) software.

164 *2.3.2 Laboratory reference analyses* Dry matter was determined in a forced-air oven

165 and corrected for the loss of VFAs, lactic acid (**LA**), alcohols and ammonia (Porter and

166 Murray, 2001) and reported as volatile-corrected oven dry matter (**VCODM**). Ash was

167 measured through combustion in a muffle oven at 550°C for 18 h. Lactic acid and other

168 volatile compound measurements (total volatile fatty acids (**TVFA**) were determined

169 using gas chromatography following extraction of representative samples in distilled

170 water (Erwin *et al.*, 1961; Givens *et al.*, 2009). Nitrogen (**N**) was measured using the

171 macro Kjeldahl method 954.01 (AOAC, 2000) and Ammonia-N (**NH$_3$-N**) was

172 determined using a calibrated ammonia ion selective electrode, which required 30 g

173 silage soaked in 150 ml of purified water for 18 h at 4°C. (McDonald *et al.,* 1981; Orion

174 Research, 1990). Both ether extract (**EE**) and water soluble carbohydrate (**WSC**) were

175 measured on dried and ground samples: EE according to AOAC method 920.29

176 (AOAC, 1990), and WSC as described previously (Fuller, 1967). Dried and ground

177 samples were subsequently passed on to Trouw Nutrition (Ashbourne, Derbyshire)

178 who performed analyses for neutral detergent fibre (**aNDF**) and acid detergent fibre

179 (**ADF**) both inclusive of residual ash using Fibrecap equipment (Foss, Hillerod,

180 Denmark) (Robertson and Van Soest, 1981; Kitcherside *et al.*, 2000; Mertens *et al.*,

181 2002). A further 200 g of silage was manually separated into clover, grass and other

182 species to determine the CC of the silage. This procedure was predominantly

183 performed by the same individual to minimise human error. Resulting fractions were

184 then dried to determine species composition on a DM basis. *In vivo* reference methods

185 were performed at CEDAR to determine silage digestibility and degradability.

186

187 *2.4. In vivo analyses.*

188 *2.4.1 In vivo Digestibility* Eighteen Mule x Texel wether sheep originating from a local

189 breeder were used to measure *in vivo* silage digestibility using a series of 3 x 3 Latin

190 square design experiments so that the final digestibility values comprised the mean of

8

191 measurements from three different animals. Each sheep was fed a silage sample *ad*

192 *libitum* (with 10% refusals) for 16 d adaption followed by a 5 d sampling period during

193 which sheep were placed in a metabolism crate for faeces and urine collection as

194 described previously (Givens *et al.*, 1989; Bratzler, 1951). All *in vivo* procedures were

195 licensed and monitored by the UK government Home Office under the Animal

196 (Scientific Procedures) Act 1986.

197 Sheep were enrolled on the study when they reached adult weight at > 30 kg.

198 Their diet was supplemented with 20 g/d of a general purpose vitamin/mineral mixture

199 for sheep (Countrywide, Evesham, Worcestershire, UK) and the weights of feed

200 offered and refused was recorded each day during the collection period. A subsample

201 of feed was taken and analysed for DM and ash to calculate OM content. Refused

202 feed was also corrected for DM. Out of the 94 samples, 4 were excluded from *in vivo*

203 analysis as there was insufficient material for the 9 week feeding schedule, but were

204 still used for all other analyses. Complete collections of faeces were taken for each

205 sheep. Each days' faecal material from the 5 d collection period was refrigerated at <

206 4°C until bulked together on d 5, thoroughly mixed and three 200 g subsamples

207 obtained. These subsamples were immediately placed in a forced air oven at 60°C for

208 72 h to determine DM content. Dried samples were then bulked, ground and a further

209 subsample was placed in a muffle oven for combustion at 500°C for 16 h for

210 determination of OM content. Digestibility results have been presented as digestible

211 organic matter in total dry matter (**DOMD**, g/kg DM).

212 *2.4.2 In situ degradability.* Degradability values were obtained using an *in situ* method

213 with rumen cannulated Holstein-Friesian dairy cattle. These cattle were housed in a

214 dedicated metabolism unit, fed a commercial grass-maize based total mixed ration

215 (**TMR**) diet once daily and milked twice daily at 0600 h and 1600 h approximately.

216 Fresh samples of each silage were placed in porous (43 µm pore size) bags that were

217 sequentially incubated in the rumen for six time intervals (3, 6, 12, 24, 48, and 72 h)

218 using a complete exchange method as described previously (Lovett *et al.* 2004).

219 Replicates were obtained by repeating the procedure with three different animals. To

220 quantify '0' hour washing loss, three further bags per silage were placed in a tub of

221 cold tap water and swirled for 5 minutes. All bags were washed (Zanussi SupeLluxe,

222 Electrolux plc, Luton, UK) on a 53 min cold wash cycle, dried (at 60°C), and weighed

223 for the determination of DM degradability, then further analysed for N (as described

224 previously). The solubility (S) of DM and N was determined by adding 1 g of DM to 30

225 ml of water and stirring for 5 minutes every half hour for a period of 2 h, the insoluble

226 material was then filtered (Whatman filter paper grade 4, Sigma-Aldrich, MO, USA)

227 (Hvelplund and Weisbjerg, 2000). The filter paper and substrate was then dried and

228 weighed to determine DM solubility by difference and residual N was measured as

229 described previously.

230   The percentage of material degraded at each time-point was used to plot a

231 degradation curve as described by Ørskov and McDonald (1979). Degradability

232 fractions termed 'a', 'b' and 'c' were obtained from the intercept, asimptote and slope

233 of the curve. Fraction 'a' contained material that is apparently degraded almost

234 immediately upon ingestion and 'b' contained the remaining insoluble but degradable

235 material with 'c' being the rate of degradation of 'b'. Two different approaches were

236 used to calculate effective degradabilitiy (**ED**) based on the above fractions. To ensure

237 the best comparison with predicted data, the ED of nitrogen (**EDN$_{FIM}$**) and of dry matter

238 (**EDDM$_{FIM}$**) were calculated using the 'Feed into Milk' (**FiM**) rationing software

239 equations (Equation 1). In this equation the outflow rate of small (k_{liq}) and large (k_f)

240 particles was standardised at 0.075 and 0.045 respectively to fairly compare against

241   predicted data. EDDM$_{FIM}$ was converted to microbial dry matter (**MDM$_{FIM}$**, g/kg DM)

242   using standard equations to convert EDDM into ATP supply as described previously

243   (FiM consortium, 2004).

244

245   Equation 1.   ED$_{FIM}$ = (0.9s/(0.9+k$_{liq}$))+(b$_D$ c/(c+k$_{liq}$))+(bc/(c+k$_f$))

246   Where s is the soluble proportion, k$_{liq}$ is the fractional outflow rate of the liquid pool (0.075), b$_D$
247   is the degradable small particle proportion, b is the degradable large particle proportion, c is
248   the fractional degradation rate of b, and k$_f$ is the fractional outflow rate of the large particle
249   pool (0.045).
250

251   A second, simpler, approach was also tested simultaneously to calculate the

252   ED of N and DM using 0.08 as the standard outflow rate (k) of all particles (**EDN$_{0.08}$**,

253   and **EDDM$_{0.08}$**) (Equation 2; Ørskov and McDonald, 1979).

254

255   Equation 2.   ED = a+bc/(c+k)

256   Where a is the rapidly degraded, b is the slowly, potentially degradable proportion, c is the
257   fractional rate of degradation of b, and k is the fractional outflow rate of material (0.08 h^{-1}).
258

259   *2.5 Statistical analysis*

260   *2.5.1 Tests of relationships and trends within the measured dataset.* Statistical

261   analysis was conducted using Genstat 16th Edition (VSNI, Hemel Hempstead, UK).

262   Composition of the silages was predicted from NIRS spectra using equations

263   developed for the UK Forage Analysis Assurance (FAA) group (www.faagroup.co.uk)

264   initially using 136 grass silage calibration samples from the studies reported by Park

265   *et al.* (1997, 1998) which were regularly updated with new spectra over time for most

266   chemical component variables other than those requiring in vivo reference analyses.

267   The measured dataset has been presented as maximum, minimum, mean and

268   coefficient of variation (CV%) values for each measured variable. The effect of CC on

269   each of the other variables was tested by grouping samples into minimal, low, medium

270   and high groups (which are equal quartiles of the dataset; representing samples within

271   the ranges of < 70, 70 - 250, 250 - 500 and > 500 g/kg DM CC respectively) which

272   were compared using analysis of variance (**ANOVA**). A post hoc Tukey test was

273   performed to determine whether there were significant differences between the means

274   of the 4 groups. The means of the observed and NIRS predicted datasets were

275   compared using a student's t-test to determine significance. Crude protein (**CP**) was

276   not directly measured or predicted but calculated using either measured or predicted

277   N and VCODM values (6.25 x Total N on a DM basis). For all dry matter values

278   throughout this study, VCODM has been used rather than DM, in accordance with the

279   industry standard used by UK laboratories. For ash, EE, WSC, ADF, and aNDF

280   (variables where the measured concentration is produced from a dry sample)

281   equations were produced that predicted concentrations on both a fresh basis and

282   directly on a DM basis.

283   *2.5.2 Tests of prediction accuracy during validation.* For the grass-based prediction

284   equation results, the difference between laboratory assays and NIRS predicted values

285   was calculated using measured minus predicted values and is henceforth termed

286   'bias'. Relative root mean square standard error of prediction (**RMSEP** as a

287   percentage of the measured mean), ratio of the standard error of prediction to the

288   standard deviation of the measured dataset (**RPD**) as recommended by Williams

289   (2014), and the R-squared value of the relationship between observed and predicted

290   data (r^2) were used to measure prediction accuracy.

291   *2.5.3 Calibration of new NIRS equations.* To create new grass-clover prediction

292   equations, different data pre-treatment methods were first assessed by varying use of

293   derivitives, gap, smoothing and scatter correction. All calibrations were performed

294   using the WinISI III v1.50 (Infrasoft International, Port Matilda, PA, USA) software.

12

295   They were carried out as Modified Partial Least Squares regressions over the range

296   1100-2498 nm using a 2 nm gap. To account for any sub-sampling error the root mean

297   square difference of each sub-sample was calculated using the WinISI III v1.50

298   software. An upper limit of 5000 was used to judge poor replication meaning any

299   sample with a root mean square greater than 5000 would be removed. None of the

300   samples in the calibration set were above this limit. Raw data and two derivatives

301   were tested in the process (Raw (0,0,1,1), 1st Derivative (1,4,4,1) and 2nd Derivative

302   (2,10,5,1)) and three scatter corrections (Standard Normal Variate Detrending

303   (**SNVD**), Normal Multiplicative Scatter Correction (**NMSC**) and Weighted Multiplicative

304   Scatter Correction (**WMSC**)) for each of the derivatives. The maximum number of

305   terms set for each equation was 11. There were three elimination passes carried out

306   and the cross validation value was set at 6 in which the calibration set was divided into

307   six groups with one group removed sequentially and predicted using a calibration

308   formed using the remaining samples. The validation errors were combined to give a

309   standard error of cross validation (**SECV**). The optimal equations were those with the

310   lowest SECV. The combination of data pre-treatment giving the optimal prediction

311   model is shown in supplementary table 1 for each variable. The optimal equation was

312   compared against the industry standard method, based on the study of Park *et al.*

313   (1997), which was taking the first derivative (1,4,4,1) with SNVD scatter correction and

314   a repeatability file (a file containing multiple spectra from the same sample measured

315   under different conditions, designed to reduce the variability caused by differing

316   environmental conditions and instruments). Differences between the optimal

317   equations and the industry standard equations were small, therefore further validation

318   was performed using the industry standard equations as these were the most likely to

319   be utilised commercially. For the purposes of a validation test, 10 samples were

320   removed from the dataset and tested using the remaining equation. These samples

321   were chosen by including the very first sample to be collected and then every tenth

322   sample in order of their arrival at CEDAR for processing.

323

324   **3. Results**

325   *3.1 Sample chemical composition*

326   The silages contained a wide range of chemical composition with LA, WSC and TVFA

327   being the nutritional characteristics with the greatest variance of those measured.

328   Volatile corrected dry matter of the silages was evenly distributed with a mean of 395

329   g/kg. Measured CP concentration (calculated from N and VCODM) ranged from 57 to

330   215 g/kg DM and with a mean of 138 g/kg DM.

331   With the exception of ash, aNDF, and WSC, the concentration of all other

332   measured variables were affected by the CC of the sample when grouped into

333   minimal, low, medium, and high clover groups (Table 2). VCODM and N were

334   significantly increased in the high clover group (>500 g/kg DM CC) relative to the other

335   three groups (both $P < 0.001$), as was CP with the exception of the medium group

336   which contained an intermediary CP concentration ($P < 0.001$). Degradability

337   parameters calculated using the Ørskov and McDonald (1979) model and DOMD were

338   lowest in the high clover group (all $P < 0.04$) and numerically highest in the low clover

339   group (60-240 g/kg DM CC), however, when degradability parameters were calculated

340   using FiM equations, differences between clover groups were non-significant.

341   Fermentation end products (LA, TFVA and TVC) decreased in concentration

342   sequentially as CC increased (all $P < 0.003$) while pH was similar for minimal, low and

343   medium groups and higher for the high clover group ($P < 0.001$). NH_3-N was also

14

344 highest in the high clover group in comparison to the minimal clover group while the

345 other two groups contained intermediate concentrations of NH_3-N (P < 0.02).

346

347 *3.2 Validation of current grass-based NIRS equations*

348 Using data from the present study grass-clover sample set to verify the accuracy of

349 the current grass-based prediction equations, a wide range of prediction accuracy was

350 observed depending on the chemical component tested (Table 3). Volatile corrected

351 dry matter and N showed good prediction accuracy with RPD values of 4.92 and 2.35

352 respectively, and no significant difference between observed and predicted means.

353 Furthermore, the relationship between the observed and predicted data for both these

354 variables closely followed a line of parity (Figure 1) especially at low concentrations.

355 However, all other variables led to RPD values that were <2 denoting inadequate

356 performance. Digestible organic matter in total dry matter, and aNDF, had low relative

357 RMSEP (both <10% of the observed mean) and no significant difference between the

358 observed and predicted means which could be considered acceptable despite having

359 an RPD value <2. For these variables the slope of the relationship between observed

360 and predicted data followed a line of parity however there was greater variability in the

361 relationship than was seen for VCODM and N (Figure 1). Crude protein prediction

362 showed a relatively high RPD value (1.58) and good correlation between predicted

363 and observed data (r^2 =0.75) however the slope of the relationship did not follow a line

364 of parity (Figure 1) leading to a significant bias (P < 0.005) for under-estimation at

365 higher concentrations with the average under-estimation being 12.4 g/kg DM.

366 Fermentation characteristics (LA, pH, TVC and TVFA) all showed intermediate

367 prediction accuracy with RPD values ranging from 1.15 to 1.22. Of these variables, LA

368 in particular had a very high relative RMSEP at 71% of the observed mean as a result

15

369 of high variability in prediction accuracy where concentration was low (Figure 1). For

370 both TVC and TVFA there was a significant bias towards over-estimation (both $P <$

371 0.001). Poor prediction accuracy (RPD value <1) was observed for NH_3-N, ADF, EE,

372 EDN_{FIM}, and MDM_{FIM} all of which showed a significant bias between the predicted and

373 observed means (all $P < 0.001$). Of special note, EDN_{FIM} and MDM_{FIM} showed the least

374 prediction accuracy of all the variables tested with a significant over-estimation for

375 EDN_{FIM} of 139 g/kg N and an under-estimation for MDM_{FIM} of 17 g/kg DM. Moreover,

376 predicted and observed data showed little correlation (Figure 1) indicated by r^2 values

377 of 0.01.

378 The degree of variation and the magnitude of bias in relation to sample CC is

379 illustrated in Figure 2 using CP and EDN_{FIM} as examples which are crucial to diet

380 formulation. In the case of CP, prediction bias in samples containing 800-1000 g/kg

381 DM CC is greater than 30 g/kg DM (Figure 2a), and similarly for EDN_{FIM}, a prediction

382 bias greater than 200 g/kg N was observed in this very high CC range. Meanwhile,

383 bias was comparatively lower in the minimal clover group (< 70 g/kg DM CC) at 6 g/kg

384 DM for CP and 103 g/kg N for EDN_{FIM} reflecting the degree of bias that might be

385 expected for a pure grass sample.

386

387 *3.3 Validation of new grass-clover equations*

388 Following production of new equations using the NIRS spectra from the grass-clover

389 silages in the sample set, a cross validation test indicated 12 out of 21 new equations

390 had a relative SECV of 10% of the observed mean, suggesting a good calibration was

391 achieved for these variables (Table 4). VCODM, pH, aNDF, ADF, and $EDDM_{0.08}$ were

392 amongst the strongest calibrations according to cross validation while TVC, WSC,

393 TVFA, Alcohol and LA were the least robust. For variables where both a fresh and a

394 DM basis equation were produced, the equation that predicted on a fresh basis gave

395 the more accurate result for ash, EE and WSC, whereas the opposite was true for

396 ADF and aNDF, where the equation that predicted concentration on a DM basis was

397 more accurate.

398 A validation test was also applied to the new grass-clover prediction equations

399 through removal of 10 samples from the calibration data-set (Table 5). Seven variables

400 gave an RPD value > 2 denoting good accuracy including VCODM, ADF, aNDF, EDN

401 and N. Additionally the RPD score of all values were improved relative to prediction

402 accuracy using the grass-based equations, which was reflected in greatly reduced

403 bias, for example, new equations reduced crude protein mean bias from -12.4 to -0.82

404 g/kg DM and EDN mean bias improved from 139 to 12 g/kg N on average. The new

405 alcohol and EE (DM basis) equations gave a low RPD value (>1) suggesting these

406 equations are unlikely to be suitable for use without further improvement.

407

408 **4. Discussion**

409 *4.1 Chemical composition and clover concentration*

410 The wide range of samples collected in this study provided a robust test for the current

411 grass-based prediction equations. The sample set was dominated by samples

412 containing predominantly grass with only a quarter of the samples obtained containing

413 > 500 g/kg DM CC. Roughly half the total number of samples obtained were below the

414 minimum optimum clover inclusion rate of 300 g/kg DM suggested by Nyfeler *et al.*

415 (2011). This may be due to the sample set comprising a greater number of first cut

416 silage samples than second, third or fourth cuts in which CC would have been greater

417 due to warmer and drier conditions in the latter half of the year (Chmelikova *et al.*,

418 2015).

419   Crude protein concentration (ranging from 57 to 215 g/kg DM with a mean

420   concentration of 138g/kg DM) indicated that, although some of the samples contained

421   very high levels of crude protein, mean concentration was similar to that expected for

422   well fertilised modern grass silages which have been observed ranging from 120-270

423   g/kg DM (Burns *et al.*, 2015). This mean is also significantly lower than those reported

424   in published feed composition tables for crude protein concentration of grass-clover

425   silages e.g. 173 g/kg DM; AFRC (1993).

426   The supply of effective degradable N (EDN) is another important factor in diet

427   formulation. High concentrations (>700 g/kg N) of rapidly degraded protein in the

428   rumen can be wasteful as there is insufficient time for bacterial N capture, which is

429   often a characteristic of legume silages (Coblentz and Grabber, 2013; Dewhurst,

430   2013). In this study, average EDN_{FIM} was 623 g/kg N so within the optimal range and

431   lower than values cited in other studies, for example, in another study, measured

432   average EDN of grass-clover silages was 880 g/kg N at an assumed passage rate of

433   0.05/hr (Hvelplund and Weisbjerg, 2000). The discrepancy may be due to clover

434   varieties in this sample set being predominantly comprised of red clovers containing

435   the enzyme poly-phenol oxidase which is thought to reduce proteolysis in the rumen

436   (Lee *et al.*, 2009). Digestibility, $EDDM_{0.08}$, and $EDN_{0.08}$ all showed a similar pattern

437   where the low group (70 – 250 g/kg DM CC) gave the highest value and the high group

438   (> 500 g/kg DM CC) the lowest suggesting inclusion of clover between 70 - 250 g/kg

439   DM is an optimal range for digestibility and degradability. Poor digestibility in the high

440   group may relate to an increased maturity of clover and grass with higher lignification

441   in samples in that range (Nousiainen *et al.*, 2009). Increasing ratios of ADF:aNDF in

442   samples with a higher CC indicates the differing fractions of fibre present in legumes

443   in comparison to grasses, especially red clover which is largely comprised of stem

444   where ADF concentration is higher than in leaves (Alstrup *et al.*, 2016). There was a

445   notable decrease in volatiles content (LA, TVC and TVFA) and an increase in pH in

446   the high group relative to the other quartiles where values were generally similar. This

447   suggests samples with a very high CC were more difficult to ensile, perhaps due to

448   reduced availability of sugar to fuel bacterial activity that may also indicate an

449   increased maturity. Using a factor of 0.0157 of DOMD, mean ME within the sample

450   set was predicted as 9.9 MJ/kg DM which is considerably lower than recently

451   measured values for modern monoculture grass silages which ranged from 12-13

452   MJ/kg DM (Burns *et al.*, 2015).

453

454   *4.2 Using grass-based NIRS equations for clover containing samples*

455   The key objective of this project was to determine whether the current grass NIRS

456   equations could be applied to grass-clover samples and predict the concentrations of

457   chemical components with good accuracy. Variates that were considered most

458   important for correct diet formation included CP, EDN, MDM and DOMD as these are

459   the variables involved in balancing rumen degradable protein and energy supply.

460   Whilst the prediction accuracy of DOMD and some other variables (including VCODM,

461   N and aNDF) could be considered suitably accurate with relative RMSEPs of <10 %

462   of the measured mean, CP, MDM, and EDN were amongst the variables with high

463   relative RMSEPs and RPD values of less than 2.0 combined with a significant bias.

464   Similar results were seen in a smaller preliminary study of 58 grass-clover silages in

465   which the same equations were tested and crude protein was significantly under-

466   estimated by 22 g/kg DM on average (Davies *et al.*, 2012). The consequence of this

467   bias would be an imbalance in microbial N supply in the ration which is likely to lead

468   to reduced N use efficiency in cattle resulting in higher levels of N excretion in urine

19

469   and faeces contributing to environmental loading (Kebreab *et al.*, 2002). Under-

470   estimation of CP in silage samples could result in farmers under-valuing grass-clover

471   silages as protein sources, and compensating through an oversupply of expensive

472   bought-in protein within the concentrate portion of the diet. For CP and EDN,

473   increasing bias correlated with increasing CC. This might be explained by samples

474   containing a high concentration of grass being more similar in composition to the

475   calibration samples used to create the current grass-based equations. Also, bias may

476   be created due to different N fractionation within clover, with some fractions present

477   that are absent (or present in different concentrations) in grass, such as the

478   concentration of non-protein N (Chrenkova *et al.*, 2014).

479   When considerring the impact of the observed inaccuracies on diet formulation

480   it is estimated metabolisable protein, and not crude protein that is often the protein

481   fraction used for diet formulation. Crude protein multiplied by EDN (as a proportion of

482   total N) is used to calculate effective rumen degradable protein (**eRDP**) which is one

483   of the factors that determines metabolisable protein (alongside digestable undegraded

484   true protein, **DUP**) in diet formulation software. The effect of CC on calculated eRDP

485   bias is shown in Figure 2c. The opposing bias in EDN and CP cancel out to some

486   extent at low CC however the overall effect is an over-estimation of eRDP that

487   increases at higher CC. This may lead to an oversupply of fermentable energy in

488   relation to available protein for microbial N capture, creating an imbalance that could

489   reduce the efficiency of dietary nutrient utilisation. This would only be further

490   compounded by the inaccuracy seen in MDM prediction which is used to determine

491   the requirement for fermentable energy.

492

493   *4.3 Performance of new grass-clover-based equations*

494 Comparing the performance of new grass-clover equations with the grass-based

495 equations for use on clover-containing samples, and using relative SECV as a

496 measure of potential performance of the calibration, some of the new grass-clover

497 equations produced in this study are likely to perform well (including important variates

498 such as VCODM, N, EDN, and DOMD) whereas others have very high errors

499 (particularly the volatile compounds) and would require further development. The

500 accuracy of prediction for volatile compounds (LA, TVC, and TVFA) is notable in all

501 equations tested (both grass and grass-clover based) for producing poor reliability,

502 and volatile concentrations showed some of the greatest variation within the measured

503 sample set. Some of the lack of reliability in these equations could be due to the

504 variability of scanning undried and unmilled material rather than presenting the sample

505 in an homogenous, dried form (Sorensen, 2004).

506   For variates where the measured value was calculated on a dry sample (ash,

507 EE, aNDF, ADF and WSC) equations have been calibrated to give both a fresh and a

508 DM basis value. In most instances, the calibration for the fresh value was more robust,

509 however, because presenting information on a DM basis is widely practiced, fresh

510 values would be transformed based on VCODM values which would introduce further

511 error. Overall however, in a small validation test, new equations were better able to

512 predict all variables when compared to the accuracy of grass-based prediction

513 equations. The prediction of $EDN_{0.08}$ and $EDDM_{0.08}$ showed marked improvement over

514 the previous prediction accuracy for EDN_{FIM} and MDM_{FIM} using the grass-based

515 equations perhaps due to the reduced complexitiy of calculating these variables from

516 measured data. For example, calculating reference values for MDM_{FIM} from measured

517 degradability at different timepoints *in vivo* is a multi-step process involving many

518 different variables (such as corrections for solubility, fatty acid content and crude

519 protein concentration) and therefore it may not be feasible to predict such a value

520 based on NIRS spectra alone. These improvements would have a significant impact

521 on the accuracy of rumen degradable protein and fermentable energy prediction.

522

523 *4.4 Implementation of new equations*

524 The implementation of new grass-clover equations requires that nutritionists, feed

525 company representatives, and farmers are widely aware of the new option, and that

526 samples are correctly identified as containing a viable quantity of clover. Additionally,

527 grass-clover mixtures are just one example of alternative forages that are currently

528 gaining popularity, and it is unlikely that new equations can be created and

529 implemented for all of them due to the time needed to collect a sufficiently large group

530 of calibration samples. Therefore, it would be more convenient if one equation (such

531 as the current grass-based equation or alternatively a seperate general 'legume'

532 equation) could be adapted to analyse many different grass and legume based

533 forages. Another solution would be to use a two step process in which the CC of the

534 sample is predicted using NIRS, and then used to apply a correction to nutritional

535 predictions. A number of previous studies have used NIRS to determine the botanical

536 composition of a mixed sample (containing two species) with success for both grass-

537 clover (Cougnon *et al.*, 2014) and lucerne-grass silages (Karayilanli *et al.*, 2016)

538 however in all instances the calibration was performed on dry samples and therefore

539 further work is required to create an analysis for fresh samples that would be

540 appropriate for use in the UK.

541

542 **5. Conclusions**

543 For some variables, notably VCODM, N, DOMD and aNDF, current UK grass-based

544 equations were able to be applied to clover-containing samples with adequate

545 accuracy. However, in general, it was concluded from the evidence observed in this

546 study that the NIRS calibration equations developed for use on grass silages, could

547 not predict a number of key chemical components (including CP and EDN) with

548 sufficient accuracy, when used for grass-clover mixture silages. This was consistent

549 with the findings of a previous study (Davies *et al.*, 2012). Therefore, we suggest two

550 possible solutions that would be appropriate for uptake by UK laboratories: (i) the

551 introduction of new grass-clover prediction equations calibrated using the sample set

552 obtained for this study or (ii) the use of a correction factor that could be applied based

553 on the CC of the sample. Furthermore, in a wider sense, this study provides some

554 evidence that caution should be used whenever NIRS equations are applied to forage

555 mixtures where only one component of the mixture was represented within the

556 equation calibration set. Where possible, using an equation based on a specific

557 calibration set that is very similar to the material requiring analysis is likely to produce

558 the most accurate predictions.

559

References

571 AFRC, 2003. Energy and protein requirements of ruminants. An advisory manual

572   prepared by the AFRC advisory committee on responses to nutrients. CAB

573   International, Wallingford, UK.

574 Alstrup, L., Soegaard, K., Weisbjerg, M. R., 2016. Effects of maturity and harvest

575   season of grass-clover silage and of forage-to-concentrate ratio on milk

576   production of dairy cows. J. Dairy Sci. 99, 328-340.

577   http://dx.doi.org/10.3168/jds.2015-9802

578 AOAC, 1990. Official Methods of Analysis, 15th ed. Association of Official Analytical

579   Chemists , Washington, DC, USA.

580 AOAC, 2000. Official Methods of Analysis, 17th ed. Association of Official Analytical

581   Chemists, Washington, DC, USA.

582 Bratzler, J. W.,1951. A Metabolism Crate for Use with Sheep. J Anim Sci 10, 592-

583   601.

584 Burns, G. A., O'Kiely, P., Grogan, D., Watson, S., Gilliland, T. J., 2015. Comparison

585   of herbage yield, nutritive value and ensilability traits of three ryegrass species

586   evaluated for the Irish Recommended List. Irish Journal of Agricultural and

587   Food Research 54, 31-40. http://dx.doi.org/10.1515/ijafr-2015-0003

588 Chmelikova, L., Wolfrum, S., Schmid, H., Hejcman, M., Hulsbergen, K. J., 2015.

589   Seasonal development of above- and below-ground organs of Trifolium

590   pratense in grass-legume mixture on different soils. Journal of Plant Nutrition

591   and Soil Science 178, 13-24. http://dx.doi.org/10.1002/jpln.201400112

592   Chrenkova, M., Ceresnakova, Z., Weisbjerg, M. R., Formelova, Z., Polacikova, M.,

593   Vondrakova, M., 2014. Characterization of proteins in feeds according to the

594   CNCPS and comparison to in situ parameters. Czech Journal of Animal

595   Science 59, 288-295.

596   Coblentz, W. K., Grabber, J. H., 2013. In situ protein degradation of alfalfa and

597   birdsfoot trefoil hays and silages as influenced by condensed tannin

598   concentration. J. Dairy Sci. 96, 3120-3137. http://dx.doi.org/10.3168/jds.2012-

599   6098

600   Copani, G., Niderkorn, V., Anglard, F., Quereuil, A., Ginane, C., 2016. Silages

601   containing bioactive forage legumes: a promising protein-rich feed source for

602   growing lambs. Grass Forage Sci. 71, 622-631.

603   http://dx.doi.org/10.1111/gfs.12225

604   Cougnon, M., Van Waes, C., Dardenne, P., Baert, J., Reheul, D., 2014. Comparison

605   of near infrared reflectance spectroscopy calibration strategies for the

606   botanical composition of grass-clover mixtures. Grass Forage Sci. 69, 167-

607   175. http://dx.doi.org/10.1111/gfs.12031

608   Davies, D. R., Davies, G. K. & Morgan, C. T. (2012). Silage analysis-comparison of

609   58 Welsh farm silages analysed either by traditional wet chemistry or wet

610   NIRs. In: Kuoppala, K., Rinne, M. & Vanhatalo, A. (eds.) Proceedings of the

611   XVI International Silage Conference. Hämeenlinna, Finland, 2-4 July 2012.

612   MTT Agrifood Research Finland, pp. 270-271.

613   DEFRA, 2015. Farm Practices Survey: Basic Payment Scheme and greening -

614   England. National Statistics.

615   Dewhurst, R. J., 2013. Milk production from silage: comparison of grass, legume and

616   maize silages and their mixtures. Agricultural and Food Science 22, 57-69.

617   Dewhurst, R. J., Delaby, L., Moloney, A., Boland, T., Lewis, E., 2009. Nutritive value

618   of forage legumes used for grazing and silage. Irish Journal of Agricultural

619   and Food Research 48, 167-187.

620   Elgersma, A., Schlepers, H., Nassiri, M., 2000. Interactions between perennial

621   ryegrass (Lolium perenne L.) and white clover (Trifolium repens L.) under

622   contrasting nitrogen availability: productivity, seasonal patterns of species

623   composition, N-2 fixation, N transfer and N recovery. Plant and Soil 221, 281-

624   299. http://dx.doi.org/10.1023/a:1004797106981

625   Erwin, E. S., Marco, G. J., Emery, E. M., 1961. Volatile Fatty Acid Analyses of Blood

626   and Rumen Fluid by Gas Chromatography. J. Dairy Sci. 44, 1768-1771.

627   http://dx.doi.org/10.3168/jds.S0022-0302(61)89956-6

628   FiM consortium, 2004. Feed into milk. A new applied feeding system for dairy cows.

629   Nottingham University Press, Nottingham, UK.

630   Fuller, K. W., 1967. Automated determination of sugars. In: Automation in analytical

631   chemistry. European Technicon Symposia, 2-4 November 1966, Paris, France,

632   pp. 57-61.

633   Givens, D. I., Everington, J. M., Adamson, A. H., 1989. The digestibility and

634   metabolisable energy content of grass silage and their prediction from

635   laboratory measurements. Anim. Feed Sci. Technol. 24, 27-43.

636   http://dx.doi.org/10.1016/0377-8401(89)90018-7

637   Givens, D. I., Humphries, D. J., Kliem, K. E., Kirton, P., Deaville, E. R., 2009. Whole

638   crop cereals 1. Effect of method of harvest and preservation on chemical

639   composition, apparent digestibility and energy value. Anim. Feed Sci.

640   Technol. 149, 102-113. http://dx.doi.org/10.1016/j.anifeedsci.2008.05.007

641   Hvelplund, T., Weisbjerg, M. R., 2000. In situ techniques for the estimation of protein

642   degradability and postrumen availability. In: Givens, D. I., Owen, E., Omed, H.

643   M., Axford, R. F. E., (Eds) Forage Evaluation in Ruminant Nutrition, CAB

644   International, Wallingford, UK, pp. 233-258.

645   Karayilanli, E., Cherney, J. H., Sirois, P., Kubinec, D., Cherney, D. J. R., 2016.

646   Botanical Composition Prediction of Alfalfa-Grass Mixtures using NIRS:

647   Developing a Robust Calibration. Crop Science 56, 3361-3366.

648   http://dx.doi.org/10.2135/cropsci2016.04.0232

649   Kebreab, E., France, J., Mills, J. A. N., Allison, R., Dijkstra, J., 2002. A dynamic

650   model of N metabolism in the lactating dairy cow and an assessment of

651   impact of N excretion on the environment. J. Anim. Sci. 80, 248-259.

652   Kitcherside, M. A., Glen, E. F., Webster, A. J. F., 2000. FibreCap: an improved

653   method for the rapid analysis of fibre in feeding stuffs. Anim. Feed Sci.

654   Technol. 86, 125-132. http://dx.doi.org/10.1016/s0377-8401(00)00153-x

655   Lee, M. R. F., Theobald, V. J., Tweed, J. K. S., Winters, A. L., Scollan, N. D., 2009.

656   Effect of feeding fresh or conditioned red clover on milk fatty acids and

657   nitrogen utilization in lactating dairy cows. J. Dairy Sci. 92, 1136-1147.

658   http://dx.doi.org/10.3168/jds.2008-1692

659   Lovett, D. K., Deaville, E. R., Mould, F., Givens, D. I. & Owen, E. (2004). Using near

660   infrared reflectance spectroscopy (NIRS) to predict the biological parameters

661   of maize silage. Animal Feed Science and Technology, 115 (1-2), 179-187.

662   https://doi.org/10.1016/j.anifeedsci.2004.02.007

663   Luescher, A., Mueller-Harvey, I., Soussana, J. F., Rees, R. M., Peyraud, J. L., 2014.

664   Potential of legume-based grassland-livestock systems in Europe: a review.

665   Grass Forage Sci. 69, 206-228. http://dx.doi.org/10.1111/gfs.12124

666 McDonald, P., Edwards, R.A., Greenhalgh, J.F.D. 1981. Animal Nutrition 3rd edition,

667   Pearson Education Limited, Edinburgh Gate, Harlow, Essex, CM20 2JE.

668 Mertens, D. R., Allen, M., Carmany, J., Clegg, J., Davidowicz, A., Drouches, M.,

669   Frank, K., Gambin, D., Garkie, M., Gildemeister, B., Jeffress, D., Jeon, C. S.,

670   Jones, D., Kaplan, D., Kim, G. N., Kobata, S., Main, D., Moua, X., Paul, B.,

671   Robertson, J., Taysom, D., Thiex, N., Williams, J., Wolf, M., 2002. Gravimetric

672   determination of amylase-treated neutral detergent fiber in feeds with refluxing

673   in beakers or crucibles: Collaborative study. Journal of AOAC International 85,

674   1217-1240.

675 Nousiainen, J., Rinne, M., Huhtanen, P., 2009. A meta-analysis of feed digestion in

676   dairy cows. 1. The effects of forage and concentrate factors on total diet

677   digestibility. J. Dairy Sci. 92, 5019-5030. http://dx.doi.org/10.3168/jds.2008-

678   1833

679 Nyfeler, D., Huguenin-Elie, O., Suter, M., Frossard, E., Luescher, A., 2011. Grass-

680   legume mixtures can yield more nitrogen than legume pure stands due to

681   mutual stimulation of nitrogen uptake from symbiotic and non-symbiotic

682   sources. Agriculture Ecosystems & Environment 140, 155-163.

683   http://dx.doi.org/10.1016/j.agee.2010.11.022

684 Orion Research, 1990. Model 95-12 Ammonia electrode instruction manual. Orion

685 Park, R. S., Gordon, F. J., Agnew, R. E., Barnes, R. J., Steen, R. W. J., 1997. The

686   use of Near Infrared Reflectance Spectroscopy on dried samples to predict

687   biological parameters of grass silage. Animal Feed Science and Technology

688   68, 235-246. http://dx.doi.org/10.1016/S0377-8401(97)00055-2

689 Park, R. S., Agnew, R. E., Gordon, F. J., Steen, R. W. J., 1998. The use of near

690   infrared reflectance spectroscopy (NIRS) on undried samples of grass silage

691   to predict chemical composition and digestibility parameters. Animal Feed

692   Science and Technology 72, 155-167 http://dx.doi.org/10.1016/S0377-

693   8401(97)00175-2

694 Park, R. S., Agnew R. E., Gordon F.J., and Barnes R. J., 1999. The development

695   and transfer of undried grass silage calibrations between near infrared

696   reflectance spectroscopy instruments. Animal Feed Science and Technology

697   78:325-340. https://doi.org/10.1016/S0377-8401(98)00261-2

698 Porter, M. G., Murray., R. S., 2001. The volatility of components of grass silage on

699   oven drying and the inter-relationship between dry-matter content estimated

700   by different analystical methods. Grass Forage Sci. 56, 405-411.

701 Robertson, J. B., Van Soest, P. J., 1981. The detergent system of analysis and its

702   application to human foods. In: James, W., Theander, O. (Eds), The analysis

703   of dietary fribe in food. Marcer Decker Inc, NY, USA.

704 Sorensen, L. K., 2004. Prediction of fermentation parameters in grass and corn

705   silage by near infrared spectroscopy. J. Dairy Sci. 87, 3826-3835.

706   http://dx.doi.org/10.3168/jds.S0022-0302(04)73522-5

707 Wachendorf, M., Ingwersen, B., Taube, F., 1999. Prediction of the clover content of

708   red clover- and white clover-grass mixtures by near-infrared reflectance

709   spectroscopy. Grass Forage Sci. 54, 87-90. http://dx.doi.org/10.1046/j.1365-

710   2494.1999.00150.x

711 Williams, P., 2014. Tutorial: The RPD statistic: a tutorial note. NIR news 25, 22-26.

712 Ørskov, E. R., McDonald, I., 1979. The estimation of protein degradability in the

713   rumen from incubation measurements weighted according to rate of passage.

714   J. Agric. Sci. 92, 499-503. http://dx.doi.org/10.1017/S0021859600063048

715

716

717 **Table 1** The means, ranges and variation coefficients (CV) of chemical components
718 measured in a set of 94 diverse grass-clover silages from UK farms (in g/kg DM
719 unless otherwise stated).

Item	Min	Max	Mean	CV, %
ADF	229	513	335	10.2
Ash	58	158	97	20.6
aNDF	299	585	447	10.0
CC	4	1000	310	91.3
CP	57	215	138	24.7
Degradability				
$EDDM_{0.08}$†	217	626	472	16.4
$EDN_{0.08}$, g/kg N†	55	821	625	18.0
MDM_{FIM}‡	60	274	146	34.8
EDN_{FIM}, g/kg N‡	297	811	623	14.3
DOMD	400	766	632	10.6
EE	14.6	42.9	26.6	26.0
LA, g/kg	0.0	64.4	13.4	91.5
pH	3.6	6.7	4.6	13.5
N, g/kg	3.6	17.7	8.8	42.2
NH_3-N, g/kg DM*100	17.5	203	62.5	42.2
TVC, g/kg§	2.3	76.1	23.6	57.8
TVFA, g/kg¶	1.1	74.3	19.7	66.0
VCODM, g/kg	182	793	395	33.4
WSC	3.9	164	41.4	86.3

720 CC = clover concentration; EDDM = effective degradability of dry matter; EDN = effective degradable
721 nitrogen; DOMD = digestible organic matter in total dry matter; EE = ether extract; LA = lactic acid;
722 MDM = microbial dry matter yield; NH_3-N = ammonia nitrogen; TVC = total volatile content, TVFA =
723 total volatile fatty acids; WSC = water soluble carbohydrates.

724 † Degradability parameters determined by *in situ* incubation in the rumen, using the model of Ørskov
725 and McDonald (1979) ED = a+b[c/(c+k)] where a = rapidly soluble material; b = non-soluble but
726 degradable material; c = rate f degradation of b; and k = an assumed outflow rate of 0.08/hr.
727 ‡ Degradability parameters determined by *in situ* incubation in the rumen, using the Feed Into Milk
728 (FIM) Consortium (2004) model $ED_{FIM} = (0.9s/(0.9+k_{liq}))+(b_D c/(c+k_{liq}))+(bc/(c+k_f))$ where s = soluble
729 proportion, k_{liq} = fractional outflow rate of the liquid pool (0.075/hr), b_D = degradable small particle
730 proportion, b = degradable large particle proportion, c = fractional degradation rate of b, and k_f is the
731 fractional outflow rate of the large particle pool (0.045/hr).
732 § TVC is the sum of acetic, butyric, lactic, propionic and valeric acids plus ethanol and propanol.
733 ¶ TVFA is calculated as for TVC but excluding ethanol and propanol.
734

735 **Table 2** Differences in chemical components in 94 grass-clover silages grouped into
736 four quartiles (Minimal (Mi), Low (L), Medium (M) and High (H)) according to their
737 clover concentration (mean of each quartile, in g/kg DM unless otherwise stated).

Item	Clover concentration quartiles†				SED	P value‡
	Mi	L	M	H		
CC	34[a]	145[b]	335[c]	743[d]	28.9	0.001
n	23	24	24	23		
Chemical components						
ADF	311[a]	329[ab]	345[b]	356[b]	12.5	0.003
Ash	91.2	94.9	103.4	96.8	5.84	0.201
aNDF	465	452	443	432	16.5	0.229
CP	122[a]	130[a]	143[ab]	158[b]	9.3	0.001
Degradability						
$EDDM_{0.08}$§	470[ab]	501[b]	478[ab]	436[a]	21.7	0.032
$EDN_{0.08}$, g/kg N§	643[b]	682[b]	640[b]	531[a]	28.6	0.001
MDM_{FIM}¶	130	135	127	122	6.3	0.218
EDN_{FIM}, g/kg N¶	627	645	629	589	25.9	0.182
DOMD	647[b]	668[b]	631[b]	581[a]	18.5	0.001
EE	26.7[ab]	28.7[b]	27.4[ab]	23.2[a]	1.99	0.044
LA, g/kg	17.6[b]	16.2[b]	14.0[ab]	5.6[a]	3.42	0.003
pH	4.45[a]	4.41[a]	4.44[a]	5.23[b]	0.155	0.001
N, g/kg	7.8[a]	7.3[a]	7.9[a]	12.3[b]	0.93	0.001
NH_3-N, g/kg DM*100	48.3[a]	56.3[ab]	68.0[ab]	77.5[b]	9.71	0.018
TVC, g/kg ‖	28.4[b]	26.8[b]	24.0[ab]	14.3[a]	3.74	0.001
TVFA, g/kg ¥	23.3[b]	22.4[b]	20.8[ab]	11.6[a]	3.65	0.001
VCODM, g/kg	397[a]	350[a]	347[a]	498[b]	35.1	0.001
WSC	56.6	39.6	32.9	38.6	10.37	0.125

738 CC = clover concentration; EDDM = effective degradability of dry matter; EDN = effective degradable
739 nitrogen; DOMD = digestible organic matter in total dry matter; EE = ether extract; LA = lactic acid;
740 MDM = microbial dry matter yield; NH_3-N = ammonia nitrogen; SED = standard error of the difference
741 between means; TVC = total volatile content, TVFA = total volatile fatty acids; WSC = water soluble
742 carbohydrates.

743 † The 94 samples were sorted by ascending clover concentration and divided into four evenly sized
744 quartiles: 0-6%DM clover (Mi); 6-24% clover (L); 25-49% clover (M); and 50-100% clover (H).
745 ‡ The probability of there being no significant difference between treatment means determined using
746 Analysis of Variance (ANOVA).
747 § Degradability parameters determined by *in situ* incubation in the rumen, using the model of Ørskov
748 and McDonald (1979) ED = a+b[c/(c+k)] where a = rapidly soluble material; b = non-soluble but
749 degradable material; c = rate f degradation of b; and k = an assumed outflow rate of 0.08/hr.
750 ¶ Degradability parameters determined by *in situ* incubation in the rumen, using the Feed Into Milk
751 (FIM) Consortium (2004) model $ED_{FIM} = (0.9s/(0.9+k_{liq}))+(b_D c/(c+k_{liq}))+(bc/(c+k_f))$ where s = soluble
752 proportion, k_{liq} = fractional outflow rate of the liquid pool (0.075/hr), b_D = degradable small particle
753 proportion, b = degradable large particle proportion, c = fractional degradation rate of b, and k_f is the
754 fractional outflow rate of the large particle pool (0.045/hr).
755 ‖ TVC is the sum of acetic, butyric, lactic, propionic and valeric acids plus ethanol and propanol.
756 ¥ TVFA is calculated as for TVC but excluding ethanol and propanol.
757 [a,b] Values within a row with different superscripts differ significantly at *P*<0.05.
758

759　**Table 3** The results of a validation in which 94 grass-clover silages were used to test
760　the prediction accuracy of grass-based NIRS equations for chemical composition
761　when used on clover-containing samples (in g/kg DM unless otherwise stated).

Item	Measured mean	Predicted mean	Bias†	P value‡	r^2§	Relative RMSEP, %¶	RPD
ADF	336	292	43.0	0.001	0.61	17.6	0.87
aNDF	448	438	9.65	0.209	0.56	8.9	1.45
Ash	96.6	91.6	5.0	0.033	0.52	16.5	1.32
CP	138	126	12.4	0.005	0.75	17.1	1.58
DOMD	632	645	-13.0	0.195	0.64	6.7	1.56
EDN$_{FIM}$, g/kg N ‖	623	762	139	0.001	0.01	24.5	0.48
EE	26.5	30.1	-3.6	0.001	0.25	25.9	0.89
LA, g/kg	13.4	14.3	-0.9	0.622	0.48	70.6	1.22
MDM$_{FIM}$‖	129	146	-17	0.003	0.01	38.1	0.39
N, g/kg	8.8	8.1	0.7	0.187	0.86	19.4	2.35
NH$_3$-N, g/kgDM*100	62.5	85.2	-22.6	0.001	0.34	45.0	0.89
pH	4.6	4.8	-0.1	0.122	0.48	10.8	1.21
TVC, g/kg ¥	23.4	30.2	-6.8	0.001	0.52	39.3	1.15
TVFA, g/kg #	19.6	25.6	-5.9	0.001	0.51	43.5	1.17
VCODM, g/kg	397	409	-12.0	0.558	0.98	6.6	4.92
WSC	41.8	48.8	-7.0	0.113	0.40	58.4	1.25

762　DOMD = digestible organic matter in total dry matter; EDN = effective degradable nitrogen; EE = ether
763　extract; LA = lactic acid; MDM = microbial dry matter yield; NH$_3$-N = ammonia nitrogen; RMSEP = root
764　mean standard error of prediction; RPD = ratio of standard deviation of the measured population to
765　the standard error of prediction; TVC = total volatile content, TVFA = total volatile fatty acids; WSC =
766　water soluble carbohydrates.

767　† Bias is the measured mean minus the predicted mean, therefore minus values indicate over-
768　estimation and positive values indicate under-estimation of the equation.
769　‡ The probability of there being no significant difference between the measured mean and the
770　predicted mean analysed using student's t-test.
771　§ Simple linear regression coefficient
772　¶ Root mean square error of prediction presented as a percentage of the measured mean for
773　standardisation
774　‖ Degradability parameters determined by *in situ* incubation in the rumen, using the Feed Into Milk
775　(FIM) Consortium (2004) model ED$_{FIM}$ = (0.9s/(0.9+k$_{liq}$))+(b$_D$ c/(c+k$_{liq}$))+(bc/(c+k$_f$)) where s = soluble
776　proportion, k$_{liq}$ = fractional outflow rate of the liquid pool (0.075/hr), b$_D$ = degradable small particle
777　proportion, b = degradable large particle proportion, c = fractional degradation rate of b, and k$_f$ is the
778　fractional outflow rate of the large particle pool (0.045/hr).
779　¥ TVFA is calculated as for TVC but excluding ethanol and propanol.
780　# TVC is the sum of acetic, butyric, lactic, propionic and valeric acids plus ethanol and propanol.
781
782

783 **Table 4** Indicators of calibration strength and prediction accuracy using cross-
784 validation for a range of optimised new NIRS equations calibrated on spectra from
785 95 diverse grass-clover silages.

Item[†]	n[‡]	SEC	r^2[§]	Relative SECV, % [¶]
ADF (DM)	183	13.4	0.90	4.49
ADF (Fresh)	181	6.22	0.98	5.71
Alcohol ‖	178	1.08	0.83	37.1
aNDF (DM)	183	18.5	0.89	4.80
aNDF (Fresh)	182	7.79	0.98	5.26
Ash (DM)	185	10.4	0.70	12.5
Ash (Fresh)	179	3.30	0.91	11.1
DOMD	172	3.10	0.83	5.47
EDDM$_{0.08}$ ¥	174	2.15	0.88	5.28
EDN$_{0.08}$ ¥	174	3.93	0.79	7.03
EE (DM)	180	2.67	0.83	11.2
EE (Fresh)	179	0.94	0.90	10.8
LA	173	4.76	0.81	41.5
N	180	0.65	0.97	8.33
NH$_3$-N	176	0.01	0.88	18.8
pH	180	0.16	0.93	4.18
TVC #	185	5.39	0.82	27.9
TVFA ††	183	5.17	0.81	31.8
VCODM	181	7.17	1.00	2.10
WSC (DM)	180	10.1	0.92	31.4
WSC (Fresh)	181	4.62	0.93	29.6

786 EDDM = effective degradability of dry matter; EDN = effective degradable nitrogen; DOMD =
787 digestible organic matter in total dry matter; EE = ether extract; LA = lactic acid; NH$_3$-N = ammonia
788 nitrogen; SEC = standard error of calibration; SECV = standard error of cross-validation; TVC = total
789 volatile content, TVFA = total volatile fatty acids; WSC = water soluble carbohydrates.

790 † For variables that are measured on a dry sample (Ash, ADF, aNDF and WSC) two equations were
791 produced, one predicting on a fresh basis and one on a DM basis.
792 ‡ The number of spectra that were included in the prediction equation.
793 § Simple linear regression coefficient
794 ¶ Standard error of cross validation presented as a percentage of the measured mean for
795 standardisation
796 ‖ Alcohol is the sum of ethanol and propanol
797 ¥ Degradability parameters determined by *in situ* incubation in the rumen, using the model of Ørskov
798 and McDonald (1979) where a = rapidly soluble material; b = non-soluble but degradable material; c =
799 rate f degradation of b; effective degradability = a+b[c/(c+k)] where k = an assumed outflow rate of
800 0.08/hr.
801 # TVC is the sum of acetic, butyric, lactic, propionic and valeric acids plus ethanol and propanol.
802 †† TVFA is calculated as for TVC but excluding ethanol and propanol.
803

Table 5 The results of a validation in which 10 grass-clover silages were used to test the prediction accuracy of new clover/grass-based NIRS equations generated from the spectra of 85 other grass-clover silages (in g/kg DM unless otherwise stated). Industry standardised data pre-treatment methods were used (1st derivative and SNVD scatter correction) in the calibration of these equations.

Item	Measured mean	Predicted mean	Bias†	r^2‡	Relative RMSEP, % §	RPD
ADF	343	352	-9.28	0.93	7.31	2.94
ADF, g/kg	162	159	1.93	0.99	6.17	8.66
Alcohol, g/kg ¶	3.5	4.9	-1.41	0.19	75.7	0.93
aNDF	459	479	-20.6	0.85	8.02	2.15
aNDF, g/kg	212	215	-2.54	0.98	5.81	7.87
Ash	88.5	87.0	1.49	0.26	16.5	1.19
Ash, g/kg	39.2	40.6	-1.38	0.73	18.8	1.82
CP	120	120	-0.82	0.74	12.3	1.92
DOMD	637	637	-0.2	0.71	9.18	1.76
EDDM$_{0.08}$ ‖	452	438	14.7	0.68	16.3	1.58
EDN$_{0.08}$, g/kg N ‖	600	588	11.9	0.92	6.74	3.43
EE	22.8	21.6	1.22	0.46	23.2	0.95
EE, g/kg	9.9	9.8	0.15	0.67	16.9	1.82
LA, g/kg	15.1	12.7	2.41	0.72	49.0	1.74
N, g/kg	8.9	9.0	-0.02	0.92	13.4	3.76
NH$_3$-N, g/kg DM*100	104	112	-8.01	0.64	25.0	1.64
pH	4.7	4.6	0.06	0.70	10.8	1.86
TVC, g/kg ¥	22.9	21.7	1.20	0.76	33.5	1.85
TVFA, g/kg #	19.5	18.4	1.03	0.73	37.0	1.86
VCODM, g/kg	451	448	2.87	0.99	2.46	14.2
WSC	51.7	58.7	-7.01	0.69	35.3	1.76
WSC, g/kg	22.4	27.2	-4.84	0.69	44.4	1.57

DOMD = digestible organic matter in total dry matter; EDN = effective degradable nitrogen; EE = ether extract; LA = lactic acid; NH$_3$-N = ammonia nitrogen; RMSEP = root mean standard error of prediction; RPD = ratio of standard deviation of the measured population to the standard error of prediction; SNVD, standard normal variate de-trending; TVC = total volatile content, TVFA = total volatile fatty acids; WSC = water soluble carbohydrates.

† Bias is the measured mean minus the predicted mean, therefore minus values indicate over-estimation and positive values indicate under-estimation of the equation.

‡ Simple linear regression coefficient

§ Root mean square error of prediction presented as a percentage of the measured mean for standardisation

¶ Alcohol is the sum of ethanol and propanol

‖ Degradability parameters determined by *in situ* incubation in the rumen, using the model of Ørskov and McDonald (1979) where a = rapidly soluble material; b = non-soluble but degradable material; c = rate f degradation of b; effective degradability = a+b[c/(c+k)] where k = an assumed outflow rate of 0.08/hr.

¥ TVC is the sum of acetic, butyric, lactic, propionic and valeric acids plus ethanol and propanol.

TVFA is calculated as for TVC but excluding ethanol and propanol.

827 **Figure captions**

828 **Figure 1** The relationship between predicted and measured values where 94 grass-
829 clover silages were utilised to assess prediction accuracy of grass-based near infra-
830 red reflectance spectrometry (NIRS) equations for 15 chemical components when
831 used on clover-containing samples. Graphs for each chemical component show a line
832 of parity. VCODM = volatile corrected oven dry matter; TVFA = total volatile fatty acids;
833 TVC = total volatile content; DOMD = digestible organic matter in total dry matter; ADF
834 = acid detergent fibre; aNDF neutral detergent fibre; MDM = microbial dry matter yield;
835 NH_3-N = ammonia nitrogen; EDN = effective degradable nitrogen; WSC = water
836 soluble carbohydrate.

837 **Figure 2** The relationship between bias and sample clover concentration in a
838 validation test where 94 grass-clover silages were utilised to assess prediction
839 accuracy of grass-based near infra-red reflectance spectroscopy (NIRS) equations
840 for a) crude protein b) effective degradable nitrogen (EDN_{FIM}) and c) calculated
841 effective rumen degradable protein (eRPD) concentration (eRDP = CP * (0.8 *
842 EDN_{FIM})). Linear lines of best fit are shown for measured (——) and NIRS predicted
843 (— —) data.
844