



RESEARCH ARTICLE OPEN ACCESS

Predicting Forage Nutritional Quality With Near-Infrared Spectroscopy

Alessandro Benelli¹ | Riccardo Primi² | Chiara Evangelista² | Raffaello Spina² | Marco Milanese¹ | Daniele Pietrucci¹ | Bruno Ronchi² | Umberto Bernabucci² | Roberto Moscetti¹

¹Department for Innovation in Biological, Agro-Food and Forest Systems, University of Tuscia, Viterbo, Lazio, Italy | ²Department of Agriculture and Forest Sciences, University of Tuscia, Viterbo, Lazio, Italy

Correspondence: Alessandro Benelli (alessandro.benelli@unitus.it) | Roberto Moscetti (rmoscetti@unitus.it)

Received: 15 May 2025 | **Revised:** 23 June 2025 | **Accepted:** 30 June 2025

Funding: The study was supported by the project AGRITECH of the European Union NextGenerationEU (grant number J83C22000830005).

Keywords: crude protein | feeding | livestock | neutral detergent fibre | pasture | PLS

ABSTRACT

The quality of green forage is crucial in pasture grazing, influencing both animal welfare, environmental sustainability, and production yield. Traditionally, the evaluation of forage composition requires time-consuming and costly chemical analysis. In this context, near-infrared spectroscopy (NIR) emerges as a promising alternative. This study adopted Fourier transform NIR (FT-NIR) spectroscopy to predict nutritional characteristics of green forages. A total of 324 samples were collected from pastures in central Italy. Partial least squares (PLS) regression models were then developed, applying variable selection methods to improve PLS model accuracy. The interval PLS (iPLS) variable selection method gave the best results for fresh forage, while the genetic algorithm (GA) performed best for dried samples. The best results from the PLS models were obtained for dry matter (DM) and crude protein (CP). The DM model for fresh forage yielded an R^2_p of 0.96 and an RMSEP of 2.95 g 100 g⁻¹ FW, while the CP model for dried forage yielded an R^2_p of 0.94 and an RMSEP of 1.84 g 100 g⁻¹ DW, with a normalised root-mean-square error of cross-validation (NRMSECV) of 3.8% and 5.6%, respectively. The results for the neutral detergent fibre (aNDF) were acceptable. NIR spectroscopy has proven to be a useful tool for assessing forage nutritional quality. Variable selection through iPLS also enabled the identification of “core” spectral regions for the development of compact and portable NIR sensors. Future research should further investigate sample preparation and moisture content effects and expand sampling to different geographical areas to enhance model robustness.

1 | Introduction

The quality of fresh forage, together with biomass availability, is a critical factor in pasture-based production systems, significantly influencing animal performance, stocking rates and methane emissions, as well as the health, well-being, and safety

and quality of the products they provide (Bewley 2010; FAO 2010; Shorten et al. 2019). Due to its essential role in sustaining productivity and profitability in meat and milk production, understanding the variability of nutritional quality is crucial for effective animal feeding management (Rivero and Lee 2022).

Abbreviations: ADF, acid detergent fibre; ADL, acid detergent lignin; aNDF, amylase-treated neutral detergent fibre; CA, crude ash; CF, crude fibre; CV, cross-validation; DM, dry matter; DW, dry weight; EE, ether extract (i.e., crude fat); EMSC, extended multiplicative scatter correction; FT-NIR, Fourier transform near-infrared; FW, fresh weight; GA, genetic algorithm; iPLS, interval partial least squares; NDVI, normalised difference vegetation index; NIR, near-infrared; NRMSECV, normalised root-mean-square error of cross-validation; PLS, partial least squares; R^2 , coefficient of determination; R^2_{CV} , coefficient of determination of cross-validation; R^2_p , coefficient of determination of prediction; RMSE, root-mean-square error; RMSECV, root-mean-square error of cross-validation; RMSEP, root-mean-square error of prediction; SEP, standard error of prediction; SNV, standard normal variate.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 The Author(s). *Journal of Sustainable Agriculture and Environment* published by Global Initiative of Sustainable Agriculture and Environment and John Wiley & Sons Australia, Ltd.

Globally, livestock production systems heavily depend on cultivated or naturally occurring pastures, as well as temporary and permanent meadows, which serve as the primary feed source for animals (Bernués et al. 2011). In 2022, the global area of permanent meadows and pastures was estimated at 3208 million hectares (approximately one-quarter of the Earth's land surface) supplemented by an additional 300 million hectares used as temporary meadows and pastures within crop rotation systems (FAO 2024). Despite the widespread use of pasture, ensuring consistent forage quantity and quality remains a complex challenge. Variability in forage quality arises from factors such as plant species composition, phenological stage, weather conditions, grazing intensity, and plant health, all of which affect the nutritional profile of the pasture (Wróbel et al. 2023). In fact, key nutritional components (e.g., crude protein, fibre, minerals, ash, digestibility of organic matter, sugars and metabolisable energy) are shaped by a range of environmental and management variables, including species diversity, plant maturity, soil characteristics, topography, and local climate conditions, among others (Holmes 2007; Pullanagari et al. 2012). Additionally, grazing management practices are a key issue for preserving pasture quality over time and maintaining its ecological functions (Teague and Kreuter 2020). Overgrazing can degrade plant diversity and soil health, reducing forage yield and nutritional value, whereas underutilisation may lead to biomass accumulation and reduced palatability (Dunn et al. 2010). Climate change further compounds these challenges, with shifting precipitation patterns, rising temperatures, and extreme weather events altering pasture productivity and composition. This variability highlights the need for meticulous and targeted pasture management to optimise animal nutrition. In this context, feedstuff analysis emerges as a crucial tool for assessing the quality of grazed forages. Therefore, it is fundamental to determine, ideally with a monthly sampling frequency, the nutritional value of pasture to implement alternative or complementary grazing/feeding practices (e.g., supplementation) to meet animal requirements (Martin et al. 2020), leading to both cost savings and a reduced environmental impact (Shalloo et al. 2018).

While conventional methods (i.e., wet chemistry) have been employed for determining the chemical composition of feedstuff, they pose significant limitations due to their cost, time-consuming nature and unsuitability for real-time analysis (González et al. 2018; Parrini et al. 2018). These factors not only increase production costs but also hinder the development of efficient quality monitoring systems. The slow turnaround time of traditional methods makes it difficult to obtain timely results that are necessary for dynamic adjustments to feeding practices in response to pasture variations in both space and time (Molle et al. 2008). Moreover, their reliance on expensive equipment, skilled personnel and chemical reagents contributes to environmental concerns (Samadi et al. 2018; Evangelista et al. 2021). Consequently, there is an increasing interest in developing alternative analytical methods that provide faster, more cost-effective, and environmentally sustainable solutions, in line with the broader objective of improving agricultural efficiency and sustainability. In this context, it is important to consider additional aspects. Specifically, the analytical technique must operate as a multi-analytical tool, allowing the simultaneous acquisition of multiple results in a single step

while minimising the use of wet-chemistry-based methods, except when their use is indispensable. In this context, while near-infrared (NIR) spectroscopy may not achieve exceptional accuracy across all measured parameters, it serves as an effective screening tool to identify samples that require further laboratory analysis. By distinguishing between samples that necessitate additional testing and those that do not, NIR can significantly reduce the overall laboratory workload.

From this perspective, NIR spectroscopy enables the rapid estimation of several feed quality parameters (Danieli et al. 2004; Cheli et al. 2012). The integration of Fourier transform near-infrared (FT-NIR) spectroscopy with chemometric models may offer high accuracy and precision in predicting the nutritional parameters of pastures (Iqbal et al. 2023). This approach allows for real-time feed analysis, providing data that are critical for adaptive pasture management. The implications of such advancements are substantial for improving animal nutrition, welfare, and overall production efficiency. By leveraging these tools, pasture-based systems can achieve greater sustainability and resilience, addressing the dual challenges of feeding a growing global population and mitigating the environmental impacts of livestock production (Dwivedi et al. 2023; Draghi et al. 2024).

Given the need for rapid and reliable fresh forage quality assessment tools, this study aims to evaluate the performance of FT-NIR spectroscopy in predicting key nutritional parameters of pasture samples. Specifically, we compare different pre-processing techniques and chemometric approaches to optimise model accuracy and robustness. Additionally, the present study aims to identify the most relevant spectral bands to simplify the model complexity as a crucial step for reducing the cost associated with the development of a field prototype. In summary, the ultimate goal is to facilitate the adoption of FT-NIR as a practical tool for routine forage analysis, enabling real-time pasture management and improved decision-making in livestock production systems.

2 | Materials and Methods

2.1 | Green Forage Sampling and Sample Preparation

The study was conducted on three beef cattle farms located in central Italy, Latium region (Figure 1), where pastures were composed of alfalfa meadows, permanent meadows, and mixed grasslands. The grass/legume ratio varied from 10:90 to 80:20, reflecting the diversity of forage composition across the sampled areas.

To ensure representative sampling across different biomass conditions, satellite-derived data were employed. Normalised difference vegetation index (NDVI) values were obtained from the most recent available Sentinel 2 A and 2B satellite images (10 × 10 m resolution), free of cloud cover before each sampling event, and processed using ArcGIS Pro 3.2.0. These NDVI values, ranging from 0.25 to 0.75, were used to stratify the study area and guide the selection of sampling locations, ensuring variability in biomass availability, phenological stages, and,

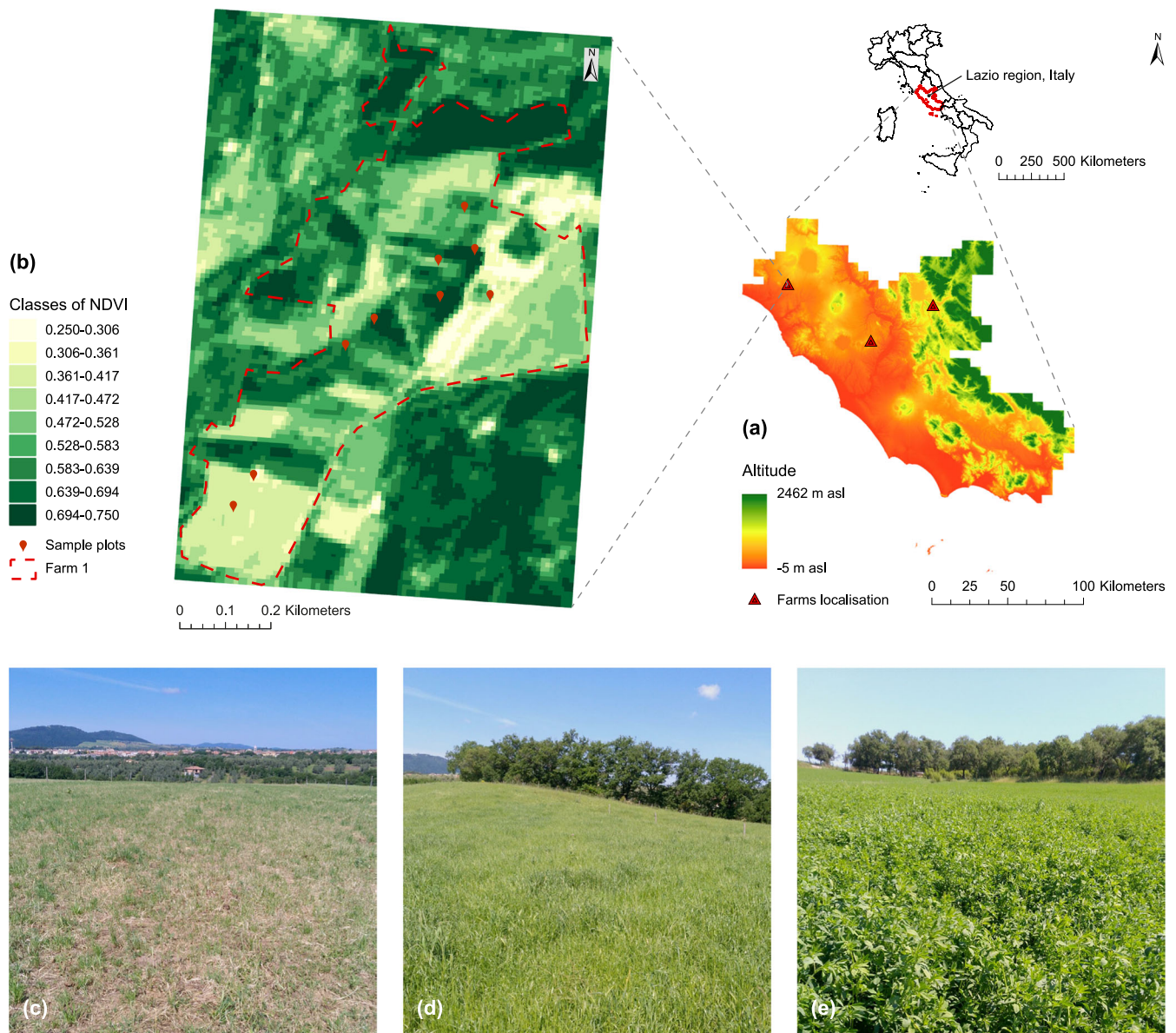


FIGURE 1 | (a) Altitude map. Data source: digital elevation model (DEM) at 20×20 m resolution, Lazio region. Available at: <https://geoportale.regione.lazio.it/>. (b) Example of stratified sampling methodology based on NDVI data. To ensure a representative range of biomass conditions, NDVI values were derived from Sentinel-2A or 2B images (10×10 m resolution) acquired on April 14, 2024. The study area was divided into nine equal-width NDVI intervals (0.25–0.75), and sampling locations were randomly selected within each interval. This approach optimises pasture quality assessment while minimising bias in vegetation representation. Photos on the bottom part are sample pasture plots with (c) low (class: 0.306–0.361), (d) medium (class: 0.472–0.528), and (e) high (class: 0.639–0.694) NDVI values.

consequently, nutritional quality characteristics. This approach improves the accuracy of the pasture quality assessment (Magney et al. 2016). The lower end of this range corresponds to sparse vegetation or early growth stages, while the upper end reflects denser and more vigorous plant cover. Following the selection of the NDVI range, spatial variability was analysed to identify distinct zones representing differences in phenological phases and vegetation conditions. To ensure an even distribution of potential sampling locations across the NDVI gradient, the range (0.25–0.75) was divided into nine equal-width intervals, each spanning 0.055 units. Within each interval, sampling locations were randomly selected to represent the corresponding vegetation conditions, ensuring balanced coverage of biomass variability. This stratified approach reduced the risk of

under- or over-representation of specific NDVI values, thereby maximising the representativeness of the dataset. Sampling points for each NDVI interval were selected within the Sentinel grid, and the geographic coordinates of the centroid of each grid cell were recorded for accurate field navigation. To account for seasonal changes, sampling was conducted at different phenological stages, capturing the full range of biomass conditions observed throughout the growing season.

Field sampling followed a geo-referenced, plot-based approach (Primi et al. 2016), using 1×5 m plots where herbaceous biomass was harvested. Sampling was carried out under dry conditions from May 2023 to May 2024, with 9 samples collected from 2 farms on each sampling day between May 2023 and

February 2024, while a third farm was sampled from March 2024 to May 2024. In total, 324 samples were collected throughout the study period (18 sampling days). Fresh biomass weights were recorded in situ using a Kern HDB dynamometer (accuracy: ± 5 g; Kern & Sohn GmbH, Germany), and cuts were made 2 cm above the ground using a double-bladed shear.

From each plot, approximately 500 g of fresh forage was placed in a sealed plastic bag, with air removed, and stored in a cool box with ice packs for transport to the laboratory. Upon arrival, samples were stored at 4°C in a cold room (Dale et al. 2017) and processed within 24 h. Finally, the material was chopped into approximately 1-cm pieces using sharpened knives, brought to room temperature, and subsequently analysed using NIR spectroscopy.

2.2 | FT-NIR Spectroscopy and Chemical Analyses

2.2.1 | FT-NIR Acquisitions on Fresh and Dried Samples

An FT-NIR spectrometer, model Antaris II (Thermo Fisher Scientific, Madison, WI, USA), was used to acquire sample scans in reflectance mode. Scans were obtained using a 5-cm-diameter sample cup spinner, with its centre of rotation shifted from the port of the integrating sphere. This configuration allowed the acquisition of spectra along a circular path during the scan, increasing the size of the scanned area and then improving the representativeness of the measurements. The resulting spectra are the average of 32 technical replicates, characterised by a resolution of approximately 1.93 cm^{-1} , with a total of 3112 features acquired over the spectral range of 4000 to $10,000\text{ cm}^{-1}$. The NIR spectrum of the fresh sample was obtained by averaging the spectra from three biological replicates, each weighing approximately 5 to 10 grams.

FT-NIR analyses were also performed on the dried samples at room temperature, with one spectrum acquired per sample. The acquisition procedure followed was the same as for fresh samples, ensuring an amount of sample to evenly cover the bottom of the sample cup.

2.2.2 | Drying Samples

Forage samples were dried in a forced-air oven (TCF 120, Argolab, Italy) at 65°C for 72 h to determine dry matter (DM) content, expressed as $\text{g } 100\text{ g}^{-1}$ on a fresh weight (FW) basis. After drying, the samples were ground using a mill (Retsch Müller, Germany) to pass through a 1-mm screen. To ensure optimal storage, the dried samples were stored in sealed polyethylene containers at room temperature.

2.2.3 | Wet Chemistry Analysis

Dried samples were analysed for crude ash (CA) content by combustion in a muffle furnace (GB 10-122, Hobersal, Spain) at

600°C for 2 h (AOAC INTERNATIONAL 2005a) using 3 g of sample. Crude protein (CP) content was determined with the Kjeldahl method (AOAC INTERNATIONAL 2005b) using 1 g of sample. CP determination consists of three phases: digestion, distillation, and titration. For the digestion process, a Speed Digester K-436 (BÜCHI, Switzerland) equipped with a Scrubber B-414 (BÜCHI, Switzerland) was used; the temperature was set at 430°C for 1 h. Then, the distillation phase was carried out in steam for approximately 5 min using a BasicKjel unit (BÜCHI, Switzerland). The last phase, titration, was performed using a Flash automatic titrator (Steroglass, Italy), calibrated with standard pH 4 and 7 solutions, using 0.1 N sulphuric acid as the titrant. Crude fat content, expressed as ether extract (EE), was determined with the Randall extraction system using 3 g of sample (AOAC INTERNATIONAL 2005c). The EE determination consists of three phases: extraction (1 h), washing (1 h), and solvent recovery (20 min). Petroleum ether was used as the extraction solvent, and the analysis was carried out with the SER 148 Semi-Automatic Solvent Extractor (VELP Scientifica, Italy), set at 130°C. Subsequently, the separated fraction, collected in a glass beaker, was dried in a vacuum oven at 82°C for 20 min. Amylase-treated neutral detergent fibre (aNDF), acid detergent fibre (ADF), and acid detergent lignin (ADL) were determined according to the procedure outlined by Van Soest et al. (Van Soest et al. 1991). Crude fibre (CF) was analysed according to AOAC method 962.09 (AOAC INTERNATIONAL 2005d), which involves acid hydrolysis followed by alkaline hydrolysis. Specifically, aNDF was analysed using procedure A (Van Soest et al. 1991), which incorporates amylase and sodium sulphite. The analyses were adapted for the Ankom 200 Fibre Analyzer (ANKOM Technology, NY), using Ankom F57 filter bags. For fibre determination, 0.5 grams of sample were weighed for both aNDF and ADF analysis, while 1 gram was used for CF determination. The Ankom 200 Fibre Analyzer was set at 100°C, with processing times of 75 min for aNDF, 60 min for ADF, and 100 min for CF (50 min each for acid and alkaline hydrolysis). The determination of ADL was performed sequentially after ADF analysis. Following ADF extraction, the sample was treated with 72% sulphuric acid for 3 h. Each sample was analysed in duplicate. All results were expressed as $\text{g } 100\text{ g}^{-1}$ on a dry weight (DW) basis.

2.3 | PLS Models for Predicting Quality Attributes

2.3.1 | Partial Least Squares Regression

Partial least squares (PLS) regression models were developed to predict each quality attribute measured through the conventional laboratory analyses described above. The software utilised for this purpose was PLS_Toolbox 8.9.1 (Eigenvector Research Inc., Manson, WA, USA), integrated in MATLAB® R2020b (The MathWorks Inc., Natick, MA, USA). The results of the PLS regression analysis were expressed in terms of the coefficient of determination (R^2) and root-mean-square error (RMSE) for both the calibration and validation phases, including cross-validation (CV) and prediction performance. The systematic error (bias) was reported to assess the prediction performance of the models.

R^2 measures the variance (i.e., the amount of variability) in the response variable (i.e., the quality attribute) that the regression model explains. An R^2 value of 1 indicates a perfect fit, meaning that the regression model explains all the variability in the quality attribute and that the predictions are in complete agreement with the measured values. An R^2 value close to 0 suggests that the model does not explain the variability in the quality attribute well, indicating a poor fit.

Then, the RMSE is calculated using the following formula:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n_p} (\hat{y}_i - y_i)^2}{n_p}}, \quad (1)$$

where $\hat{y}_i$ is the predicted value, y_i is the observed value, and n_p is the total number of observations. RMSE quantifies the differences between the predictions made by the PLS model and the actual observed values of the quality attribute. RMSE provides a measure of the expected prediction error for the model; lower RMSE values indicate better predictive accuracy.

The root-mean-square error of cross-validation (RMSECV), the number of latent variables, and the coefficient of determination of cross-validation (R^2_{CV}) were considered to identify the best PLS models for each quality attribute. The models were selected based on their performance across these metrics.

2.3.2 | Data Splitting

Before PLS modelling, the dataset was split into two subsets, one for calibration and CV and the other for external validation. The Onion algorithm (Gallagher and O'sullivan 2020) was employed, allocating 70% of the data for calibration and CV and 30% for external validation (fraction = 0.7), ensuring a balanced representation of sample variability. The Onion method starts by selecting a ring of samples that are furthest from the mean, like the outermost layer of an onion, and assigning them to the calibration subset (loopfraction = 0.1). In the next step, a ring of samples closer to the mean is assigned to the validation subset. This process is repeated twice, resulting in three sets of samples for calibration and three for validation (nonion = 3). Finally, any remaining samples are randomly divided between the two subsets. Overall, these proportions and this method of splitting the original dataset were chosen to achieve a fair trade-off between the risk of underfitting and overfitting during model calibration.

The Onion algorithm provided a balanced representation of sample variability across calibration/cross-validation and external validation sets, although the samples were drawn from the same overall dataset. However, it should be noted that the dataset includes samples from different locations, phenological stages and forage types, and that external validation evaluates the performance of the models on the variability characterising these data. A “true” external validation with data from entirely new and independent conditions could be the focus of future work to assess a potential extrapolation capacity of the models.

2.3.3 | Preprocessing

To reduce and correct instrumental noise and light scattering effects, 48 combinations of preprocessing techniques were applied to the spectra. The preprocessing methods included Savitzky-Golay smoothing with three different window widths (15, 23 and 31), either used alone or combined with other techniques; Savitzky-Golay first derivative (window widths: 31, 41 and 51; polynomial order: 2) and second derivative (window widths: 31, 41, 51, 61 and 71; polynomial order: 2); standard normal variate (SNV), extended multiplicative scatter correction (EMSC) and detrending (polynomial order: 1) to eliminate linear offset (Rinnan et al. 2009). All combinations were subsequently processed through mean centring or autoscaling, both of which were also evaluated independently.

2.3.4 | Cross-Validation and Outliers

The PLS models were then developed and cross-validated by applying the various preprocessing combinations to the subset of data for calibration and CV. To ensure robust models by systematically excluding the effects of individual sampling days from the results, leave-1-day-out cross-validation was adopted. Then, the optimal number of latent variables for each PLS model was selected when the RMSECV reached the first local minimum, preferentially choosing the model with the fewest latent variables. This approach enables for the development of more interpretable models with improved generalisability while minimising the risk of under- and overfitting. Finally, outliers were identified and removed in the case of consistent systematic errors in the predicted values generated by the regression model.

2.3.5 | Variable Selection

Variable selection algorithms help in reducing the dimensionality of datasets and assist in identifying variables (i.e., predictors like wavenumbers) most closely associated with the chemical fingerprints of samples. For these reasons, the effect of two variable selection techniques on model performance was tested. The purpose of these two methods was to extract relevant information from different spectral regions, focusing on significant features while removing interference from the irrelevant ones. The first technique adopted was the interval partial least squares (iPLS) algorithm (Nørgaard et al. 2000), which begins by selecting the spectral interval that results in the lowest RMSECV. The algorithm then iteratively adds another interval in “forward” mode until the RMSECV starts to increase, indicating potential overfitting. Three different iPLS settings were chosen: (i) 120 variables × 9 windows; (ii) 60 variables × 6 windows; and (iii) 30 variables × 3 windows (Table 1). The second variable selection technique adopted was the genetic algorithm (GA). GA generates a population of random subsets of variables (i.e., individuals). Each individual is evaluated based on its fitness, calculated by PLS regression, which predicts the actual quality attribute of interest and is expressed in terms of RMSECV. The algorithm then discards the lower half of the individuals based on their fitness scores

TABLE 1 | Variable selection methods that were applied in the study.

Variable selection method	Bandwidth (# variables)	Bands selected	Variables selected	Acronym
iPLS	30	3	90	iPLS_30
	60	6	360	iPLS_60
	120	9	1080	iPLS_120
GA	30	Undefined	Undefined	GA_30
	40	Undefined	Undefined	GA_40

Note: Variable is a spectral feature.

Abbreviations: GA, genetic algorithm; iPLS, interval partial least squares.

TABLE 2 | Descriptive statistics of the quality attributes of the green forage.

	Mean	SD	SE	1st qu.	Median	3rd qu.	Min	Max	Range	Skew	Kurtosis
DM	29.53	15.68	0.87	18.86	24.79	34.04	7.57	90.56	82.99	1.69	2.74
CA	10.47	3.18	0.18	8.19	10.29	12.06	4.39	26.01	21.62	1.10	2.51
CP	16.06	7.64	0.42	9.64	15.06	21.23	2.53	35.17	32.64	0.44	−0.78
EE	1.91	0.68	0.04	1.42	1.83	2.33	0.28	4.41	4.13	0.44	0.14
aNDF	52.48	11.52	0.64	45.66	52.56	60.23	25.85	83.37	57.53	0.01	−0.28
ADF	38.27	10.05	0.56	31.04	36.01	45.53	18.06	64.17	46.12	0.47	−0.56
ADL	9.03	3.64	0.20	6.28	8.41	11.32	2.75	23.73	20.98	0.92	0.99
CF	29.50	9.09	0.50	23.04	28.20	35.27	9.56	63.41	53.85	0.69	0.83

Note: **Quality attributes.** DM, dry matter; CA, crude ash; CP, crude protein; EE, ether extract (i.e., crude fat); aNDF, amylase-treated neutral detergent fibre; ADF, acid detergent fibre; ADL, acid detergent lignin; CF, crude fibre. **Headers.** SD, standard deviation; SE, standard error; 1st qu., first quartile; 3rd qu., third quartile. **Units.** FW, fresh weight; DW, dry weight. DM, g 100 g^{−1} FW; CA, CP, EE, aNDF, ADF, ADL, CF, g 100 g^{−1} DW. Mean, SD, SE, 1st qu., median, 3rd qu., min, max and range are expressed in the same unit as the relative quality attribute.

and evolves the population through reproduction and mutation mechanisms. This iterative process continues until predefined termination criteria are met (Leardi et al. 1992). For GA, two different variable window widths (30 and 40, corresponding to GA_30 and GA_40), were set, allowing an unlimited number of selectable bands to ensure flexibility in variable selection while avoiding excessive model sparsity (Table 1).

2.3.6 | Comparison of the Best Models and External Validation

All the PLS models were obtained by adopting several combinations of preprocessing and variable selection methods. To compare the top-performing models related to each quality attribute, normalised root-mean-square error of cross-validation (NRMSECV) was adopted. This metric provides a standardised way to evaluate and rank models based on their predictive accuracy relative to the scale of the observed data. The NRMSECV is calculated as follows:

$$\text{NRMSECV} = \frac{\text{RMSECV}}{y_{\text{range}}} = \frac{\sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n_p}}}{y_{\text{max}} - y_{\text{min}}}, \quad (2)$$

where $y_{\text{range}} = y_{\text{max}} - y_{\text{min}}$ represents the difference between the maximum and minimum observed values of the quality attribute.

To ensure the robustness of the models, they were finally subjected to external validation using the specifically prepared subset of data.

3 | Results and Discussion

Data from wet chemistry analysis of forage quality attributes and spectra from FT-NIR spectroscopy constitute the two datasets for the development of PLS prediction models. Descriptive statistics of data from wet chemistry analysis are reported in Table 2 and are in line with pasture composition. The extended DM range from 7.57 to 90.56 g 100 g^{−1} FW could be an effect of fluctuations in forage composition in terms of species, phenological stage, phytosanitary status and environmental condition. The high mean value of DM, close to 30% on FW (29.53 g 100 g^{−1} FW), coupled with the high variability (SD = 15.68 g 100 g^{−1} FW), underscores the effect of seasonality, particularly during summer droughts. Furthermore, DM has the highest kurtosis value (2.74), indicating heavier tails compared to a normal distribution and suggesting a higher probability of extreme values. The positive skewness (1.69), which results in a right-skewed distribution, further indicates the tendency for a higher DM content in some samples.

The high mean CP value of 16.06 g 100 g^{−1} DW is indicative of a relatively high protein content in the forage, which is essential for animal growth and milk production. Furthermore, the first

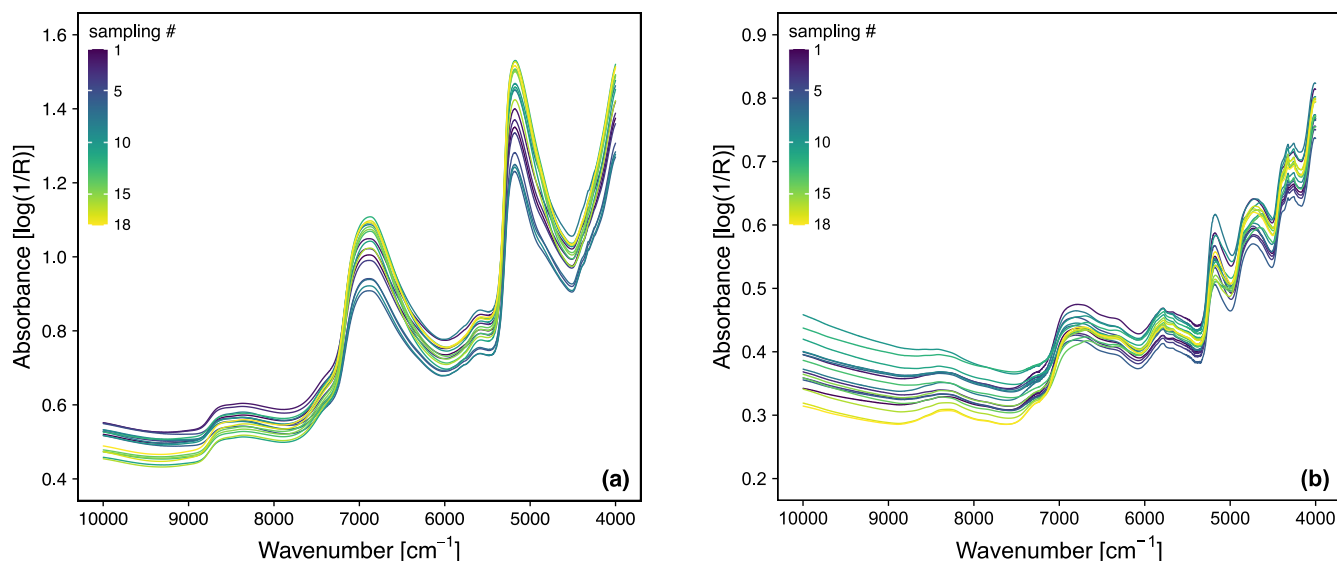


FIGURE 2 | Mean absorbance FT-NIR spectra per day of sampling related to (a) fresh forage and (b) dried-milled forage samples.

quartile (Q1) is $9.64 \text{ g } 100 \text{ g}^{-1} \text{ DW}$, while the third quartile (Q3) is $21.23 \text{ g } 100 \text{ g}^{-1} \text{ DW}$. These statistics indicate that half of the samples fall within a protein range that is suitable for cattle feeding (Pereira-Crespo et al. 2022). Differing from DM, a skewness of 0.44 suggests a more symmetrical distribution, indicating that the protein content is relatively consistent across samples. Prolonged drought conditions have been shown to accelerate forage maturation, leading to increased DM content and reduced CP levels, which in turn negatively affect digestibility and nutrient availability for livestock (Muzzo et al. 2025). These shifts in forage quality can impact livestock feeding strategies, requiring adjustments in dietary supplementation to maintain optimal animal performance.

aNDF, ADF and ADL generally follow normal distributions, with some asymmetry observed in ADF and ADL (0.47 and 0.92, respectively), suggesting greater variability. ADF is about half the concentration of aNDF, implying relevant digestible content. The mean ADF/ADL ratio (4.2) suggests relatively high levels of lignin, which could suggest potential digestibility issues. The data show a balance between mineral and fibre content, expressed as CA and CF, respectively. The high ranges and skewness in CA and CF suggest that different forage types could affect the overall quality, leading to differences in nutritional content and digestibility. As expected, EE (i.e., crude fat) ($1.91 \pm 0.68 \text{ g } 100 \text{ g}^{-1} \text{ DW}$) was low, indicating it does not contribute substantially to the overall energy intake.

The second set of data acquired for the development of the PLS prediction models is composed of the spectra from FT-NIR spectroscopy. In Figure 2, the mean FT-NIR spectra for fresh and dried forage are reported, organised by sampling day. The spectra of both kinds of samples show a baseline shift, which is typical in NIR spectroscopy. In addition, a multiplicative effect is evident in the spectra of the dried samples in the $10,000\text{--}7000 \text{ cm}^{-1}$ region. This effect could arise from physical phenomena like light scattering caused by the particle size distributions of the powdery sample, highlighting the importance of sample preparation. Furthermore, the spectra are significantly influenced by the presence of water. In particular, the

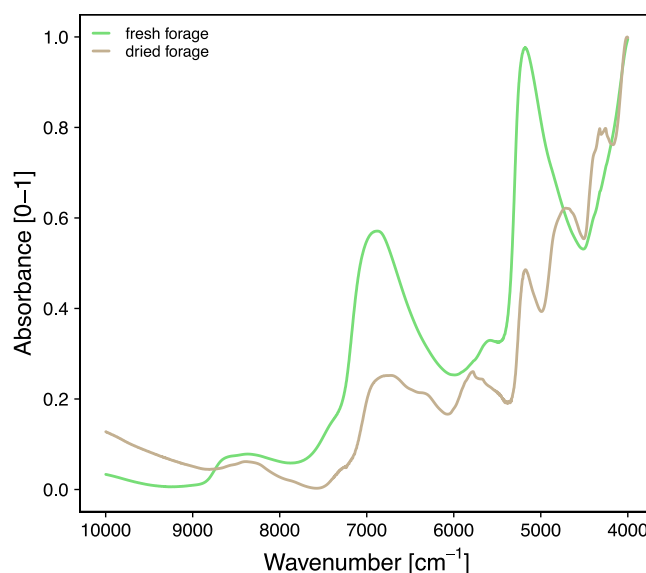


FIGURE 3 | Comparison of overall mean FT-NIR absorbance spectra between fresh and dried forage samples. Spectra are min-max normalised (arbitrary units).

absorbance varies in intensity and is generally higher in the spectra of the fresh forage samples. Additionally, some peaks appear only in the spectra of the dried samples.

To enhance the visualisation of the differences between the overall mean spectra of fresh and dried samples, the absorbance units in Figure 3 are normalised using a min-max scaling approach (corresponding to the actual value minus the minimum value, all divided by the difference between the maximum and minimum values). This normalisation procedure emphasises the impact of moisture content on the bands observed at approximately 6900 cm^{-1} (1st overtone of O-H stretching) and 5180 cm^{-1} (combination band). Characteristic lipid-related peaks are also visible in the mean spectrum of dried forage at $\sim 4320\text{--}4250 \text{ cm}^{-1}$ (2nd overtone of C-H bending; C-H stretching and C-C stretching combination bands, respectively) and $\sim 5785\text{--}5680 \text{ cm}^{-1}$ (1st overtone of C-H stretching), even

TABLE 3 | Results of the PLS models for prediction of the quality attributes of the green forage.

LVs	X-Preprocessing	X VarSel method	X Size	Y range CV	RMSEC (Cal)	RMSECV (CV)	RMSEP (Pred)	NRMSECV	R ² _c (Cal)	R ² _{cv} (CV)	R ² _p (Pred)	Bias (Pred)
Fresh forage												
DM 6	EMSC, AS	GA_30	227 × 630	82.99	3.020	3.165	2.951	0.038	0.965	0.962	0.960	0.521
DM 3	S (w:31), EMSC, AS	iPLS_30	227 × 90	82.99	3.337	3.474	3.455	0.042	0.957	0.954	0.944	0.388
DM 5	EMSC, MC	—	227 × 3112	82.99	3.327	3.501	3.086	0.042	0.958	0.953	0.954	0.451
CA 6	2ndD (w:31), MC	iPLS_120	227 × 1080	21.62	1.835	2.127	1.693	0.098	0.694	0.591	0.641	0.038
CA 6	1stD (w:31), MC	—	227 × 3112	21.62	2.131	2.433	1.534	0.113	0.588	0.470	0.708	0.266
CP 3	2ndD (w:71), MC	iPLS_120	227 × 1080	32.64	3.677	4.007	3.676	0.123	0.773	0.732	0.756	−0.276
CP 5	2ndD (w:61), MC	—	227 × 3112	32.64	3.596	4.134	3.563	0.127	0.783	0.716	0.770	−0.298
EE 5	S (w:15), SNV, MC	iPLS_120	227 × 1080	4.13	0.480	0.502	0.488	0.121	0.522	0.479	0.445	−0.048
EE 2	2ndD (w:31), AS	—	227 × 3112	4.13	0.502	0.532	0.520	0.129	0.477	0.412	0.377	−0.064
aNDF 7	2ndD (w:31), AS	iPLS_60	227 × 360	56.48	4.373	5.253	6.082	0.093	0.843	0.776	0.762	0.611
aNDF 3	2ndD (w:31), MC	iPLS_30	227 × 90	56.48	5.250	5.679	6.993	0.101	0.774	0.736	0.686	0.789
aNDF 7	2ndD (w:41), MC	—	227 × 3112	56.48	4.624	5.618	6.371	0.099	0.825	0.744	0.739	0.584
ADF 3	2ndD (w:41), AS	iPLS_60	227 × 360	45.41	4.614	5.088	5.123	0.112	0.796	0.753	0.715	0.393
ADF 3	2ndD (w:31), AS	—	227 × 3112	45.41	4.498	5.488	5.337	0.121	0.807	0.715	0.695	0.659
ADL 4	2ndD (w:61), AS	GA_40	227 × 560	20.98	2.189	2.500	2.297	0.119	0.657	0.556	0.588	−0.098
ADL 3	2ndD (w:51), AS	—	227 × 3112	20.98	2.355	2.824	2.404	0.135	0.603	0.434	0.522	−0.190
CF 3	S (w:23), MC	iPLS_120	227 × 1080	47.27	6.293	6.598	7.170	0.140	0.516	0.469	0.392	−1.102
CF 3	2ndD (w:61), AS	—	227 × 3112	47.27	5.852	6.626	7.314	0.140	0.581	0.469	0.376	−0.926
Dried forage												
CP 2	EMSC, MC	GA_40	189 × 720	32.64	1.775	1.815	1.841	0.056	0.951	0.949	0.939	0.202
CP 2	EMSC, MC	iPLS_30	189 × 90	32.64	1.884	1.958	1.865	0.060	0.945	0.940	0.937	0.169
CP 2	S (w:23), EMSC, MC	—	189 × 3112	32.64	1.967	2.103	2.075	0.064	0.940	0.931	0.923	0.264
EE 7	2ndD (w:31), MC	GA_40	227 × 480	4.13	0.349	0.403	0.348	0.098	0.745	0.663	0.731	0.073
EE 7	2ndD (w:31), MC	—	227 × 3112	4.13	0.353	0.442	0.361	0.107	0.738	0.596	0.700	0.030
aNDF 5	2ndD (w:71), AS	GA_40	227 × 560	57.09	3.913	4.499	4.202	0.079	0.885	0.849	0.865	−0.463
aNDF 6	S (w:31), DET, MC	iPLS_30	227 × 90	57.09	4.520	4.974	4.508	0.087	0.847	0.815	0.867	−0.911
aNDF 4	2ndD (w:31), AS	—	227 × 3112	57.09	3.662	4.711	4.317	0.083	0.900	0.834	0.862	−0.878
ADF 3	2ndD (w:61), MC	GA_30	227 × 480	46.12	3.854	4.160	3.529	0.090	0.858	0.835	0.862	−0.060

(Continues)

TABLE 3 | (Continued)

	LVs	X-Preprocessing	X VarSel method	X Size	Y range CV	RMSEC (Cal)	RMSECV (CV)	RMSEP (Pred)	NRMSECV	R ² _c (Cal)	R ² _{cv} (CV)	R ² _p (Pred)	Bias (Pred)
ADF	3	2ndD (w:61), AS	—	227 × 3112	46.12	3.936	4.562	3.586	0.099	0.852	0.802	0.856	−0.030
ADL	4	2ndD (w:61), AS	GA_40	226 × 600	20.98	1.951	2.172	2.029	0.104	0.708	0.639	0.699	−0.151
ADL	4	2ndD (w:71), AS	—	226 × 3112	20.98	2.064	2.357	2.060	0.112	0.673	0.576	0.621	−0.092
CF	4	2ndD (w:51), AS	iPLS_60	227 × 360	49.02	5.709	6.373	5.645	0.130	0.643	0.558	0.517	−0.493
CF	4	2ndD (w:61), AS	—	227 × 3112	49.02	5.5644	6.8897	5.6552	0.141	0.661	0.490	0.520	−0.580

Note: **Samples.** DM, dry matter; CP, crude protein; CA, crude ash; EE, ether extract (i.e. crude fat); aNDF, amylase-treated neutral detergent fibre; ADF, acid detergent fibre; ADL, acid detergent lignin; CF, crude fibre. **Headers.** X is referred to as the spectral variable, and Y is referred to as the quality attribute variable. LVs, latent variables; X VarSel method, X variable selection method; Y range CV, Y variable range of the cross-validation subset; RMSE, root-mean-square error of calibration (RMSEC), cross-validation (RMSECV), and prediction (RMSEP); NRMSECV, normalised root-mean-square error of cross-validation, NRMSECV = RMSECV/Y range CV (expressed as %); R², coefficient of determination of calibration (R²_C), cross-validation (R²_{CV}), and prediction (R²_P). **Units.** FW, fresh weight; DW, dry weight; DM, g 100 g^{−1} FW; CA, CP, EE, aNDF, ADF, ADL, CF, g 100 g^{−1} DW. Y, RMSE, and Bias have the same unit as their respective quality attribute. X size: number of samples × number of spectral variables. **Preprocessing.** 1stD: Savitzky-Golay 1st derivative; 2ndD: Savitzky-Golay 2nd derivative; DET, detrending; EMSC, extended multiplicative scatter correction; S, Savitzky-Golay smoothing; SNV, standard normal variate; AS, autoscaling; MC, mean centre; w, smoothing window. **X variable selection method.** GA, genetic algorithm; iPLS, interval PLS. The number following the underscore is the window size.

though lipids are present in modest amounts (Kays 2004; Workman and Weyer 2007; Barbin et al. 2014).

The best results of PLS models for predicting the quality attributes of green forage are reported in Table 3. To compare model performances, the NRMSECV was used to normalise RMSECV by considering the range of the Y variable (i.e., the quality attribute) and expressing the error as a percentage of its variability. Therefore, the two best models for fresh forage were DM and aNDF, with NRMSECV of 3.8% and 9.3%, respectively. For dried samples, the two best models were CP and aNDF, with NRMSECV of 5.6% and 7.9%, respectively (Table 3).

The Savitzky-Golay second derivative (2ndD) was the pre-processing step most frequently used in the best models. The 2ndD method employs different window sizes (ranging from 31 to 71, with increments of 10), followed by either mean centring or autoscaling. This method effectively removes irrelevant baseline signals from the spectra. Additionally, it suggests that the variables may be strongly correlated with each other and that adjacent variables might contain similar correlated signals. The EMSC was adopted for the best DM models of fresh samples and CP models of dried samples. EMSC corrected multiplicative effects in the spectra, which were primarily caused by light dispersion resulting from the physical characteristics of the samples. When considering fresh samples, the irregular arrangement of the approx. 1-cm fragments may have led to misalignment or overlapping during FT-NIR analysis, potentially creating empty space at the bottom of the sample cup, which could have influenced the measurements. In contrast, for dried and ground samples, the scattering effect can probably be attributed to their coarse and dusty texture, which results in a more substantial multiplicative effect observed in the spectra.

PLS models based on selected variables obtained through the two adopted spectral variable selection methods, iPLS and GA, generally outperformed models that included the full set of variables, as evidenced by the superior metrics in both CV and external set prediction (Table 3). Noteworthy, the iPLS method was particularly effective for fresh forage models, achieving the best results in 6 out of 8 cases. In contrast, the GA method was more effective for dried forage models, achieving the best outcomes in 5 out of 6 cases. Figure 4 compares the results of all variable selection methods applied to obtain the four best PLS models in terms of NRMSECV, highlighting the wavenumber bands selected. The GA_30 and GA_40 methods were set to provide a variable number of wavenumber bands, which may highlight the presence of synergies between bands but may also lead to more complex models. In contrast, the iPLS methods were set to provide a fixed number of bands: nine bands for iPLS_120, six for iPLS_60, and three for iPLS_30. The comparison of the results of the five variable selection methods revealed a trend in the selection of wavenumber bands in some spectral regions. Except for the aNDF model for fresh samples (Figure 4b), GA yielded the best results, represented by blue bars intersecting the red spectrum. Regarding DM (Figure 4a), two regions selected by at least four methods are approximately at 5500–5440 cm^{−1} and 4230–4150 cm^{−1}, which may be associated with lignin, typically found at ~6250–5550 cm^{−1} and ~4550–4150 cm^{−1} (Shenk and Westerhaus 1994). Similar observations can be made for the aNDF model of fresh samples

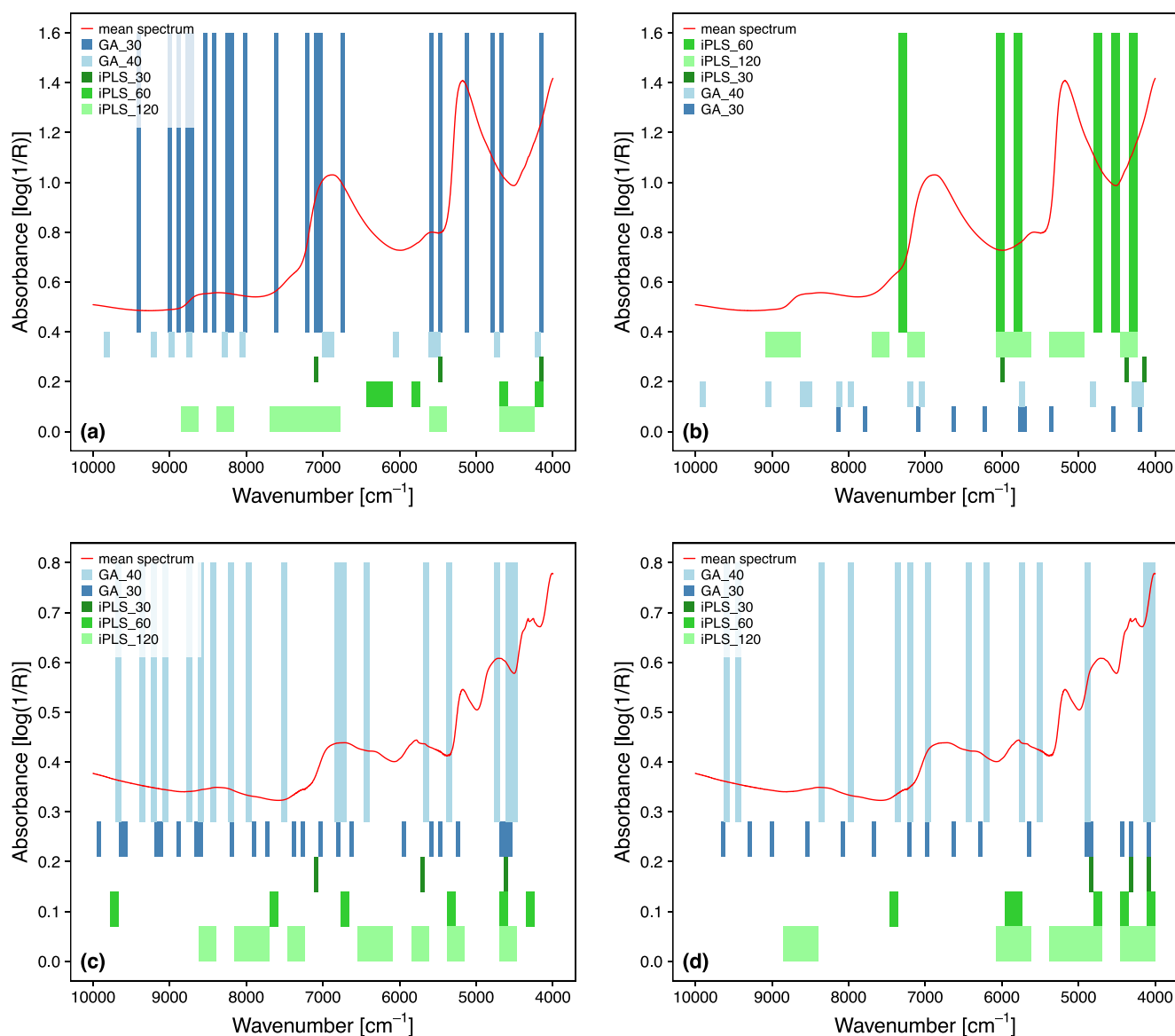


FIGURE 4 | Comparison of the selected wavenumber bands obtained through genetic algorithm (GA) and interval partial least squares (iPLS) variable selection methods for the best four PLS models related to (a) dry matter (DM) and (b) amylase-treated neutral detergent fibre (aNDF) for fresh forage samples, and (c) crude protein (CP) and (d) aNDF for dried samples. The bars crossing the red spectrum indicate the top-performing variable selection methods. The number following the underscore in the legend denotes the window width (i.e., the number of features included in each window).

(Figure 4b), with the two regions $\sim 4400\text{--}4120\text{ cm}^{-1}$ and $\sim 5850\text{--}5620\text{ cm}^{-1}$ that could be associated with lignin and cellulose. Specifically, cellulose is one of the constituents of fibre, which absorbs in the $\sim 4350\text{--}4270\text{ cm}^{-1}$ region. CP for dried samples (Figure 4c) presents a common region at $\sim 4770\text{--}4460\text{ cm}^{-1}$, which could be associated with proteins, which absorb at $\sim 4610\text{--}4590\text{ cm}^{-1}$ (combination bands of N-H bend and C-H/C=O stretch) (Shenk and Westerhaus 1994). Finally, aNDF for dried samples (Figure 4d) presents interesting absorption bands in the $\sim 5000\text{--}4000\text{ cm}^{-1}$ region, specifically at $\sim 4150\text{--}4000$, $\sim 4460\text{--}4290$ and $\sim 4920\text{--}4690\text{ cm}^{-1}$, which can be associated with lignin and cellulose. Finally, to identify a set of “core” spectral regions that can be used for the development of portable instruments, the results of the best models obtained by the iPLS₃₀ variable selection method are reported in Table 3. Regarding CP in dried samples, the

iPLS₃₀ model presented prediction performance perfectly in line with the best model with variable selection GA₄₀, with RMSEP of 1.865 and 1.841 g 100 g⁻¹ DW, and R²_P of 0.937 and 0.939 for the iPLS₃₀ and GA₄₀ models, respectively. In contrast, the predictive error performance of the other three iPLS₃₀ models (DM and aNDF for fresh samples, and aNDF for dried samples) were slightly lower. These results are a good compromise between accuracy and a large reduction of variables (from 3112 to 3 bands of 30 variables). Furthermore, this large selection of variables significantly reduces the instrumental and computational complexity required, making the models ideally implementable on compact sensors and devices, with benefits in terms of cost, power consumption and acquisition speed. The wavenumber bands of the respective iPLS₃₀ predictive models are reported in Figure 4.

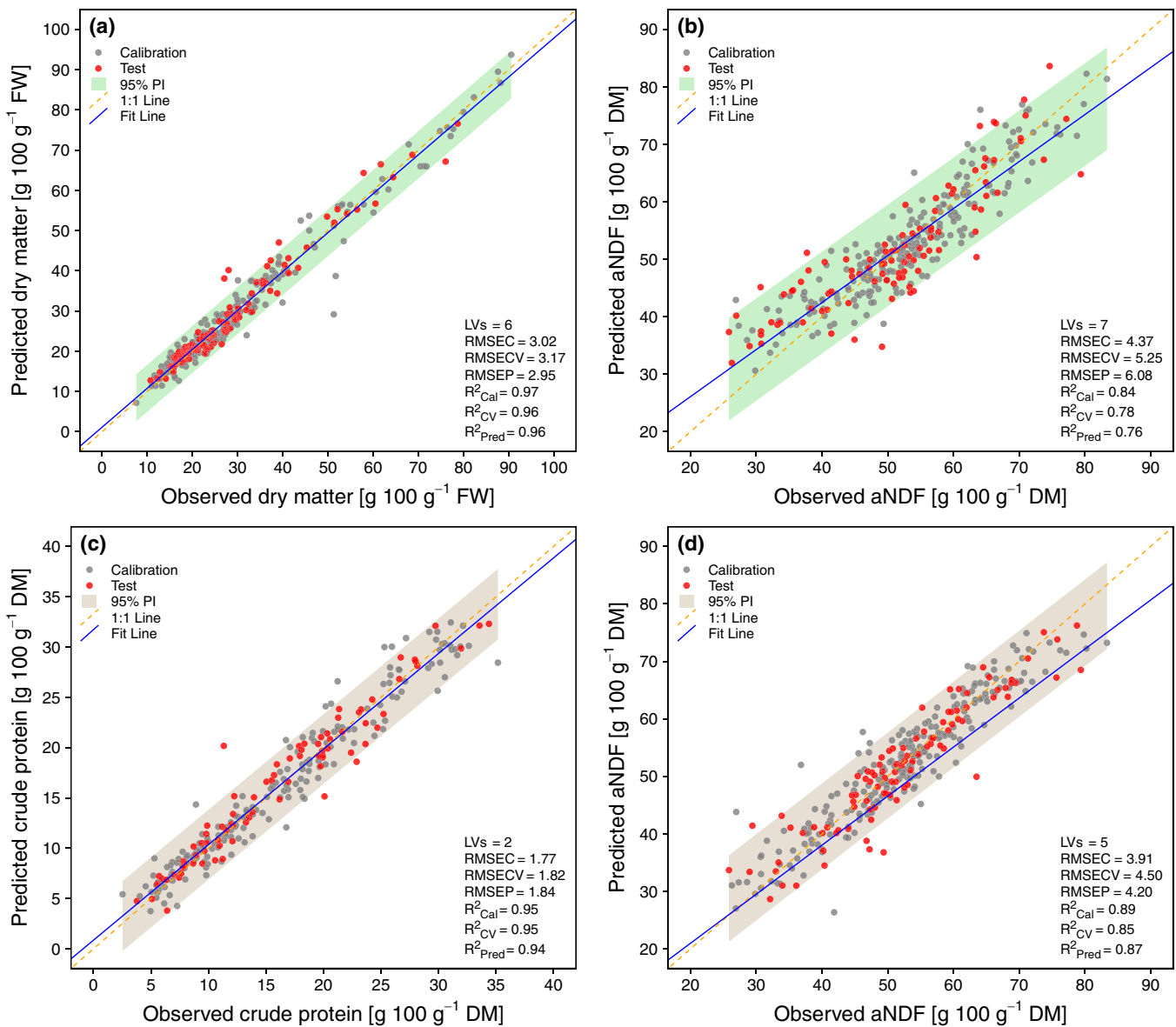


FIGURE 5 | Observed versus predicted values from the best four PLS regression models related to (a) dry matter (DM) and (b) amylase-treated neutral detergent fibre (aNDF) for fresh forage samples, (c) crude protein (CP) and (d) aNDF for dried-milled forage samples. PI is the prediction interval; LVs is the number of latent variables.

The measured versus predicted values for the top four PLS models, previously discussed, are presented in Figure 5. Among these, DM for fresh samples and CP for dried samples (Figure 5a,c, respectively) stand out in terms of metric performances. The DM model achieves a coefficient of determination of prediction (R^2_P) of 0.96 and a root-mean-square error of prediction (RMSEP) of 2.95 g 100 g⁻¹ FW, while the CP model yields an R^2_P of 0.94 and an RMSEP of 1.84 g 100 g⁻¹ DW. Moreover, the CP model, which is characterised by only two latent variables, is simpler and more robust. Regarding aNDF, for fresh and dried samples (Figure 5b,d, respectively), the performance of PLS models was worse compared with DM and CP models. The dried samples model yielded better results, with fewer latent variables (five vs. seven) and superior metrics, achieving an R^2_P of 0.87 versus 0.76 for fresh samples and an RMSEP of 4.20 g 100 g⁻¹ DW versus 6.08 g 100 g⁻¹ DW. The relatively lower prediction performance for aNDF may be caused by the intrinsic structural variability of the forage

samples and its complex composition, which consists of a heterogeneous fraction of cellulose, hemicellulose and lignin. This can result in overlapping spectral absorption characteristics, affecting model performance. Furthermore, a high water content in fresh samples may cover or dilute the naturally variable spectral characteristics of the aNDF components. Finally, sample preparation, for example, drying and grinding, can influence the physical properties of the heterogeneous fibrous material, introducing spectral differences (e.g., shifts in peak positions, changes in peak intensity or spectral offsets).

In a recent study, Murphy et al. (Murphy et al. 2022), developed calibration models for DM and CP from FT-NIR spectra of fresh perennial ryegrass from Irish pastures. The homogeneity of the sample, composed of a single grass species, may have positively influenced the results. Prediction errors were expressed as standard error of prediction (SEP), which is not directly comparable to the RMSE that includes systematic error (BIAS). The

SEP resulted in $2.038 \text{ g } 100 \text{ g}^{-1} \text{ DW}$ for CP and $0.946 \text{ g } 100 \text{ g}^{-1}$ for DM. When normalised across the prediction measurement range, the errors corresponded to 9.6% for CP and 7.9% for DM. These findings are comparable to the NRMSECV values obtained in the current study, 12.3% for CP of fresh green forage and 3.8% for DM, indicating a slightly worse result for CP and a better result for DM (Table 3). In a previous study, Danieli et al. (Danieli et al. 2004), analysed dried forage from dairy sheep farms in central Italy over a period of 3 years. The results from their cross-validation models, expressed as the standard error of cross-validation normalised within the observed range of the quality attribute (NSECVC), are comparable to those of the present study. However, the number of outliers removed was not reported. The error for CP was equivalent to that observed in the present study (NSECVC and NRMSECV are equal to 5.6%). Regarding aNDF, the error was slightly better (NSECVC of 6.2%, vs. an NRMSECV of 7.9%) (Table 3).

Overall, PLS models provided good predictive performance for some nutritional quality attributes of green forage. This result could lead to the development of real-time monitoring systems, which could improve livestock management. However, it is important to acknowledge some limitations of the study: (i) the relatively small dataset and (ii) the limited geographical sampling area only partially represent the diversity of forages, which is also determined by environmental conditions and species variability. These limitations reduce the generalisability of the predictive models. Furthermore, also considering the numerosity of the wavenumber bands of some prediction models, these limitations could contribute to overfitting in practical applications. Future research should focus on long-term studies supported by deep learning techniques that could improve the performance of predictive models. In addition, this approach could be combined with complementary spectroscopic techniques to provide better insights into forage nutritional quality dynamics.

4 | Conclusions

The present study provided predictive models for the evaluation of the nutritional quality of green forage by NIR spectroscopy as a rapid and cost-effective alternative to traditional wet chemistry. The best PLS models resulted for DM in fresh forage ($R^2_P = 0.96$, RMSEP = $2.95 \text{ g } 100 \text{ g}^{-1} \text{ FW}$) and CP in dried samples ($R^2_P = 0.94$, RMSEP = $1.84 \text{ g } 100 \text{ g}^{-1} \text{ DW}$). Good results were also obtained for aNDF. Variable selection methods (interval PLS and genetic algorithms) significantly improved model accuracy; interval PLS performed better with fresh forage, while genetic algorithm performed better with dried samples. Furthermore, interval PLS demonstrated its potential in selecting a minimal set of “core” spectral bands with acceptable accuracy, suitable for the development of compact and cost-effective portable NIR sensors for use in the field. Sample characteristics and preparation, particularly the heterogeneity of the forage species and their growth stage, and sample preparation, as drying and milling processes, influenced the NIR spectra and consequently the performance of the predictive models. Future research should focus on expanding the dataset to include larger geographical areas with specific climatic

conditions and related forage species at different phenological stages to improve the robustness and generalisability of the predictive models for wider adoption in forage quality management.

The study illustrates the potential of NIR spectroscopy in improving forage and pasture quality management, which impacts livestock sustainability. The use of FT-NIR spectroscopy provided the basis for low-cost, real-time forage and pasture monitoring, facilitating adaptive grazing practices.

Author Contributions

Alessandro Benelli: conceptualisation, formal analysis, investigation, methodology, software, validation, visualisation, writing – original draft, writing – review and editing. **Riccardo Primi:** conceptualisation, funding acquisition, methodology, software, writing – original draft, writing – review and editing. **Chiara Evangelista:** investigation, methodology, writing – original draft. **Raffaello Spina:** investigation, writing – original draft, writing – review and editing. **Marco Milanesi:** resources. **Daniele Pietrucci:** writing – review and editing. **Bruno Ronchi:** supervision. **Umberto Bernabucci:** conceptualisation, funding acquisition, project administration, supervision, writing – review and editing. **Roberto Moscetti:** conceptualisation, funding acquisition, methodology, software, supervision, writing – review and editing.

Acknowledgements

This study was carried out in the framework of the project AGRITECH of the European Union NextGenerationEU (grant number J83C22000830005) within the “National Research Centre for Agricultural Technologies” research programme. This article reflects only the authors’ views and opinions; neither the European Union nor the European Commission can be considered responsible for them. Moreover, our sincere thanks to Gloria Bernabucci, Irene Bottoni, Amar Brunetti, Stefano Ceci, Francesca Cellitti, Alessio Cocchi, Tommaso Del Signore, Giuseppe Fabbri, Pedro Girotti, and Eleonora Taormina for their valuable help with sampling and wet chemical analyses. Open access publishing facilitated by Università degli Studi della Tuscia, as part of the Wiley - CRUI-CARE agreement.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

- AOAC INTERNATIONAL. 2005a. *Official Method 942.05*. 18th ed., 8. Official Methods of Analysis of AOAC INTERNATIONAL.
- AOAC INTERNATIONAL. 2005b. *Official Method 984.13*. 18th ed., 31. Official Methods of Analysis of AOAC INTERNATIONAL.
- AOAC INTERNATIONAL. 2005c. *Official Method 2003.06*. 18th ed., 42–44. Official Methods of Analysis of AOAC INTERNATIONAL.
- AOAC INTERNATIONAL. 2005d. *Official Method 962.09*. 18th ed., 44–46. Official Methods of Analysis of AOAC INTERNATIONAL.
- Barbin, D. F., A. L. S. M. Felicio, D. W. Sun, S. L. Nixdorf, and E. Y. Hirooka. 2014. “Application of Infrared Spectral Techniques on Quality and Compositional Attributes of Coffee: An Overview.” *Food Research International* 61: 23–32. <https://doi.org/10.1016/j.foodres.2014.01.005>.

- Bernués, A., R. Ruiz, A. Olaizola, D. Villalba, and I. Casasús. 2011. "Sustainability of Pasture-Based Livestock Farming Systems in the European Mediterranean Context: Synergies and Trade-Offs." *Livestock Science* 139, no. 1–2: 44–57. <https://doi.org/10.1016/j.livsci.2011.03.018>.
- Bewley, J. Precision Dairy Farming: Advanced Analysis Solutions for Future Profitability. The First North American Conference on Precision Dairy Management. 2010 p. 16. https://www.researchgate.net/publication/267711814_Precision_Dairy_Farming_Advanced_Analysis_Solutions_for_Future_Profitability.
- Cheli, F., D. Battaglia, L. Pinotti, and A. Baldi. 2012. "State of the Art in Feedstuff Analysis: A Technique-Oriented Perspective." *Journal of Agricultural and Food Chemistry* 60, no. 38: 9529–9542. <https://doi.org/10.1021/jf302555b>.
- Dale, A. J., A. W. Gordon, J. Archer, and C. P. Ferris. 2017. "Impact of Sampling and Storage Technique, and Duration of Storage, on the Composition of Fresh Grass When Analysed Using Near-Infrared Reflectance Spectroscopy." *Grass and Forage Science* 72, no. 4: 614–625. <https://doi.org/10.1111/gfs.12266>.
- Danieli, P. P., P. Carlini, U. Bernabucci, and B. Ronchi. 2004. "Quality Evaluation of Regional Forage Resources by Means of Near Infrared Reflectance Spectroscopy." *Italian Journal of Animal Science* 3, no. 4: 363–376. <https://doi.org/10.4081/ijas.2004.363>.
- Draghi, S., N. E. Fehri, F. Ateş, et al. 2024. "Use of Hair as Matrix for Trace Elements Biomonitoring in Cattle and Roe Deer Sharing Pastures in Northern Italy." *Animals: An Open Access Journal From MDPI* 14, no. 15: 2209. <https://doi.org/10.3390/ani14152209>.
- Dunn, B. H., A. J. Smart, R. N. Gates, et al. 2010. "Long-Term Production and Profitability From Grazing Cattle in the Northern Mixed Grass Prairie." *Rangeland Ecology and Management* 63, no. 2: 233–242. <https://doi.org/10.2111/rem-d-09-00042.1>.
- Dwivedi, V., J. Ahokas, J. Viljanen, et al. 2023. "Optical Assessment of the Spatial Variation in Total Soil Carbon Using Laser-Induced Breakdown Spectroscopy." *Geoderma* 436: 116550. <https://doi.org/10.1016/j.geoderma.2023.116550>.
- Evangelista, C., L. Basiricò, and U. Bernabucci. 2021. "An Overview on the Use of Near Infrared Spectroscopy (NIRS) on Farms for the Management of Dairy Cows." *Agriculture (London)* 11, no. 4: 296. <https://doi.org/10.3390/agriculture11040296>.
- FAO. Greenhouse Gas Emissions From the Dairy Sector—A Life Cycle Assessment. Rome, Italy; 2010. <https://www.fao.org/4/k7930e/k7930e00.pdf>.
- FAO. Land statistics 2001–2022. Global, Regional and Country Trends. Rome, Italy: FAOSTAT Analytical Briefs No. 88; 2024. <https://doi.org/10.4060/cd1484en>.
- Gallagher, N. B., and D. O'sullivan Selection of Representative Learning and Test Sets Using the Onion Method. White Pap Eig. 2020. https://eigenvector.com/wp-content/uploads/2020/01/Onion_SampleSelection.pdf.
- González, L. A., I. Kyriazakis, and L. O. Tedeschi. 2018. "Review: Precision Nutrition of Ruminants: Approaches, Challenges and Potential Gains." Supplement, *Animal* 12, no. s2: s246–s261. <https://doi.org/10.1017/s1751731118002288>.
- Holmes, C. W. 2007. *Milk Production From Pasture*, 602. Massey University.
- Iqbal, M. W., I. Draganova, P. C. H. Morel, and S. T. Morris. 2023. "Associations of Grazing and Rumination Behaviours With Performance Parameters in Spring-Calving Dairy Cows in a Pasture-Based Grazing System." *Animals: An Open Access Journal From MDPI* 13, no. 24: 3831. <https://doi.org/10.3390/ani13243831>.
- Kays, S. E. 2004. "Analysis of Cereal Food Products." In *Near-Infrared Spectroscopy in Agriculture*, edited by C. A. Roberts, J. Workman, and J. B. Reeves, 411–438. American Society of Agronomy, Inc. Crop Science Society of America, Inc. Soil Science Society of America, Inc. <https://doi.org/10.2134/agronmonogr44.c16>.
- Leardi, R., R. Boggia, and M. Terrile. 1992. "Genetic Algorithms as a Strategy for Feature Selection." *Journal of Chemometrics* 6, no. 5: 267–281. <https://doi.org/10.1002/cem.1180060506>.
- Magney, T. S., J. U. H. Eitel, D. R. Huggins, and L. A. Vierling. 2016. "Proximal NDVI Derived Phenology Improves In-Season Predictions of Wheat Quantity and Quality." *Agricultural and Forest Meteorology* 217: 46–60. <https://doi.org/10.1016/j.agrformet.2015.11.009>.
- Martin, G., K. Barth, M. Benoit, et al. 2020. "Potential of Multi-Species Livestock Farming to Improve the Sustainability of Livestock Farms: A Review." *Agricultural Systems* 181: 102821. <https://doi.org/10.1016/j.agsy.2020.102821>.
- Molle, G., M. Decandia, A. Cabiddu, S. Y. Landau, and A. Cannas. 2008. "An Update on the Nutrition of Dairy Sheep Grazing Mediterranean Pastures." *Small Ruminant Research* 77, no. 2–3: 93–112. <https://doi.org/10.1016/j.smallrumres.2008.03.003>.
- Murphy, D. J., B. O' Brien, M. O' Donovan, T. Condon, and M. D. Murphy. 2022. "A Near Infrared Spectroscopy Calibration for the Prediction of Fresh Grass Quality on Irish Pastures." *Information Processing in Agriculture* 9, no. 2: 243–253. <https://doi.org/10.1016/j.inpa.2021.04.012>.
- Muzzo, B. I., R. D. Ramsey, and J. J. Villalba. 2025. "Changes in Climate and Their Implications for Cattle Nutrition and Management." *Climate [Internet]* 13, no. 1: 1. <https://doi.org/10.3390/cli13010001>.
- Nørgaard, L., A. Saudland, J. Wagner, J. P. Nielsen, L. Munck, and S. B. Engelsen. 2000. "Interval Partial Least-Squares Regression (IPLS): A Comparative Chemometric Study With an Example From Near-Infrared Spectroscopy." *Applied Spectroscopy* 54, no. 3: 413–419. <https://doi.org/10.1366/0003702001949500>.
- Parrini, S., A. Acciaioli, A. Crovetto, and R. Bozzi. 2018. "Use of FT-NIRS for Determination of Chemical Components and Nutritional Value of Natural Pasture." *Italian Journal of Animal Science* 17, no. 1: 87–91. <https://doi.org/10.1080/1828051x.2017.1345659>.
- Pereira-Crespo, S., A. Botana, M. Veiga, et al. 2022. "Prediction of Quality of Total Mixed Ration for Dairy Cows by Near Infrared Reflectance Spectroscopy and Empirical Equations." *Journal of Applied Animal Research* 50, no. 1: 69–79. <https://doi.org/10.1080/09712119.2021.2022488>.
- Primi, R., G. Filibeck, A. Amici, et al. 2016. "From Landsat to Leafhoppers: A Multidisciplinary Approach for Sustainable Stocking Assessment and Ecological Monitoring in Mountain Grasslands." *Agriculture, Ecosystems and Environment* 234: 118–133. <https://doi.org/10.1016/j.agee.2016.04.028>.
- Pullanagari, R. R., I. J. Yule, M. P. Tuohy, M. J. Hedley, R. A. Dynes, and W. M. King. 2012. "In-Field Hyperspectral Proximal Sensing for Estimating Quality Parameters of Mixed Pasture." *Precision Agriculture* 13, no. 3: 351–369. <https://doi.org/10.1007/s11119-011-9251-4>.
- Rinnan, Å., F. Berg, and S. B. Engelsen. 2009. "Review of the Most Common Pre-Processing Techniques for Near-Infrared Spectra." *TrAC, Trends in Analytical Chemistry* 28, no. 10: 1201–1222. <https://doi.org/10.1016/j.trac.2009.07.007>.
- Rivero, M. J., and M. R. F. Lee. 2022. "A Perspective on Animal Welfare of Grazing Ruminants and Its Relationship With Sustainability." *Animal Production Science* 62, no. 18: 1739–1748. <https://doi.org/10.1071/an21516>.
- Samadi, S. Wajizah, and A. A. Munawar. 2018. "Fast and Simultaneous Prediction of Animal Feed Nutritive Values Using Near Infrared Reflectance Spectroscopy." In *IOP Conference Series: Earth and Environmental Science*, 012112. IOP Publishing. <https://doi.org/10.1088/1755-1315/122/1/012112>.
- Shalloo, L., M. O' Donovan, L. Leso, et al. 2018. "Review: Grass-Based Dairy Systems, Data and Precision Technologies." Supplement, *Animal* 12, no. s2: s262–s271. <https://doi.org/10.1017/s175173111800246x>.

- Shenk, J. S., and M. O. Westerhaus. 1994. "The Application of Near Infrared Reflectance Spectroscopy (NIRS) to Forage Analysis." In *Forage Quality, Evaluation, and Utilization*, edited by G. C. Fahey, 406–449. Wiley. <https://doi.org/10.2134/1994.foragequality.c10>.
- Shorten, P. R., S. R. Leath, J. Schmidt, and K. Ghamkhar. 2019. "Predicting the Quality of Ryegrass Using Hyperspectral Imaging." *Plant Methods* 15, no. 1: 1–12. <https://doi.org/10.1186/s13007-019-0448-2>.
- Van Soest, P. J., J. B. Robertson, and B. A. Lewis. 1991. "Methods for Dietary Fiber, Neutral Detergent Fiber, and Nonstarch Polysaccharides in Relation to Animal Nutrition." *Journal of Dairy Science* 74, no. 10: 3583–3597. [https://doi.org/10.3168/jds.S0022-0302\(91\)78551-2](https://doi.org/10.3168/jds.S0022-0302(91)78551-2).
- Teague, R., and U. Kreuter. 2020. "Managing Grazing to Restore Soil Health, Ecosystem Function, and Ecosystem Services." *Frontiers in Sustainable Food Systems* 4: 534187. <https://doi.org/10.3389/fsufs.2020.534187>.
- Workman, J., and L. Weyer. 2007. *Practical Guide to Interpretive Near-Infrared Spectroscopy*. 1st ed., 344. CRC Press. <https://doi.org/10.1201/9781420018318>.
- Wróbel, B., W. Zielewicz, and M. Staniak. 2023. "Challenges of Pasture Feeding Systems—Opportunities and Constraints." *Agriculture (London)* 13, no. 5: 974. <https://doi.org/10.3390/agriculture13050974>.