Use of principal component analysis to classify forages and predict their calculated energy content

A. Gallo, M. Moschini, C. Cerioli and F. Masoero

Feed & Food Science and Nutrition Institute, Faculty of Agriculture, Università Cattolica del Sacro Cuore, Via Emilia Parmense 84, 29100 Piacenza, Italy

(Received 11 April 2012; Accepted 16 November 2012; First published online 9 January 2013)

A set of 180 forages (47 alfalfa hays, 26 grass hays, 52 corn silages, 35 small grain silages and 20 sorghum silages) were randomly collected from different locations of the Po Valley (Northern Italy) from 2009 to 2010. The forages were characterised for chemical composition (11 parameters), NDF digestibility (five parameters) and net energy for lactation (NE\textsubscript{L}). The latter was calculated according to the two approaches adopted by the 2001 Nutrient Research Council and based on chemical parameters either alone (NE\textsubscript{L,3x-Lig}) or in combination with 48 h NDF degradability in the rumen (NE\textsubscript{L,3x-48h}). Thereafter, a principal component analysis (PCA) was used to define forage populations and limit the number of variables to those useful for obtaining a rapid forage quality evaluation on the basis of the calculated NE\textsubscript{L} content of forages. The PCA identified three forage populations: corn silage, alfalfa hay and a generic population of so-called ‘grasses’, consisting of grass hays, small grain and sorghum silages. This differentiation was also confirmed by a cluster analysis. The first three principal components (PC) together explained 79.9% of the total variation. PC1 was mainly associated with protein fractions, ether extract and lignin, PC2 with ash, starch, NDF and indigestible NDF (iNDF) and PC3 with NDF digestibility. Moreover, PC2 was highly correlated to both NE\textsubscript{L,3x-Lig} (r = 0.84) and NE\textsubscript{L,3x-48h} (r = 0.94). Subsequently, forage-based scores (FS) were calculated by multiplying the original standardised variables of ash, starch, NDF and iNDF with the scoring factors obtained from PCA (0.112, 20.141, 0.227 and 0.170, respectively). The FS showed a high determination coefficient for both NE\textsubscript{L,3x-Lig} (R\textsuperscript{2} = 0.86) and NE\textsubscript{L,3x-48h} (R\textsuperscript{2} = 0.73). These results indicate that PCA enables the distinction of different forage classes and appropriate prediction of the energy value on the basis of a reduced number of parameters. With respect to the rumen in situ parameters, iNDF was found to be more powerful at discriminating forage quality compared with NDF digestibility at different rumen incubation times or rates of NDF digestion.

Keywords: nutritive value, net energy for lactation, forage, multivariate principal component analysis

Implications

An accurate evaluation of the forage energy value is crucial to formulate diets that meet the energy requirements of ruminants, to maximise animal performances and reduce nutrient losses into the environment. Dynamic models to evaluate the energy value of forages require a large number of variables, some of them being poorly repeatable, time-consuming and expensive. The principal component analysis (PCA) allows to convert the original variables into a few latent components without much loss of information. In the current paper, the PCA was used to cluster forages in specific populations and predict their calculated energy content.

Introduction

Nowadays, more than 50% of dairy cows, producing about 70% of Italy’s milk, are in the Po Valley of Northern Italy. Whole-plant corn, alfalfa and cool-season crops, either as silage or hay, are among the main forages fed to dairy cattle in this area and represent from 30% to 80% of the dietary dry matter (DM). A multi-year rotation consisting of alfalfa, corn and wheat crops is a frequent farming system. However, annual double crops, such as Italian ryegrass–corn or barley/triticale–sorghum crops, represent opportunities for maximising forage yield.

Forages are extremely variable in chemical composition, digestibility and intake potential (Oba and Allen, 1999; Cherney, 2000). Therefore, an accurate and rapid estimate of these parameters is critical to formulate diets that meet animal energy requirements (De Boever et al., 1996; Weiss, 1998; Aguiar et al., 2011). Historically, the energy content of forages was estimated using empirical single-component equations based on specific nutrients such as NDF, ADF or crude fibre (Mertens, 1992; Rotz et al., 1999; Ferreira and Mertens, 2005). A summative approach was adopted by dynamic nutritional models (Sniffen et al., 1992;
National Research Council (NRC), 2001; Robinson et al., 2004) to estimate the net energy for lactation (NE_L) content of forage on the basis of their total digestible nutrient (TDN) values and considering the level of intake of animals (Weiss et al., 1992; Weiss, 1998). The empirical and summative approaches differ mainly in the number of parameters required to estimate NE_L. The former is based on a single laboratory determination and specific equations are used for each forage type, resulting in different regression coefficients for the same predictor (Mertens, 1992). In contrast, in the summative approach, a unique equation is used to estimate the NE_L of different feeds, but a large number of parameters are needed. A major disadvantage of complex models is the large number of laboratory parameters, with some being poorly repeatable, time-consuming and expensive (Cherney, 2000; Spanghero et al., 2010). In addition to chemical parameters, biological parameters such as in situ rumen NDF degradation characteristics have been proposed to properly evaluate forages used in dairy cow diets (Van Amburgh et al., 2003; Robinson et al., 2004; Krämer et al., 2012). In particular, it has been proved that biological measurements can be useful to improve the prediction of in vivo nutrient digestibility and the energy values of forages (De Boever et al., 1996; Robinson et al., 2004).

Therefore, a viable reduction process of parameters, at the same time capable of classifying forages and properly predicting their energy content, could be useful to address this issue (Aguiar et al., 2011). However, many parameters in a data set with a large number of chemical/biological measurements are correlated, and the use of correlated parameters in a prediction equation often results in a non-robust model. With principal component analysis (PCA), it is possible to convert the original variables into a small number of principal components (PC), which are not correlated (O’Rourke et al., 2005; Stevens, 2009).

The aims of the current experiment were to use PCA, on the one hand, to characterise forage populations and, on the other, to reduce the number of variables to those useful for obtaining a rapid and accurate prediction of the calculated energy value.

Material and methods

A set of 180 forages were randomly collected from different locations of the Po Valley, Northern Italy, in the 2009 to 2010 harvest seasons. Samples collected were: 47 alfalfa (including first alfalfa cutting) hays, 26 grass hays (i.e. cool-season annual and perennial grass forages), 52 corn silages, 35 small grain silages and 20 forage sorghum silages.

Sampling procedure

The forages were sampled by trained people. In particular, hays from small square, large square or large round bales and silages from wrapped bales were identified and collected from at least three bales randomly selected.

The silages in the bunkers were sampled at least 3 weeks after ensiling and from different sites (i.e. bottom, middle and top) of freshly cut face. Samples were taken with a cylindrical probe (35 × 5 cm; Quas Franco, Spilimbergo, Pordenone, Italy), and at least three probes were pooled in a single bag and mixed thoroughly. Samples were sent to a laboratory within 24 h for analysis.

Chemical analysis

The whole sample was dried (60°C in a ventilated oven until constant weight), ground through a 1-mm screen using a laboratory mill (Thomas-Wiley, Arthur H. Thomas Co., Philadelphia, PA, USA) and stored until analysis. All samples were assayed in duplicate. The DM was determined by gravimetric loss of free water by heating at 105°C for 3 h (Association of Official Analytical Chemists (AOAC), 1995, method 945.15); ash was determined as gravimetric residue after incineration at 550°C for 2 h (AOAC, 1995, method 942.05) and ether extract (EE) was obtained following the method 920.29 of AOAC (1995). The CP (N × 6.25) was determined using the Kjeldahl method (AOAC, 1995, method 984.13). The soluble fraction of CP (expressed on a CP basis) was determined according to Licitra et al. (1996). The NDF, ADF and ADL were determined using the Ankom® Fiber Analyzer (Ankom Technology Corporation, Fairport, NY, USA). The NDF analyses utilised a neutral detergent solution containing sodium sulphite and a heat stable amylase (activity = 17 400 Liquefon units/ml, Ankom Technology). The NDF, ADF and ADL contents were corrected for the residual ash content. The neutral (i.e. NDICP) and acid (i.e. ADICP) detergent-insoluble CP fractions were determined as residual nitrogen of Ankom fibre bags (Van Soest et al., 1991). Starch was measured by polarimetry (Polax 2L, Atago®, Tokyo, Japan) according to European Community official method of analysis (EC, 1999). The non-fibre carbohydrates (NFC) were calculated as: 100 − (CP + ash + EE + (NDF − NDICP)).

In situ digestible and indigestible NDF (iNDF) determinations

The digestibility of NDF (dNDF) was measured in situ by incubating nylon bags in the rumen of two cannulated dry cows for 12, 24, 48 and 288 h (Spanghero et al., 2010). Animal care was in accordance with the EC Council Directive guidelines for animals used for experimental and other scientific purposes (EC, 1986). The cows were daily fed 10 kg DM of a total mixed ration daily consisting of alfalfa hay, ryegrass hay, corn silage and concentrates (i.e. 300, 300, 300 and 100 g/kg DM, respectively) in two portions at 0800 and 1800 h. The diet contained 120 g CP and 550 g NDF/kg DM.

The rate of NDF digestion (kd, h⁻¹) was calculated from the undigestible fractions at different time points using a first-order model (Mertens, 1993b) and assuming a constant lag time of 3 h (Van Amburgh et al., 2003). The undigestible fraction after 288 h of incubation was considered as the iNDF.
Net energy calculations
The energy content of forages was calculated using equations (1)–(4) of NRC (2001) for the truly digestible (td, g/100 g DM) NFC, tdCP, td fatty acid (FA) and tdNDF determinations. The FA content of forages can be estimated as (EE − 1):

\[ \text{tdNFC} = 0.98 \times (100 - ((\text{NDF} - \text{NDICP}) + \text{CP} + \text{EE} + \text{Ash})) \times \text{processing adjusting factor} \]

(1.00 for all forages, with the exception of 0.94 used for corn silages) \hspace{1cm} (1)

\[ \text{tdCP} = \text{CP} \times e^{-1.2 \times (\text{ADICP/CP})} \]

(2)

\[ \text{tdFA} = \text{FA}(\text{if EE < 1, then FA = 0}) \]

(3)

\[ \text{tdNDF} = 0.75 \times ((\text{NDF} - \text{NDICP}) - \text{ADL}) \times (1 - (\text{ADL}/(\text{NDF} - \text{NDICP})))^{0.667} \] \hspace{1cm} or tdNDF = 48 hNDF.

(4)

Two values for digestible energy at maintenance (DE1x) were estimated for each tested forage using equation (5) of NRC (2001): the first (i.e. DE1x-Lig) was based only on chemical parameters, whereas the second (i.e. DE1x-48 h) was based on both chemical parameters and dNDF evaluated in situ at 48 h incubation.

\[ (\text{DE1x}, \text{ Mcal/kg}) = (\text{tdNFC}/100) \times 4.2 + (\text{tdCP}/100) \times 5.6 + (\text{tdFA}/100) \times 9.4 + (\text{tdNDF}/100) \times 4.2 - 0.3 \]

(5)

Thereafter, applying equation (6) (i.e. discount for diet containing 74% of TDNx and intake level of three times maintenance level of intake), equations (7)–(9), two values for NEi three times maintenance (i.e. NE\text{Ei}\text{Lig} and NE\text{Ei}\text{Lig-48h}) were estimated from DE\text{Ei}\text{Lig} and DE\text{Ei}\text{Lig-48h}, respectively:

\[ \text{Discount} = ((\text{TDN}_{i,x} \text{ of diet} - ((0.18 \times \text{TDN}_{i,x} \text{ of diet}) - 10.3)) \times 2)/\text{TDN}_{i,x} \text{ of diet} \]

\hspace{1cm} (6)

\[ \text{Metabolizable energy at three times maintenance} \]

\[ (\text{ME}_{3x}, \text{ Mcal/kg}) = ((1.01 \times (\text{DE}_{1x} \times \text{discount}) - 0.45) + 0.0046 \times (\text{EE} - 3) \]

\hspace{1cm} (7)

if EE < 3, NE\text{Ei}\text{Lig} (Mcal/kg) = (0.703 \times \text{ME}_{3x}) - 0.19 \hspace{1cm} (8)

if EE > 3, NE\text{Ei}\text{Lig} (Mcal/kg) = (0.703 \times \text{ME}_{3x}) - 0.19 + (((0.097 \times \text{ME}_{3x} + 0.19)/97) \times (\text{EE} - 3)) \hspace{1cm} (9)

Statistical analysis
Data were analysed as a completely randomised design using the GLM procedure of Statistical Analytical System (SAS, 2003) and according to the model described below

\[ Y_{ij} = \mu + \alpha_{i} + e_{ij} \]

where \( Y_{ij} \) is the dependent variable on the \( j \)th subject assigned to forage class \( i \) (i.e. alfalfa hay, grass hay, corn silage, small grain silage and sorghum silage), \( \mu \) is the overall mean, \( \alpha_{i} \) is the fixed effect of forage class \( (i = 1 \text{ to } 5) \) and \( e_{ij} \) is the residual error. The minimum significant difference (MSD) was generated from Tukey's test at \( P < 0.05 \) and it was used for comparison among means.

A PCA was performed using the FACTOR procedure of SAS (2003). The PRIN method with Kaiser's criterion (i.e. eigenvalue >1.0, Stevens, 2009) and the orthogonal Varimax rotation (ROTATE option) were used to extract latent constructs. This procedure generated loading (correlations between common latent factors and original variables) vectors. The forages were characterised for their position on the first two rotated extracted components. Then, a hierarchical cluster analysis was performed using the unweighted pair group mean with the arithmetic averages (UPGMA) method by the CLUSTER procedure of SAS (2003) to confirm population differentiation. The CORR procedure of SAS (2003) was used to study simple relationships between the rotated extracted components and the two NE\text{i} values (i.e. NE\text{Ei}\text{Lig} and NE\text{Ei}\text{Lig-48h}). The variables with loading vectors of more than 0.50 were retained and interpreted to obtain a forage subject-based score (FS) characterising each tested forage. In particular, the threshold value to retain variables was obtained by multiplying the critical value of a correlation coefficient at \( \alpha = 0.01 \) for a two-tailed test (i.e. 0.2) by 2.5 (Stevens, 2009). The NE\text{i} values were regressed on FS with the REG procedure of SAS (2003).

Results
Chemical composition, NDF digestibility, iNDF and energy evaluations of forages
The range in chemical composition of the analysed samples could be considered typical for each forage class (Table 1). In particular, the average forage CP contents ranged from 76 in corn silages to 166 g/kg DM in alfalfa hays, with the soluble fraction higher in silages than in hays (MSD = 67 g/kg CP, \( P < 0.05 \)). The average ash content was over 80 g/kg DM for all forages, except for corn silages. The ash, EE and CP intra-forage coefficients of variation (CVs) were generally lower than 25%, except for the CVs of EE in small grain (i.e. 32%) and sorghum silages (i.e. 33%).

As expected, grass hays, small grain and sorghum silages were characterised by higher (MSD = 42 g/kg DM, \( P < 0.05 \)) NDF mean values than corn silages and alfalfa hays. The corn silages had the lowest ADF and ADL mean values (MSD = 30 and 10 g/kg DM, \( P < 0.05 \)). The CV associated with NDICP ranged from 25% to 56%, respectively, for
Table 1  Chemical parameters of alfalfa hays (n = 47), grass hays (n = 26), corn silage (n = 52), small grain silages (n = 35) and sorghum silages (n = 20)

<table>
<thead>
<tr>
<th></th>
<th>Alfalfa hays</th>
<th>Grass hays</th>
<th>Corn silages</th>
<th>Small grain silages</th>
<th>Sorghum silages</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean ± s.d.</strong></td>
<td>908 ± 31</td>
<td>850 ± 39</td>
<td>340 ± 33</td>
<td>340 ± 33</td>
<td>295 ± 58</td>
</tr>
<tr>
<td><strong>Mean ± s.d.</strong></td>
<td>288 ± 43</td>
<td>343 ± 39</td>
<td>76 ± 7</td>
<td>43 ± 46</td>
<td>218 ± 54</td>
</tr>
<tr>
<td><strong>Range</strong></td>
<td>(176–379)</td>
<td>(30–55)</td>
<td>(57–94)</td>
<td>(0–151)</td>
<td>(0–228)</td>
</tr>
<tr>
<td><strong>Mean ± s.d.</strong></td>
<td>82 ± 14</td>
<td>34 ± 4</td>
<td>29 ± 3</td>
<td>22 ± 7</td>
<td>21 ± 7</td>
</tr>
<tr>
<td><strong>Mean ± s.d.</strong></td>
<td>80 ± 19</td>
<td>51 ± 20</td>
<td>82 ± 20</td>
<td>34 ± 20</td>
<td>86 ± 15</td>
</tr>
<tr>
<td><strong>Mean ± s.d.</strong></td>
<td>380 ± 36</td>
<td>303 ± 48</td>
<td>234 ± 48</td>
<td>58 ± 9</td>
<td>40 ± 34</td>
</tr>
<tr>
<td><strong>Mean ± s.d.</strong></td>
<td>10 ± 12</td>
<td>51 ± 35</td>
<td>21 ± 13</td>
<td>13 ± 5</td>
<td>37 ± 13</td>
</tr>
<tr>
<td><strong>Range</strong></td>
<td>(7–38)</td>
<td>(9–51)</td>
<td>(8–24)</td>
<td>(3–13)</td>
<td>(6–17)</td>
</tr>
<tr>
<td><strong>Mean ± s.d.</strong></td>
<td>14 ± 4</td>
<td>306 ± 48</td>
<td>183 ± 68</td>
<td>137 ± 55</td>
<td>147 ± 34</td>
</tr>
</tbody>
</table>

**Chemical parameters**
- **DM (g/kg)**
- **Starch (g/kg DM)**
- **Ash (g/kg DM)**
- **EE (g/kg DM)**
- **CP (g/kg DM)**
- **Soluble CP (g/kg CP)**
- **NDF (g/kg DM)**
- **ADF (g/kg DM)**
- **ADL (g/kg DM)**
- **NDICP (g/kg DM)**
- **NFC (g/kg DM)

**DM** = dry matter; **nd** = not determined; **EE** = ether extract; **ADICP** = acid detergent insoluble CP; **NDICP** = neutral detergent insoluble CP; **NFC** = non-fibre carbohydrate.

P-value of the model (one-way ANOVA and post hoc Tukey's test)

**Table 1**

In Table 1, the correlation coefficients between the two NE-Lig values and the three main PCs are presented. The data set was considered suitable for PCA (Cerny and Kaiser, 1977). As a matter of fact, three components, with the highest coefficients for PC2, were loaded on PC2. A two-dimensional plot characterised each sample for the three main PCs, being 31.9%, 26.2%, 23.0% and 9.1% for 12, 24, 48 and 288 h dNDF, respectively. In addition, the CVs calculated for 48 h dNDF reduced numerically the NE-Lig-48 h in small grain and sorghum silages, whereas the lowest average values were obtained for small grain silages (i.e. 4.0%) and sorghum silages (i.e. 11%). The NFC contents differed among forages (MSD = 32 ± kg DM), from 27% to 33%, respectively. The NFC contents differed among forages (MSD = 32 ± kg DM), from 27% to 33%, respectively. The NFC contents differed among forages (MSD = 32 ± kg DM), from 27% to 33%, respectively. The NFC contents differed among forages (MSD = 32 ± kg DM), from 27% to 33%, respectively. The NFC contents differed among forages (MSD = 32 ± kg DM), from 27% to 33%, respectively.
Similarly, PC2 was highly correlated with both NE\textsubscript{L3x-Lig} and NE\textsubscript{L3x-48h} when analysis was carried out within forage populations. However, significant (P < 0.05) correlations were also observed between NE\textsubscript{L3x-48h} and PC1, with higher correlation coefficients observed for alfalfa and grasses.

### Table 2

Digestibility parameters and energy evaluations of alfalfa hays (n = 47), grass hays (n = 26), corn silages (n = 52), small grain silages (n = 35) and sorghum silages (n = 20)

<table>
<thead>
<tr>
<th>Items</th>
<th>Alfalfa hays</th>
<th>Grass hays</th>
<th>Corn silages</th>
<th>Small grain silages</th>
<th>Sorghum silages</th>
<th>(\sqrt{\text{MSE}})</th>
<th>MSD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean ± s.d.</td>
<td>Range</td>
<td>Mean ± s.d.</td>
<td>Range</td>
<td>Mean ± s.d.</td>
<td>Range</td>
<td></td>
</tr>
<tr>
<td>Digestibility parameters(^a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 h dNDF (g/kg DM)</td>
<td>109 ± 33</td>
<td>(44–171)</td>
<td>147 ± 33</td>
<td>(95–246)</td>
<td>87 ± 29</td>
<td>(40–160)</td>
<td></td>
</tr>
<tr>
<td>24 h dNDF (g/kg DM)</td>
<td>193 ± 33</td>
<td>(85–339)</td>
<td>262 ± 35</td>
<td>(184–340)</td>
<td>198 ± 61</td>
<td>(108–329)</td>
<td></td>
</tr>
<tr>
<td>48 h dNDF (g/kg DM)</td>
<td>228 ± 53</td>
<td>(128–358)</td>
<td>346 ± 38</td>
<td>(267–422)</td>
<td>285 ± 50</td>
<td>(202–399)</td>
<td></td>
</tr>
<tr>
<td>iNDF (g/kg DM)</td>
<td>271 ± 59</td>
<td>(156–404)</td>
<td>233 ± 47</td>
<td>(150–352)</td>
<td>121 ± 38</td>
<td>(41–200)</td>
<td></td>
</tr>
<tr>
<td>(kd) (h(^{-1}))</td>
<td>0.042 ± 0.013</td>
<td>(0.027–0.090)</td>
<td>0.033 ± 0.008</td>
<td>(0.021–0.048)</td>
<td>0.030 ± 0.008</td>
<td>(0.014–0.047)</td>
<td></td>
</tr>
</tbody>
</table>

### Energy evaluations\(^b\)

<table>
<thead>
<tr>
<th>NE\textsubscript{L3x-Lig} (Mcal/kg)</th>
<th>Alfalfa hays</th>
<th>Grass hays</th>
<th>Corn silages</th>
<th>Small grain silages</th>
<th>Sorghum silages</th>
<th>(\sqrt{\text{MSE}})</th>
<th>MSD</th>
</tr>
</thead>
<tbody>
<tr>
<td>12 h dNDF</td>
<td>1.14 ± 0.10</td>
<td>(0.92–1.37)</td>
<td>1.17 ± 0.11</td>
<td>(0.91–1.34)</td>
<td>1.43 ± 0.08</td>
<td>(1.00–1.61)</td>
<td></td>
</tr>
<tr>
<td>24 h dNDF</td>
<td>1.20 ± 0.17</td>
<td>(0.85–1.52)</td>
<td>1.48 ± 0.10</td>
<td>(1.16–1.71)</td>
<td>1.04 ± 0.12</td>
<td>(0.65–1.33)</td>
<td></td>
</tr>
<tr>
<td>48 h dNDF</td>
<td>1.23 ± 0.15</td>
<td>(0.80–1.50)</td>
<td>1.20 ± 0.17</td>
<td>(0.85–1.52)</td>
<td>1.48 ± 0.10</td>
<td>(1.16–1.71)</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\)\(\text{dNDF}\) = digestible NDF; DM = dry matter; \(\text{iNDF}\) = indigestible NDF; \(kd\) = rate of NDF degradability.  
\(^b\)P-value of the model (one-way ANOVA and post hoc Tukey's test), 0.05 for all parameters analysed. 
\(^c\)Net energy for lactation values estimated according to National Research Council (2001) and based either on chemical parameters (NE\textsubscript{L3x-Lig}) or both chemical parameters and 48 h dNDF (NE\textsubscript{L3x-48h}).

### Table 3

Loading vectors of original variables on Varimax rotated extracted PC as estimated by multivariate analysis

<table>
<thead>
<tr>
<th>PC1</th>
<th>PC2</th>
<th>PC3</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\text{NDICP})</td>
<td>0.87*</td>
<td>0.11</td>
</tr>
<tr>
<td>(\text{CP})</td>
<td>0.85*</td>
<td>0.02</td>
</tr>
<tr>
<td>(\text{ADL})</td>
<td>0.81*</td>
<td>0.32</td>
</tr>
<tr>
<td>(\text{Soluble CP})</td>
<td>0.81*</td>
<td>0.02</td>
</tr>
<tr>
<td>(\text{kd})</td>
<td>0.73*</td>
<td>0.41</td>
</tr>
<tr>
<td>(\text{ADICP})</td>
<td>0.71*</td>
<td>0.26</td>
</tr>
<tr>
<td>(\text{EE})</td>
<td>0.70*</td>
<td>0.35</td>
</tr>
<tr>
<td>(\text{ADF})</td>
<td>0.24</td>
<td>0.92*</td>
</tr>
<tr>
<td>(\text{NFC})</td>
<td>0.07</td>
<td>0.88*</td>
</tr>
<tr>
<td>(\text{NDF})</td>
<td>0.21</td>
<td>0.85*</td>
</tr>
<tr>
<td>(\text{Starch})</td>
<td>0.58</td>
<td>0.71*</td>
</tr>
<tr>
<td>(\text{iNDF})</td>
<td>0.48</td>
<td>0.66*</td>
</tr>
<tr>
<td>(\text{Ash})</td>
<td>0.52</td>
<td>0.65*</td>
</tr>
<tr>
<td>24 h dNDF</td>
<td>0.14</td>
<td>0.31</td>
</tr>
<tr>
<td>12 h dNDF</td>
<td>0.08</td>
<td>0.19</td>
</tr>
<tr>
<td>48 h dNDF</td>
<td>0.43</td>
<td>0.25</td>
</tr>
</tbody>
</table>

\(\text{PC}\) = principal components; \(\text{NDICP}\) = neutral detergent insoluble CP; \(\text{kd}\) = rate of NDF degradability; \(\text{ADICP}\) = acid detergent insoluble CP; \(\text{EE}\) = ether extract; \(\text{NFC}\) = non-fibre carbohydrate; \(\text{iNDF}\) = indigestible NDF; \(\text{dNDF}\) = digestible NDF.  
\(\ast\)Variables loaded on extracted components (i.e. loading vectors higher than 0.50).
$r = -0.65$ and $r = -0.50$, respectively) than for corn populations ($r = 0.35$). PC3 was related ($P < 0.05$) to NE$_{\text{L3x-Lig}}$ in the corn population and to NE$_{\text{L3x-48h}}$ in grasses.

The FS were calculated using loadings (Table 3) associated with the variables with the greatest weight on PC2, these being ash, starch, NDF and iNDF (scoring factors of 0.112, $-0.141$, 0.227 and 0.170, respectively).

The highest ($P < 0.05$) FS was obtained for corn silages (i.e. $0.75 \pm 0.21$), whereas the lowest ($P < 0.05$) was obtained for grass hays, sorghum and small grain silages (i.e. $-0.37 \pm 0.20$, $-0.36 \pm 0.17$ and $-0.34 \pm 0.20$, respectively). The alfalfa hays were characterised with an intermediate FS value (i.e. $-0.24 \pm 0.26$). The relationships between FS and NE$_L$ evaluations are shown in Figure 3. The highest coefficient of determination (i.e. $R^2 = 0.86$, $P < 0.05$) and the lowest standard errors of intercept (i.e. 0.0058) and slope (i.e. 0.0130) were obtained when EN$_{\text{L3x-Lig}}$ was regressed on FS. Moreover, FS accounted for 73% of the variation of EN$_{\text{L3x-48h}}$ when analysis was carried out on all data.

Figure 2  Dendrogram by cluster analysis showing the average distance between forages belonging to alfalfa, corn and grass populations (AA, CS and GF, respectively).

**Discussion**

Chemical composition, NDF digestibility, iNDF and energy evaluations of forages

In general, both mean and standard deviations of chemical parameters could be considered similar to those reported by the NRC (2001) for specific forage classes. However, some differences related mainly to structural carbohydrates (i.e. NDF, ADF and ADL) and CP contents were observed, suggesting that the tested forages were presumably harvested from a mid to an advanced stage of maturity. These differences were not observed for corn silages, having a numerically lower ADF when compared with NRC (2001) ‘normal’ corn silages.

The 48 h in situ assay is the time suggested by NRC (2001) as an alternative to calculate the tdNDF value and, being less affected by both lag time and rate of degradation than earlier incubation times (Mertens, 1992; Spanghero et al., 2010), it could be used to compare different forages. However, some authors evaluated the opportunity to reduce
the length of the rumen incubation to 30, 24 h or even shorter (Ferreira and Mertens, 2005; Spanghero et al., 2010; Tagliapietra et al., 2011). Despite an increased data variability (Spanghero et al., 2010), a reduced rumen incubation time could better describe the real NDF degradation dynamic in the rumen of high-producing dairy cows at high levels of intake (Goës and Combs, 2009; Tagliapietra et al., 2011).

The forages were also characterised by iNDF after 288 h of rumen incubation. Its importance was highlighted by Mertens (1993a), suggesting that ‘the estimation of indigestible fraction is not a mathematical or modelling contrivance, but is a critical biological principle’ needed to estimate the potentially available degradable NDF pool. A way to rapidly estimate the iNDF is based on the concept that lignin represents the first factor limiting NDF digestibility and it is highly correlated to the iNDF fraction (Weiss, 1998; Ferreira and Mertens, 2005; Vieira et al., 2012). Currently, the Cornell Net Carbohydrate and Protein System (CNCPS) nutritional model uses a lignin content that is 2.4 times lignin content of NDF to calculate the iNDF (Sniffen et al., 1992; Vieira et al., 2012). This simple equation was originally based on long-term (60 to 90 days) methane yield on some waste materials, and then tested on several forages by Traxler et al. (1998). As reported by these authors, the equation was able to correctly predict iNDF of the different forages.

Recently, some authors (Kraemer et al., 2010; Kraemer et al., 2012; Vieira et al., 2012) reported that the use of a constant coefficient could result in an inaccurate estimate of the iNDF among different forages. In particular, Kraemer et al. (2010) observed variable iNDF/ADL ratios among forages, ranging from 1.8 to 3.6 in grass hays and red clover, respectively.

Our results support the idea that the use of a fixed factor for different forages could result in deviating iNDF estimates.

In addition, the rate of NDF digestion (i.e. kd) in the rumen could be useful to discriminate between forage quality (Van Amburgh et al., 2003). A higher kd (P < 0.05) was observed in alfalfa hays compared with other forages, in accordance with Mertens (1993b) reporting that alfalfa had a more rapid rate of NDF digestion than other forages, but to a lower extent, showing a fibre digestion plateau at a lower level. For corn silages, Van Amburgh et al. (2003) reported the kd of NDF of 21 samples (three were brown midrib hybrids), estimated with multiple time point calculations (i.e. from 6 to 36 h of rumen incubation) and variable lag time, ranging from 0.035 to 0.071 h \(^{-1}\). In our model, the 48 h dNDF value was included and this value, being near the inflection point of the degradation curve in the slow pool of neutral detergent residue, could be associated with a reduction of kd (Mertens, 1993b; Ferreira and Mertens, 2005).

Referring to energy evaluations, different authors have debated on which of these approaches could be considered the most accurate (Robinson et al., 2004; Tagliapietra et al., 2011). Our data are inadequate to evaluate the methods of calculating NE\(_L\). However, we observed some discrepancies...
Nutritive and energy values of forages

dNDF to compute NE L calculation numerically increased the yield and lactation persistency of dairy cows (Macciotta 1996). More recently, some authors used multivariate analysis to evaluate the variability and range of variation of energy evaluation of the Graminaceae population between corn silage and cutting hays. PC2 seemed more selective in differentiating forages composed of either alfalfa or alfalfa/grass (i.e. first alfalfa observed, probably because the alfalfa population was identified by the first two constructs: corn silage, alfalfa hay and grasses). No additional populations were identified by PC3, suggesting that it was not as critical as the previous PCs in population differentiation.

Up to now, the multivariate analysis has had only limited use in agricultural research, particularly in feed evaluation. In the past, linear regression analysis based on both chemical and biological (i.e. in vitro rumen fluid or enzymatic digestibility) parameters was used to predict the energy value of feeds and forages (Givens et al., 1989; De Boever et al., 1996). More recently, some authors used multivariate analysis techniques to extract latent variables linked to the milk yield and lactation persistency of dairy cows (Macciotta et al., 2006), to study the influence of quality parameters on the methanogenic potential of different forages (Jayanegara et al., 2011) or the between-goat variability in feeding behaviour (Desnoyers et al., 2011).

The PCA allowed the characterisation of each sample for its position on extracted components, and the affiliation of single forage to a specific population was visually defined. Thus, in our database, three main populations were identified by the first two constructs: corn silage, alfalfa hay and ’grasses’, the latter made up of grass hays, small grain and sorghum silages. In particular, PC1 defined two populations: alfalfa and Graminaceae forages. However, a partial overlap between poor-quality alfalfa hays and grass hays was observed, probably because the alfalfa population was composed of either alfalfa or alfalfa/grass (i.e. first alfalfa cutting) hays. PC2 seemed more selective in differentiating the Graminaceae population between corn silage and grasses. No additional populations were identified by PC3, suggesting that it was not as critical as the previous PCs in population differentiation.

A cluster analysis was run on all forages to confirm the results obtained by PCA. As shown in Figure 2, three clusters were identified. In particular, all the corn silages were grouped into one cluster, whereas some forages classified as grasses (16.7%) were included in the alfalfa cluster, confirming the observations discussed above.

Another application of PCA analysis is to narrow down the number of variables to those mainly related to a specific parameter of interest (i.e. NE L content of forages). In this context, it should be mentioned that the computation of NE L values was done by using some parameters having an important weight on PCs. The PC2 was highly correlated to NE L and it included ash, NFC, starch, NDF, ADL and IDF. Even if other extracted factors explained a significant amount of total variation, their coefficients of correlation with NE L ranged from low to zero. Thus, only variables loading on this component were retained to obtain the FS of each tested forage. The NFC was removed by FS estimate because it is calculated from other analytical determinations (i.e. CP, EE, ash, NDF and NDICP). In addition, the remaining selected variables were correlated with each other (Supplementary Table S1) with very high coefficients of correlation measured for NDF, ADL and starch (r = 0.82 and r = −0.82, respectively; P < 0.05). Starch is routinely measured on corn silage and it could be useful as a discriminating measurement. Even if ash was highly correlated to starch, it was retained because it could be extremely valuable in specific forages because of its dilution effect on energy value (e.g. high soil contamination), whereas the NDF rather than ADL was retained in FS calculation to avoid an additional analytical (i.e. acid detergent treatment) step to obtain a value that is correlated to NDF (De Boever et al., 1996).

Therefore, the original parameters retained to calculate FS were ash, starch, NDF and IDF. Even if this approach may not be completely perfect as not all the original variables were used to compute the factor score (O’Rourke et al., 2005), the adopted solution is reasonable from a practical point of view too. In particular, the approach allowed for a limited number (i.e. three in hays or four in silages) of parameters used in FS calculation, thus resulting in a less expensive and time-consuming preliminary quality forage evaluation.

Both NE L evaluations (Y-axis) were related to FS (X-axis), with a higher relationship observed for NE L3x−48h (i.e. R² = −0.86, P < 0.05) than for NE L3x−lig (i.e. R² = −0.73, P < 0.05). In our condition, the high coefficients of determination supported the idea that the FS resulting from the PCA of the complete database was able to differentiate between and within forage classes. Lower coefficients of determination and higher regression bias were observed when the number of variables was reduced.

When relationships between FS and NE L evaluations were ascertained within populations (i.e. corn silage, alfalfa hay and grasses), coefficients of determination from 0.54 to 0.89 were measured (P < 0.05). However, coefficients of determination of 0.37 and 0.45 (P < 0.05) were calculated between NE L3x−48h, and FS in grasses and corn silage populations, respectively. A possible explanation could be related to the reduced number of samples and the close range of variation of measurements, as well as the variability measured in 48 h dNDF.
The results obtained through the PCA on the current database confirmed that the use of biological measurement other than chemical parameters could be useful to better discriminate between poor- and high-quality forages (Givens et al., 1989; De Boever et al., 1996; Weiss, 1998). In our experimental condition, the use of iNDF, instead of other biological measurements (i.e. 12, 24, 48 hNDF or kd), was critical for forage characterisation, suggesting that the extent of fibre digestion is useful for characterising forage quality (Mertens, 1993a; Weiss, 1998). However, the methods currently adopted to measure iNDF (i.e. 288 h of rumen incubation) are troublesome and rely on the use of fistulated animals (Krämer et al., 2012). Other methodological approaches, such as in vitro enzymatic or gas production techniques, could represent valid and attractive alternatives to the in situ determination of iNDF (De Boever et al., 1996; Aguiar et al., 2011). Alternatively, the development of a specific near-infrared spectroscopy (NIRS) calibration curve could be used to predict the iNDF of forages (Krämer et al., 2012). In particular, NIRS measurements have been successfully used to predict fibre digestibility, INDF, energy content and voluntary intake of forages (De Boever et al., 1996; Lundberg et al., 2004; Andueza et al., 2011; Krämer et al., 2012).

In conclusion, the PCA applied on the current data set allowed for the detection of three latent constructs associated with tested forages, such as chemical components (i.e. protein fractions, EE and lignin), carbohydrate and their indigestible fraction or NDF digestibility. Furthermore, this approach allowed for a development of a synthetic FS on the basis of a reduced number of variables (i.e. ash, starch, NDF and iNDF) capable of differentiating between high- and poor-quality forage both among and within forage populations. However, the findings obtained through PCA analysis should be reproduced over a larger database collected over a wider range of years to confirm actual observation.

Acknowledgements
The authors thank Sara Bruschi for her invaluable help in laboratory analysis. Financial support was provided by the AGROSCENARIO project ‘Scenari di adattamento dell’agricoltura italiana ai cambiamenti climatici’ of MIPAAF (Ministero delle politiche agricole alimentari e forestali – Italy) and by the PRIN project no. 2007PB8MJW1 of MIUR (Ministry of Education, University and Research – Italy).

Supplementary materials
For supplementary material referred to in this article, please visit http://dx.doi.org/doi:10.1017/S1751731112002467

References
Nutritive and energy values of forages


Robinson PH, Givens DI and Getachew G 2004. Evaluation of NRC, UC Davis and ADAS approaches to estimate the metabolizable energy values of feeds at maintenance energy intake from equations utilizing chemical assays and in vitro determinations. Animal Feed Science and Technology 114, 75–90.


