Computers and Electronics in Agriculture 251 (2026) 112081

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

# Estimating grass-clover mixtures quality using field spectroradiometry and stacking-based ensemble learning models

Shaohui Zhang [a,*], Martin Leon Gnyp [b], Ping Xin [a,*], Petra Junklewitz [b], Jörg Jasper [b], Chariton Kalaitzidis [c], Kiril Manevski [a]

[a] Department of Agroecology, Aarhus University, Blichers Allé 20, Tjele 8830, Denmark
[b] Research Centre Hanninghof, Yara International ASA, 48249 Dülmen, Germany
[c] Department of Geoinformation in Environmental Management, Mediterranean Agronomic Institute of Chania 73100 Chania, Crete, Greece

ARTICLE INFO

*Keywords:*
Grass quality
Stacking ensemble learning
Reflectance
Spectra derivatives
Standard Normal Variate

ABSTRACT

In this study, we designed a stacking ensemble machine learning method for accurate estimation of grass-clover biomass quality from high-resolution hyperspectral field data. Basic learners were support vector machine (SVM), K-Nearest Neighbor (KNN), random forest (RF), decision tree (DT), partial least square regress (PLS), shallow artificial- (ANN) and deep multi-layer perceptron (MLPs) neural networks. Three meta learners in the second layer were RF (stackingRF), eXtreme Gradient Boosting (stackingXGB), and CatBoost (stackingCat). We used 6 different datasets for the estimation, including the raw spectral reflectance, its 1st and 2nd derivatives, and their Standard Normal Variate (SNV). Quality indices were crude ash, crude protein, crude fat, crude fiber, neutral detergent fiber (NDF), acid detergent fiber (ADF), water-soluble carbohydrates (WSC), sugar, and fructans.

The results showed that spectral feature selection by the PLS was useful in preprocessing. All machine learning methods predicted well quality from the raw reflectance, with PLS performing best ($R^2$ = 0.64–0.87) for all indices except crude ash. The stacked ensemble learning benefited the predictions, especially with stackingRF and stackingCat as meta learners (R2 increased by 2–8 % compared with the average of the four best machine learning algorithms in the first layer). SNV generally performed better than the original reflectance across most datasets. In particular, SNV applied to the 2nd derivative, combined with either PLS or StackingRF, yielded the best predictions for carbohydrates (sugar, fructans, WSC) and fiber (crude fiber, NDF, ADF), achieving the highest $R^2$ values (0.74–0.87) and the lowest nRMSE (5–15 %) among all pre-processing options. The red-edge region (700–780 nm) is the most important for all quality indices, except for ADF.

This is among the first studies to systematically evaluate stacking ensemble learning, using the outputs of multiple machine-learning algorithms as inputs, for predicting biomass quality and supporting field management in complex grass–clover mixtures. PLS-based feature selection combined with stacking ensemble learning on second-derivative SNV spectra appears particularly promising and warrants testing at larger spatial and temporal scales.

## 1. Introduction

Grasslands account for a substantial portion of the terrestrial surface and play a crucial role in a range of biophysical and environmental processes, as well as function as a significant feed source for livestock (Bengtsson et al., 2019; O'Mara, 2012). Knowledge of their biochemical and nutrient composition is vital for effective decisions on management and assessments. In the past, laboratory spectroscopy techniques have emerged as alternatives to traditional methods for analyzing biomass quality indices (Bec et al., 2020). Among these methods, near-infrared spectroscopy (NIRS) covering the 450 to 2500 nm spectrum has been notably applied to assess biomass protein, fiber, lignin, and in vitro dry matter in grass and other crops (Vasseur et al., 2022). However, these analyses necessitate prior sample preparation and standardization, being affected by particle size and other external factors (Batten, 1998). Such preparation processes can also be time-intensive, costly, and

* Corresponding authors at: Department of Agroecology, Aarhus University, Blichers Allé 20, Tjele 8830, Denmark.
*E-mail addresses:* shaohui.zhang@agro.au.dk (S. Zhang), ping.xin@agro.au.dk (P. Xin).

https://doi.org/10.1016/j.compag.2026.112081
Received 22 April 2025; Received in revised form 10 April 2026; Accepted 16 June 2026
Available online 19 June 2026
0168-1699/© 2026 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

**Table 1**
Nitrogen management in the GLN field with grassland in Germany. Nitrogen form 1–10 is 0 – no nitrogen;1 – calcium ammonium nitrate (CAN) + sulphur (24 %N, 6 % S); 2 – ammonium sulphate (21 %N, 24 %S); 3 – ammonium sulphate + nitrification inhibitor (21 %N, 24 %S); 4 – urea (46 %N); 5 – urea + urease inhibitor (46 %N); 6 – ammonium sulphate (1st cut)/ CAN + sulphur to later cuts; 7 – ammonium sulphate (1st and 2nd cut) / CAN + sulphur to later cuts; 8 – ammonium sulphate (1st to 3rd cut) / CAN + sulphur to later cuts; 9 – CAN (27 %N); 10 – Ammonium sulphate nitrate (26 %N, 13 %S).

| | | | Nitrogen rate (kg/ha) | | | | | |
|---|---|---|---|---|---|---|---|---|
| Treatment | Nitrogen form | Total nitrogen rate (kg/ha) | Cut 1 | Cut 2 | Cut 3 | Cut 4 | Cut 5 | Organic fertilizer |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | slurry to 1st and second cut (28 $m^3$) |
| 2 | 10 | 177 | 62 | 35 | 35 | 27 | 18 | slurry to 1st and second cut (28 $m^3$) |
| 3 | 1 | 54 | 19 | 11 | 11 | 8 | 5 | slurry to 1st and second cut (28 $m^3$) |
| 4 | 1 | 177 | 62 | 35 | 35 | 27 | 18 | slurry to 1st and second cut (28 $m^3$) |
| 5 | 1 | 300 | 105 | 60 | 60 | 45 | 30 | slurry to 1st and second cut (28 $m^3$) |
| 6 | 1 | 422 | 148 | 84 | 84 | 63 | 42 | slurry to 1st and second cut (28 $m^3$) |
| 7 | 2 | 177 | 62 | 35 | 35 | 27 | 18 | slurry to 1st and second cut (28 $m^3$) |
| 8 | 3 | 54 | 19 | 11 | 11 | 8 | 5 | slurry to 1st and second cut (28 $m^3$) |
| 9 | 3 | 177 | 62 | 35 | 35 | 27 | 18 | slurry to 1st and second cut (28 $m^3$) |
| 10 | 3 | 422 | 148 | 84 | 84 | 63 | 42 | slurry to 1st and second cut (28 $m^3$) |
| 11 | 4 | 177 | 62 | 35 | 35 | 27 | 18 | slurry to 1st and second cut (28 $m^3$) |
| 12 | 5 | 54 | 19 | 11 | 11 | 8 | 5 | slurry to 1st and second cut (28 $m^3$) |
| 13 | 5 | 177 | 62 | 35 | 35 | 27 | 18 | slurry to 1st and second cut (28 $m^3$) |
| 14 | 5 | 422 | 148 | 84 | 84 | 63 | 42 | slurry to 1st and second cut (28 $m^3$) |
| 15 | 6 | 177 | 62 | 35 | 35 | 27 | 18 | slurry to 1st and second cut (28 $m^3$) |
| 16 | 7 | 177 | 62 | 35 | 35 | 27 | 18 | slurry to 1st and second cut (28 $m^3$) |
| 17 | 8 | 177 | 62 | 35 | 35 | 27 | 18 | slurry to 1st and second cut (28 $m^3$) |
| 18 | 9 | 177 | 62 | 35 | 35 | 27 | 18 | slurry to 1st and second cut (28 $m^3$) |
| 19 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | none |

laborious. In contrast, field spectroradiometry is a non-destructive analytical approach for real-time and accurate assessments of land cover, including grass quality (Fernández-Habas et al., 2022; Zhou et al., 2019). Portable handheld field spectroradiometers are used to collect canopy hyperspectral reflectance data to study properties and correlations with quality and biochemistry (Shaver et al., 2011; Cao et al., 2015; Fernández-Habas et al., 2022; Gámez et al., 2024). To ensure accurate retrieval of the desired parameters, data-driven modelling, often called "artificial intelligence", including machine learning (ML) offers a plethora of advanced algorithms to be tested.

Machine learning indeed becomes an essential tool in agronomy and environmental science, including biomass quality investigations from hyperspectral data. Support Vector Machines (SVM), Random Forests (RF), Partial Least Squares (PLS), and Artificial Neural Networks (ANN) facilitate the extraction of spectral features that correlate with important quality parameters, including dry matter and nutrient contents (Oliveira et al., 2020). A recent study effectively utilized the visible to near-infrared (VNIR) and shortwave infrared (SWIR) spectrum to assess various grass parameters with normalized root mean square errors (nRMSE) below 20 % (Oliveira et al., 2023). The typical workflow involves feature extraction, where relevant spectral features are identified, which in turn optimizes the dataset for model training. This step is particularly important because the number of spectral bands is extensive, autocorrelated and thus redundant. Selecting the most informative features enhances the performance of the model (Oliveira et al., 2023; Manevski et al., 2011, 2012). Moreover, ML-based algorithms can be ensembled into a learning process utilizing the unique strengths of each individual model for greater accuracy and generalization compared to the use of a single ML model. Ensemble learning, such as boosting, bagging, and stacking is *meta*-learning, i.e., "learning to learn" and aims to minimize model variances and overfitting and provide better predictions from complex datasets, such as hyperspectral, where the spectral features often correlate nonlinearly and intricately with observed properties such as biomass quality (Mutanga and Skidmore, 2004). Stacking combines the predictions of multiple "base" ML models into a *meta*-learner by assigning weights to each base model based on its performance (Du et al., 2024). The stacking process also creates new features derived from the predictions of the base models, providing further insights and improving the overall model performance (Wolpert, 1992; Liu et al., 2026).

Preprocessing the hyperspectral data is often useful for effective model training and prediction as it emphasizes spectral peaks and dips related to specific biochemical features of the biomass (Viña et al., 2011; Næs et al., 2002). It can also significantly reduce the influence of background vegetation on canopy spectral information, thus enhancing the correlation between spectral variables and target variables. For example, Passos et al. (2019) developed a range of differential spectra and their derivative transformations to estimate soluble solids content in pears. Similarly, Zhang et al. (2018) used derivative spectra to successfully estimate grass crude protein content. The findings of these and other comparable studies (Fernández-Habas et al., 2022; Fu et al., 2021; Gámez et al., 2024) indicate that differential spectral variables outperform other spectral variables for estimating biomass properties and can achieve high correlations ($R^2$ values exceeding 0.80) compared to a multivariate regression model. The 1$^{st}$ derivative technique effectively removes horizontal baselines of varying levels, highlighting points of inflection or areas where rapid changes occur in the hyperspectral signature. The 2$^{nd}$ derivative eliminates linear baselines of differing levels and slopes, emphasizing even more regions where the rate of change itself varies. A limitation of derivatives is their inability to correct for scatter or noise present in the data (Barnes et al., 1989). Scatter-the random deviations in the data can be present and can obscure meaningful trends or signals. To address these issues, derivative data can be combined with scatter correction methods, such as standard normal variate normalization (Næs et al., 2002).

This study aimed to evaluate the efficacy of spectral information obtained from a portable spectroradiometer in estimating biomass quality indices. The specific objectives were to: (i) preprocess the spectral reflectance to six different datasets comprising the original raw reflectance, reflectance normalized using the standard normal variate, the first derivative of reflectance, the first derivative of reflectance normalized with the standard normal variate, and the respective second derivatives, ii) select most relevant wavelengths by feature selection for estimating biomass quality; and (iii) develop a predictive model for grass quality indices that integrates base machine learning with stacked ensemble learning techniques.

## 2. Methods

### 2.1. Study site, experimental design, and field data collection

This study was conducted in two grassland fields coded "GLN" and

**Table 2**
Nitrogen management in the GLT field with grassland in Germany. Nitrogen form 9 denotes calcium ammonium nitrate (CAN; 27 %N).

| Treatment | Nitrogen form | Total nitrogen rate (kg/ha) | Nitrogen rate (kg/ha) Cut 1 | Cut 2 | Cut 3 | Cut 4 | Pasture species |
|---|---|---|---|---|---|---|---|
| 1 | 9 | 0 | 0 | 0 | 0 | 0 | Grass |
| 2 | 9 | 0 | 0 | 0 | 0 | 0 | Grass-Clover |
| 3 | 9 | 105 | 30 | 30 | 30 | 15 | Grass |
| 4 | 9 | 105 | 30 | 30 | 30 | 15 | Grass-Clover |
| 5 | 9 | 210 | 60 | 60 | 60 | 30 | Grass |
| 6 | 9 | 210 | 60 | 60 | 60 | 30 | Grass-Clover |
| 7 | 9 | 315 | 90 | 90 | 90 | 45 | Grass |
| 8 | 9 | 315 | 90 | 90 | 90 | 45 | Grass-Clover |
| 9 | 9 | 420 | 120 | 120 | 120 | 60 | Grass |
| 10 | 9 | 420 | 120 | 120 | 120 | 60 | Grass-Clover |

**Table 3**
Timing of fertilizer, harvest, and hyperspectral field data collection for GLN and GLT grassland fields in Germany. Dash denotes no field sampling and hyperspectral data collection due to drought.

| Field | Cut | Date Fertilizer | Harvest | Biomass sampling | Hyperspectral data collection |
|---|---|---|---|---|---|
| GLN | Cut1 | 03/28/2023 | 05/03/2023 | 05/03/2023 | 05/03/2023 |
| GLN | Cut2 | 05/11/2023 | 05/31/2023 | 05/31/2023 | 05/30/2023 |
| GLN | Cut3 | 06/05/2023 | 07/10/2023 | 07/10/2023 | 07/10/2023 |
| GLN | Cut4 | 07/17/2023 | 08/14/2023 | 08/14/2023 | 08/14/2023 |
| GLN | Cut5 | 08/18/2023 | 10/11/2023 | 10/11/2023 | 10/11/2023 |
| GLT | Cut1 | 03/13/2023 | 04/26/2023 | 04/26/2023 | 04/26/2023 |
| GLT | Cut2 | 05/19/2023 | 05/17/2023 | – | – |
| GLT | Cut3 | 07/30/2023 | 08/14/2023 | 08/14/2023 | 08/14/2023 |
| GLT | Cut4 | 08/18/2023 | 09/25/2023 | 09/25/2023 | 09/19/2023 |

"GLT" (51°50'N, 7°15'E) located at the Yara Hanninghof research center (Yara International ASA, Oslo, Norway) in Dülmen, Germany. Sowing for the GLT fields occurred in September 2022. The GLN field, characterized by perennial grassland with a sward age of over 5 years, was also established in 2022 and is routinely reseeded. Effective data collection for both fields took place in 2023. The GLN field consisted of plots mainly with pure grass (ryegrass-dominated) without legumes, studied at different annual nitrogen fertilizer levels (0, 54, 177, 300, 422 kg $ha^{-1}$) and different fertilizer types (inc. nil and organic), totally 19 treatments with 4 replicates (Table 1). In GLT, grass-clover mixture was also treated with different nitrogen fertilizer levels (0, 105, 210, 315, 420 Kg $ha^{-1}$), in total of 10 treatments with 4 replicates (Table 2). The plot size in both the GLT and GLN fields was 3 × 12 m. The proportion of clover in the mixture in the GLT field decreased with increasing N fertilization and as the growing season progressed. The GLT field was cut 4 times in 2023, with no data collection in the second cut due to severe drought; the GLN field was cut 5 times in 2023 (Table 3). In total, 500 biomass data points were collected.

### 2.2. Determination of biomass quality indices

For each harvest time, aboveground biomass samples were analyzed for crude ash, crude protein, crude fat, crude fiber, neutral- (NDF) and acid-detergent fiber (ADF), water soluble carbohydrates (WSC), sugar, and fructans. These quality parameters were analyzed at Landwirtschaftliche Kommunikations- und Servicegesellschaft's laboratory following the VDLUFA methods (Naumann et al., 1997) for NIRS-chemical analysis of feedstuff. The front probe was illuminated by a beam of light with a well-defined wavelength (380 to 1690 nm), generating a reflective measurement point from a wavelength of 440 nm. The results of the measurement are based on the quantitative analysis of water and organic matter. Dieterle et al. (2003) indicated that the correlation between NIRS measurements and reference methods helps to create statistical results that are quality and verified by rigorous, irresistible tests.

### 2.3. Field spectroscopy data collection and measurement geometry

A handheld field spectroradiometer (*tec5USA's HandySpec,* Steinbach, Germany) was used to measure canopy hyperspectral reflectance in the 500–1700 nm spectrum shortly before each harvest. The device consists of two sensors with slightly different resolutions, i.e., 2.6 nm for the VNIR range 400–1000 nm and 6 nm for the NIR range 1000–1700 nm. The final spectral resolution was resampled to 10 nm to simplify the calculation of the inter-band correlation matrix and improve visualization and interpretation of the NDI contour maps, as well as to reduce computation time for the deep multi-layer perceptron (MLP) models. For each plot, the spectroradiometer head was positioned above the central part of the plot so that its circular field of view was filled by the grass canopy (no soil, tramlines, or neighboring plots). Grass canopy reflectances were measured on a plant-by-plant basis under clear-sky conditions, within two hours of local solar noon. The pistol was held perpendicular to the center of the grass canopy at a constant distance of 50–80 cm. We randomly acquired 15–20 individual spectra per plot, moving the sensor slightly from one plot side to the other side to capture small-scale spatial variability while staying well inside the plot boundaries. These 15–20 spectra were then averaged to represent the mean spectral reflectance of that plot. The device consists of two optical heads: one pointed upward to the sky to quantify incoming irradiance (reference channel), and one pointed downward at the canopy with a nadir view angle of 0° (canopy channel), ensuring a vertical viewing geometry into the grass sward. Measurements were conducted under stable daylight conditions, avoiding periods with rapidly changing cloud cover. Before each plot measurement, the system was calibrated using a 30 % grey reference panel (SphereOptics, Herrsching, Germany) to compensate for changes in ambient illumination. To account for the strong influence of water on canopy spectra, measurements were taken on rain-free days after dew had evaporated, and plots with visible surface water or wet soil patches were excluded. In addition, spectra with abnormal shape or reflectance outside the expected physical range were discarded as part of the quality-assurance procedure, ensuring that only high-quality reflectance data were used for model calibration and validation. The collected spectra were compiled into a dataset-spectral library and stored for further analyses, i.e., pre- and main processing.

### 2.4. Spectra preprocessing and PLS feature selection

The spectral data of this study were centered for wavelengths at their mean, but not scaled, using standard normal variate (SNV) as described by Barnes et al. (1989). The SNV pre-processing correction was applied to the raw reflectance, its 1st and 2nd derivatives, yielding a total of 6 preprocessed dataset (including the raw spectra and their two derivatives before SNV transformation). SNV transformation was applied spectrum–wise. For each spectrum, we subtracted the mean reflectance across all wavelengths and divided by the corresponding standard

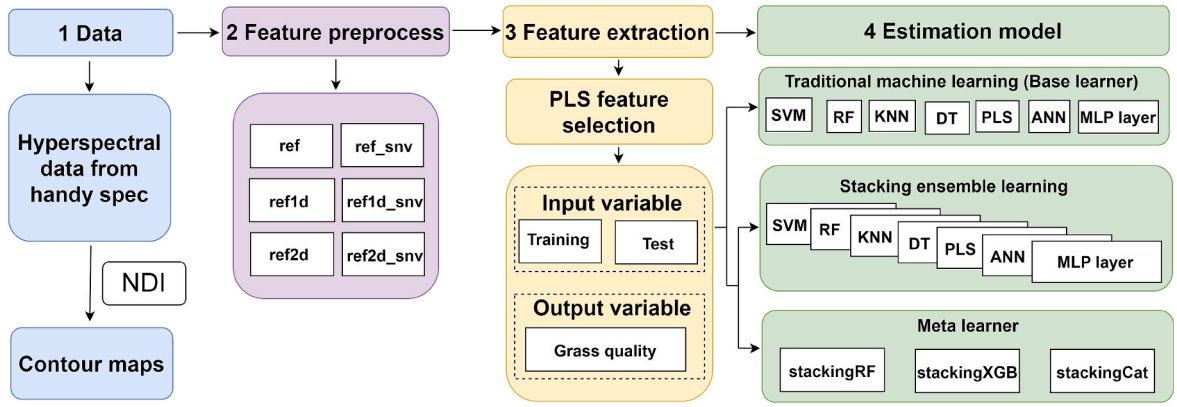


**Fig. 1.** Workflow in the stacking ensembled model framework for grass biomass quality estimation. The collected hyperspectral field data were preliminary used to create contour maps correlating Normalized Difference Index (NDI) and the measured biophysical parameters. The hyperspectral field data were preprocessed to six datasets: original reflectance (ref), its 1st (ref1d) and 2nd derivatives (ref2d), and their standard normal variate (ref_snv, ref1d_snv and ref2d_snv). Each of these spectral libraries undergo spectral feature selection by partial least square regression (PLS) for dimension reduction, to further test several estimation models in a first layer: support vector machine (SVM), K-Nearest Neighbor (KNN), random forest (RF), decision tree (DT), PLS, shallow artificial neural networks (ANN) and deep multi-layer perceptron neural networks (MLP, including 2–6 layers, MLP layer2-MLP layer6). Their predictions were combined into three "stacking" *meta*-learners in a second layer to boost predictability: random forest (stackingRF), eXtreme Gradient Boosting (stackingXGB), and CatBoost (stackingCat).

deviation (no weighting or window was used). First and second derivatives of the spectra were calculated using a Savitzky–Golay filter with a 2nd–order polynomial. For the first derivative we used a window length of 11 bands, and for the second derivative a window length of 13 bands, applied to the (optionally SNV–transformed) reflectance prior to modelling. The original reflectance also compiled for contouring the correlation of all two-bands combination with the biomass quality parameters using the Normalized Difference Index (NDI; Cao et al., 2017):

$$NDI = \frac{R\lambda 1 - R\lambda 2}{R\lambda 1 + R\lambda 2} \quad (1)$$

where Rλ1 and Rλ2 are the reflectance values at wavebands λ1 and λ2 in the 500–1700 nm range. This yielded a total number of 11,664 combinations to be investigated.

Main processing: In order to decrease the number of collinear variables in the spectra (neighbour wavelengths that vary the same way), we applied a 'wrapper method' to perform PLS feature selection to all data sets. Wrapper methods involve iteratively filtering features and refitting the model to optimize parameters. This kind of procedure has been shown to improve the model's predictability by removing redundant and noisy features from the data (Passos et al., 2019).

### 2.5. Base machine learning algorithms screening and selection

Several ML algorithms were tested for their predictive power, including SVM, K-Nearest Neighbor (KNN), RF, Decision Tree (DT), PLS, shallow ANNs, MLP (including 2 to 6 layers in this study, abbreviated MLP layer2 to MLP layer6, respectively). For the SVM, the grid search and cross-validation were selected by the penalty parameter of the error term C and the kernel of the SVM from given values. For the PLS, calibration was done by calculating the huber loss 5-fold cross-validation and then testing with the components. For the KNN, the n_neighbors, weights, and algorithm were selected by the grid search and cross-validation from given values. For the DT, max_depth, max_features, and n_estimators were selected by the grid search and cross-validation from given values. For ANNs and MLP, activation, alpha, hidden_layer_sizes, learning_rate, and solvers were selected by the grid search and cross-validation from given values. The hyperparameters for each traditional ML algorithm (SVM, KNN, RF, DT, PLS, ANN, MLP layer2 to MLP layer6) are all determined through 5-fold cross-validation for the training datasets. Based on the preliminary runs, three base learners, i.e., RF, eXtreme Gradient Boosting (XGB), and gradient boosting on decision trees (CatBoost) were used as stacking meta learners, abbreviated, respectively, stackingRF, stackingXGB and stackingCat.

### 2.6. Ensembled stacking learning method and models evaluation

The workflow of the proposed stacking learning method is shown in Fig. 1. The main processing includes several steps:

(1) Build contour maps of the correlation coefficient ($R^2$) between grass quality indices and NDI by using the original spectral reflectance.

(2) Feature preprocess including applied the Standard Normal Variate (SNV) correction to the

reflectance, 1st and 2nd derivatives spectra, yielding a total of 6 preprocessed data types (the original.

absorbance spectra, the 1st and 2nd derivatives, and their SNV counterparts).

(3) Feature main preprocessing by feature selection using PLS to discard the lowest correlation wavelengths. These wavelengths are the ones that typically worsen the quality of the model training and validation, therefore discarding them effectively improves the metrics associated with the main prediction.

(4) Model training (1) Data partitioning: The original dataset (500 collected samples) is randomly divided into training and validation sets in a 3:1 ratio; (2) Cross-validation: the training set is divided into five equal parts, with one part used as the test set and the remaining four parts used as the training set. The hyperparameters for each traditional ML algorithm (SVM, KNN, RF, DT, PLS, ANN, MLP layer2 to MLP layer6) are determined through 5-fold cross-validation.

(5) Base learner prediction (first layer): Eleven machine-learning algorithms were applied separately to the PLS-selected feature dataset as model input, using a randomly selected 75 % of the total samples for training.

(6) Meta learner prediction (second layer) of stacking nature from Random Forest (stackingRF), eXtreme Gradient Boosting (stackingXGB), and CatBoost (stackingCat). The best predictions from four of the eleven first-layer base learners were combined to form a new training set for the meta-model. This meta-model was then trained on that set, and the stacking ensemble was evaluated on the remaining 25 % of the data used as an independent test set. The performance of each ML and stacking ensembled models was evaluated for the test dataset by the coefficient of determination ($R^2$), root mean square error (RMSE), and its normalized value (nRMSE).

**Table 4**
Average contents, standard deviation (SD), range, and coefficients of variation (CV) of quality indices (g $Kg^{-1}$ DM) in grass-clover samples from GLN and GLT fields. n is the sample number.

| Quality indices (g $Kg^{-1}$ DM) | GLN(n = 360) | | | | GLT(n = 120) | | | |
|---|---|---|---|---|---|---|---|---|
| | Mean | Range | SD | CV (%) | Mean | Range | SD | CV (%) |
| Crude ash | 95.5 | 78.4–120.6 | 7.3 | 7.6 | 95.1 | 86.7–106.9 | 3.7 | 3.9 |
| Crude protein | 152.7 | 94.6–259.3 | 32.3 | 21.2 | 153.4 | 112.7–239 | 27.9 | 18.2 |
| Crude fat | 34.1 | 21–44.9 | 4.1 | 12.1 | 33.6 | 31–35.3 | 0.9 | 2.7 |
| Crude fiber | 239.6 | 193.6–277.5 | 14.9 | 6.2 | 240.4 | 215.3–253.6 | 6.5 | 2.7 |
| NDF | 506.6 | 398.6–577.2 | 32.7 | 6.5 | 506 | 463.4–527.5 | 11 | 2.2 |
| ADF | 277.9 | 231.6–318.5 | 18.5 | 6.6 | 278.5 | 254.3–294.7 | 7.5 | 2.7 |
| Sugar | 131.8 | 51.9–261.2 | 40.2 | 30.5 | 132.9 | 85.4–175.5 | 18.1 | 13.6 |
| Fructans | 68.6 | 27.5–176.9 | 22.9 | 33.4 | 68.9 | 41–113.5 | 15.1 | 21.9 |
| WSC | 200.4 | 84.9–419 | 60.2 | 30.1 | 201.8 | 128–289 | 32.5 | 16.1 |

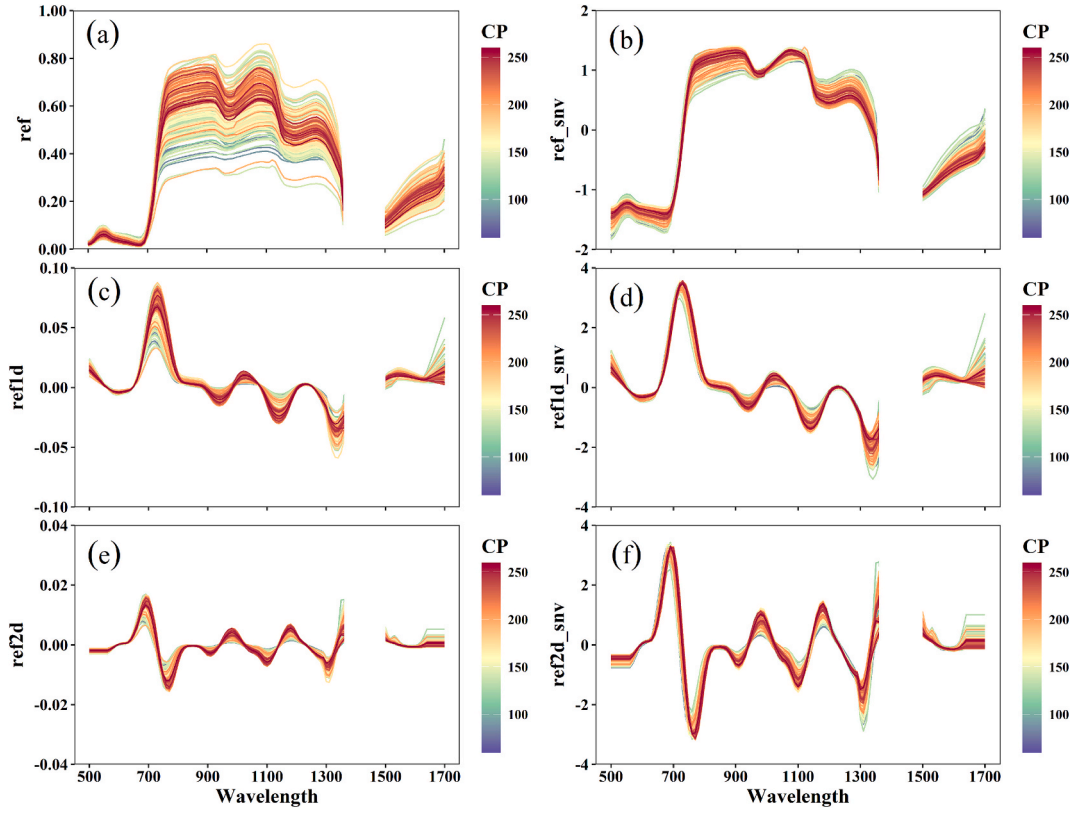


**Fig. 2.** Grassland spectral libraries of (a) ref: original reflectance, (b) ref_snv: its standard normal variate, (c) ref1d: 1st derivative and (d) ref1d_snv: 1st derivative standard normal variate, (e) ref2d: 2nd derivative and (f) ref2d_snv: 2nd derivative standard normal variate. CP is crude protein. Wavelengths of 1360–1500 nm were removed due to noise from leaf water and air water vapor.

## 3. Results

### 3.1. View on the grassland quality indices and spectral datasets

The Average contents, standard deviation (SD), range, and coefficients of variation (CV) of quality indices in grass-clover samples from GLN and GLT fields are presented in Table 4. A wide range of variation was observed for crude protein and carbohydrate, considering the samples in both fields, especially the highest coefficients of variation were found at all three carbohydrate contents. Fig. 2 shows the 6 different spectral libraries, including the raw and the derivatives reflectance and their SNV obtained from grassland plots differing in crude protein (CP) contents as an example. The SNV considerably reduced the spectral variation as it removed the multiplicative interferences of scatter and surface roughness/particle size; however, this variation is not necessarily undesirable in spectral discrimination and therefore relevant to compare with their non-SNV transformed spectra. Also from Fig. 2, it is evident that a non-linear relationship exists between reflectance values and CP, with peak concentrations observed at intermediate reflectance levels, prompting the need for effective assessment tools.

### 3.2. Correlation between two-band hyperspectral indices and grass biomass quality indices

To initially identify optimal spectral indices for predicting biomass quality indices, a two-band difference spectral index (NDI) was estimated using a calibration dataset, and optimal indices were defined as spectral ranges exhibiting the highest $R^2$ value when correlating the index with the measured indices. The resulting contour map (Fig. 3) indicated different spectral bands for different biochemical indices.

For instance, no spectral regions were identified strong enough to predict grass crude ash (CA) and CP, but crude fat, fiber (crude fiber, NDF, ADF), and carbohydrates content (sugar, fructans, WSC) could be

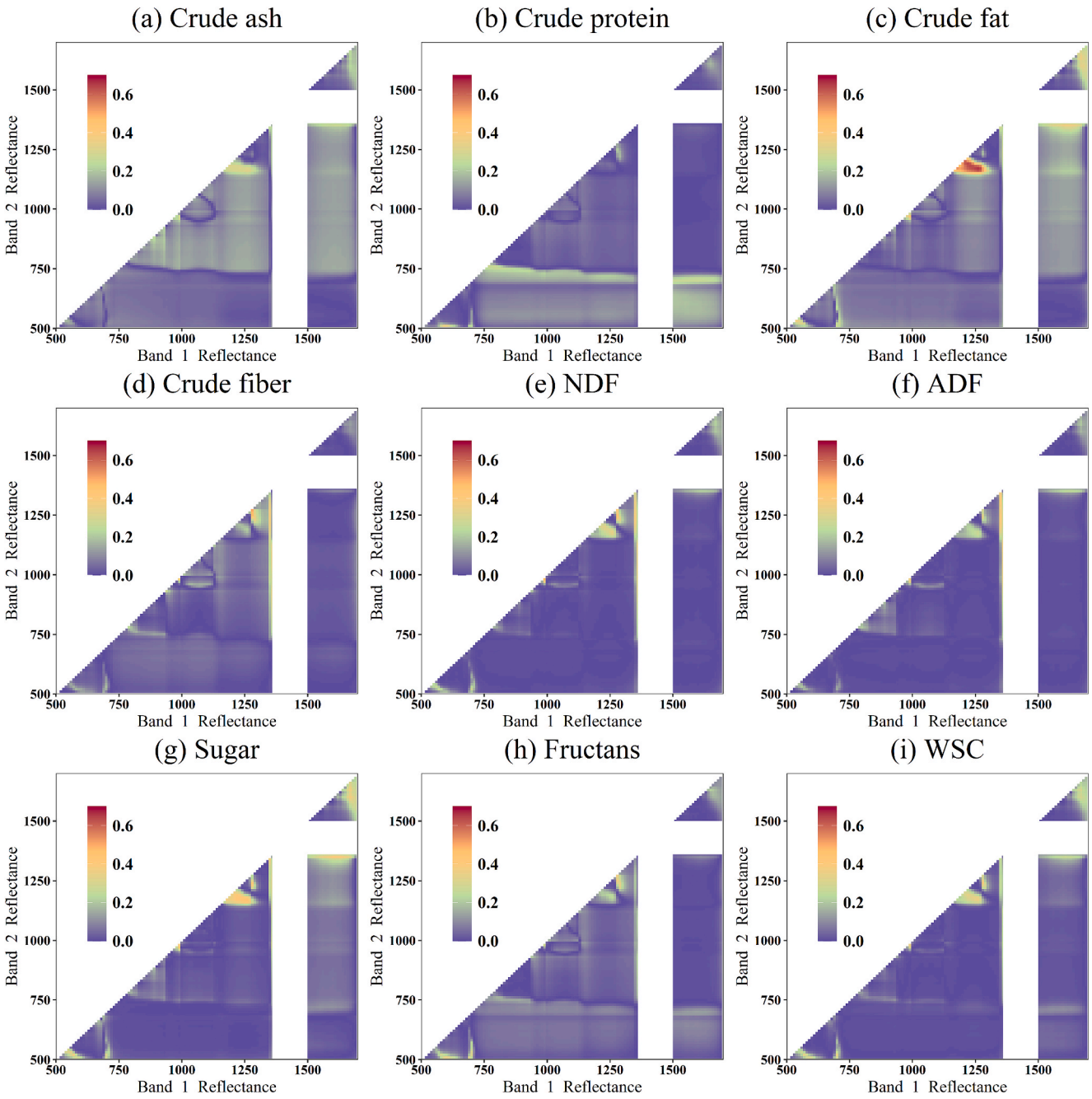


**Fig. 3.** Contour maps showing correlation coefficient ($R^2$) between grass quality indices and all possible normalized difference indices (NDI) calculated by the hyperspectral field data acquired at harvest. Wavelength at 1360–1500 nm was removed due to noise. Only results bellow matrix diagonal are shown as $R^2$ values are symmetrical. NDF is neutral detergent fiber, ADF is acid detergent fiber, WSC is water-soluble carbohydrates.

predicted mostly around 1150 to 1360 nm. While the contour map of $R^2$ depicts optimal spectral band combination for NDI predicting biomass quality parameters, it still contains redundant spectral information and requires focus on only the most relevant spectral features with the highest predicting power.

### 3.3. *Dimension reduction through spectral feature selection*

The Fig. 4 illustrates the PLS selected wavebands for each quality parameter of the 6 different spectral datasets (ref, ref_snv, ref1d, ref1d_snv, ref2d, ref2d_snv). The results mostly showed great variation in the selected bands as a function of spectral data input and parameters evaluated. However, common broad spectral regions could be seen in some cases for most parameters, such as those between 1200 and 1300 nm and around 1600 nm. The summary of the total numbers of selected wavelengths by PLS on those 6 datasets is shown in the Table. 5. For the majority of preprocessed datasets, the PLS feature selection reduced the nRMSE except for WSC in the ref1d dataset, fructans and WSC in the ref1d_snv dataset (Fig. 5). PLS feature selection proved to improve not only the PLS model metrics but also the other tested models. Based on this, all the results presented from this point onward are based on data sets that were pre-processed with PLS feature selection. The importance ranking of spectral wavelengths for the original reflectance dataset is shown in Fig. 6. The red-edge region (700–780 nm) is the most important for most quality indices, except for ADF, for which the 1200–1700 nm range is most relevant. The visible region (500–700 nm) is also

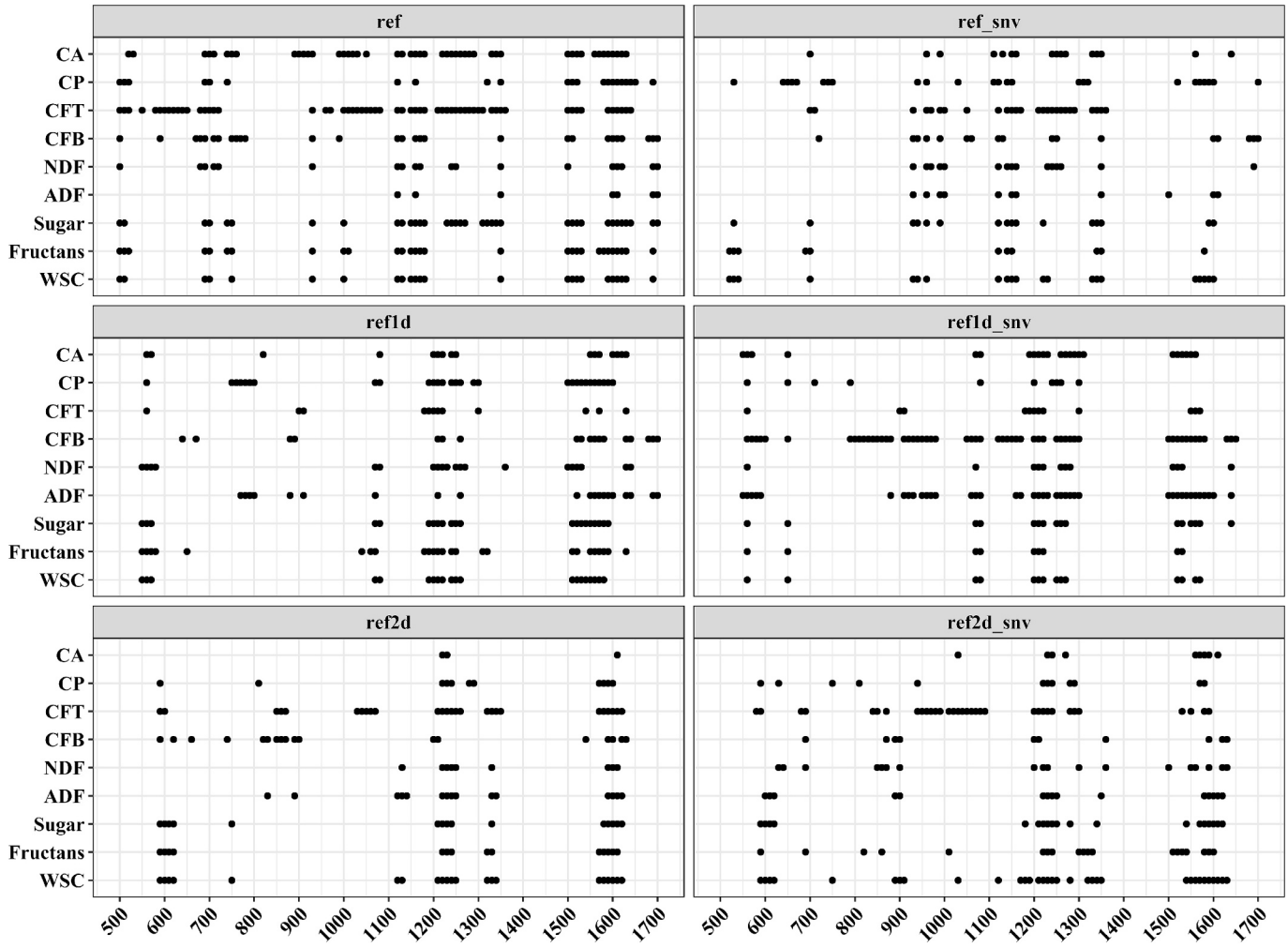


**Fig. 4.** Feature selection by Partial Least Squares conducted on the original reflectance (ref), its 1st (ref1d) and 2nd derivatives (ref2d), and their standard normal variate (ref_snv, ref1d_snv and ref2d_snv). Dots show wavelengths selected for each quality parameter. Only the selected wavelengths are used for model optimizations. Wavelengths of 1360–1500 nm were removed due to noise from leaf water and air water vapor. The total number of selected wavelengths is shown in Table 5.

**Table 5**
Summary of total numbers of selected wavelengths by Partial Least Squares conducted on the original reflectance (ref), its 1st (ref1d) and 2nd derivatives (ref2d), and their standard normal variate (ref_snv, ref1d_snv and ref2d_snv).

| Quality | ref | ref_snv | ref1d | ref1d_snv | ref2d | ref2d_snv |
|---|---|---|---|---|---|---|
| ADF | 7 | 11 | 20 | 40 | 15 | 15 |
| CA | 48 | 16 | 16 | 23 | 3 | 9 |
| CFB | 28 | 17 | 18 | 55 | 18 | 10 |
| CFT | 60 | 26 | 12 | 12 | 26 | 34 |
| CP | 22 | 25 | 29 | 10 | 11 | 12 |
| Fructans | 29 | 11 | 25 | 9 | 14 | 19 |
| NDF | 19 | 15 | 20 | 12 | 9 | 18 |
| Sugar | 36 | 16 | 21 | 16 | 15 | 19 |
| WSC | 24 | 21 | 20 | 14 | 21 | 33 |

important for crude fiber, NDF, fructans, and WSC. The NIR (Near infrared: 780–1360 nm) and SWIR region (Short-wavelength infrared: 1500–1700 nm) are related to all the quality indices (Fig. 4; Fig. 6).

### 3.4. Comparison of different quality indices predicted by machine learning and stacking learning

The performance of ML and stacking learning for quality indices estimation by using original reflectance (ref) is shown in Fig. 7. SVM and PLS were among the most stable algorithms in the base models, with $R^2$ higher than 0.6 except for crude ash. Also, nRMSE was lower than 20 % for all indices when using the best models for estimation. Fiber (including crude fiber, NDF, ADF) and carbohydrate (sugar, fructan, WSC) were predicted very well when using PLS, ANN, MLP layer2 and ensembled stacking models (R2 > 0.8, nRMSE < 18 %).

Comparison across algorithms and datasets showed that CA and also often CP were the two quality indices the most difficult to predict with good accuracy regardless of dataset and algorithm (Fig. 8). Otherwise, the rest of the quality indices could be predicted fairly well with most datasets and algorithms, with a clear tendency for the stacking meta learner algorithms to increase accuracy, especially stakingRF and stackingCat. There were a few exceptions, such as the low predictive power of ANN for almost all quality indices when using SNV-modified reflectance (Fig. 8).

### 3.5. Comparison of grass quality estimation performance using different datasets and stacking ensemble learning

Calculating 1st and 2nd derivatives and standard normal variates (SNV) are very common and effective methods for spectroscopy analysis. In this study, we found that there is a considerable difference between those datasets in most quality indices, while there are no significant differences in WSC in stackingRF (Fig. 9). In most cases, using

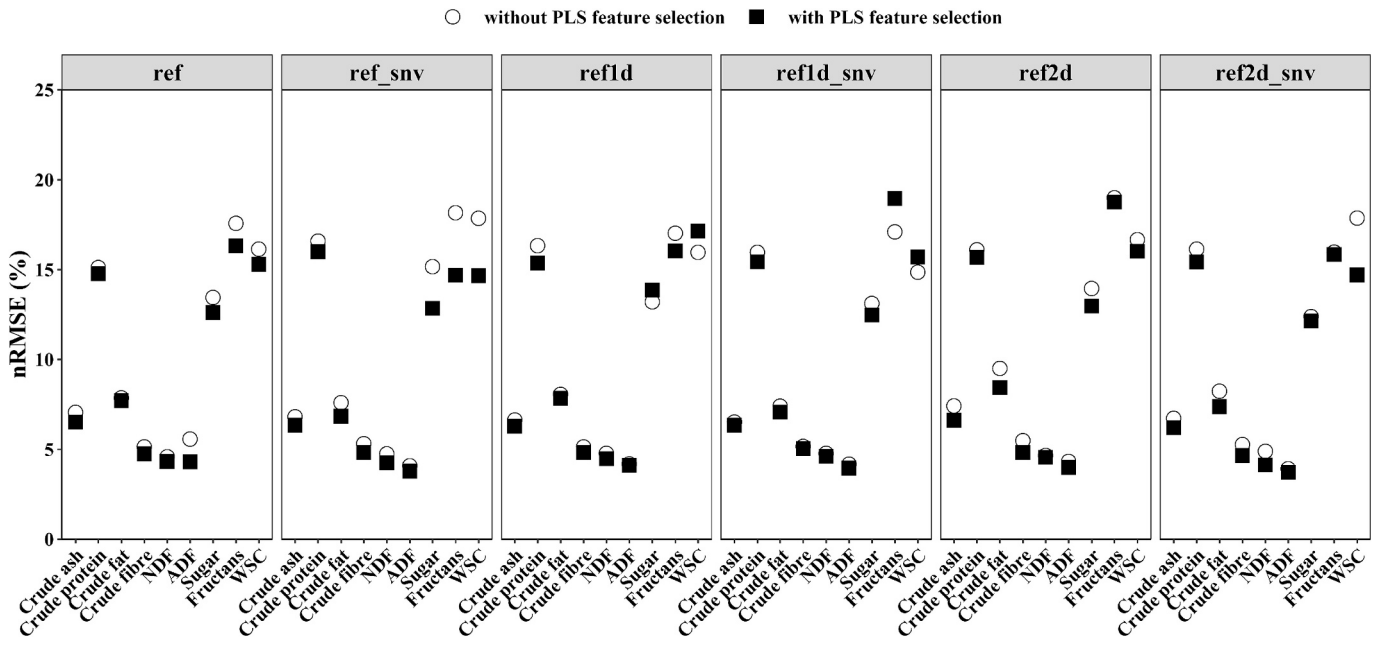


**Fig. 5.** Comparison of nRMSE across all reflectance datasets, with and without the PLS feature selection procedure, for PLS estimation of the different quality indices. The wavelength after PLS feature selection is shown in Fig. 4.

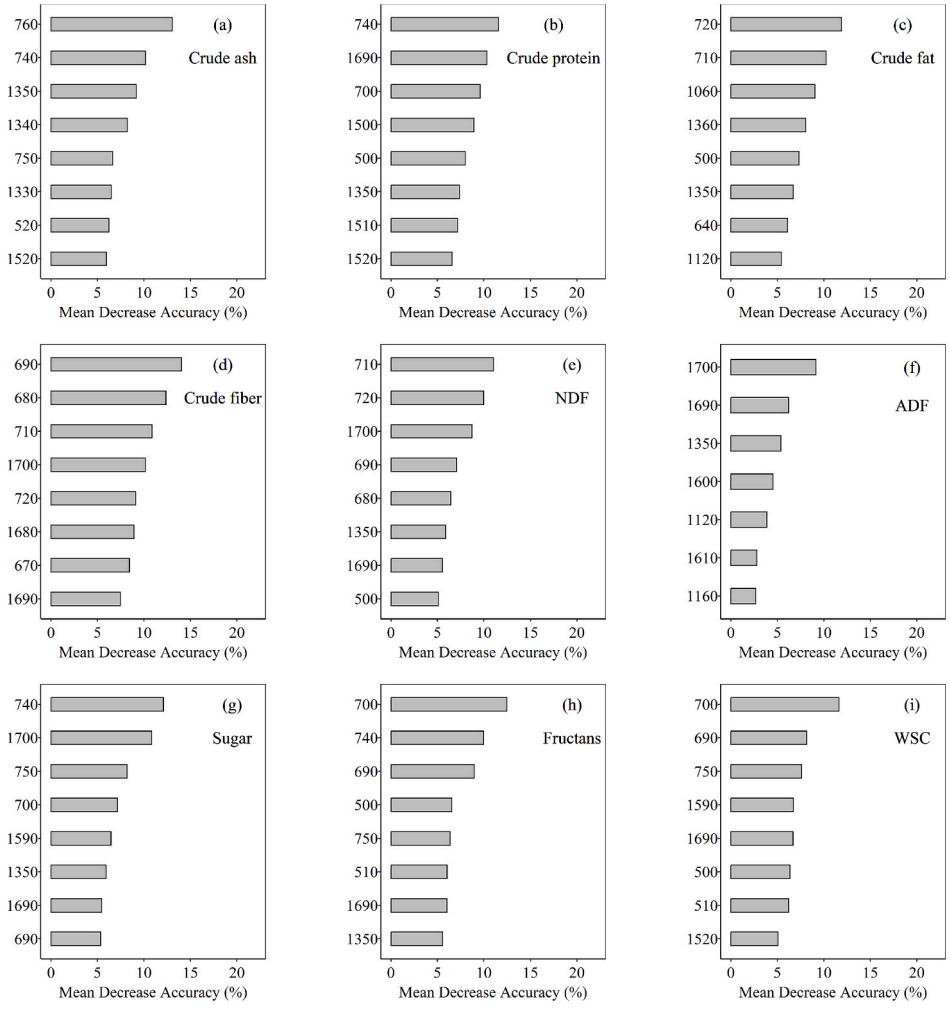


**Fig. 6.** Wavelength importance rank according to the random forest (RF) analysis after the feature selection using original reflectance (ref) in explaining grass quality. NDF is neutral detergent fiber, ADF is acid detergent fiber, WSC is water-soluble carbohydrates.

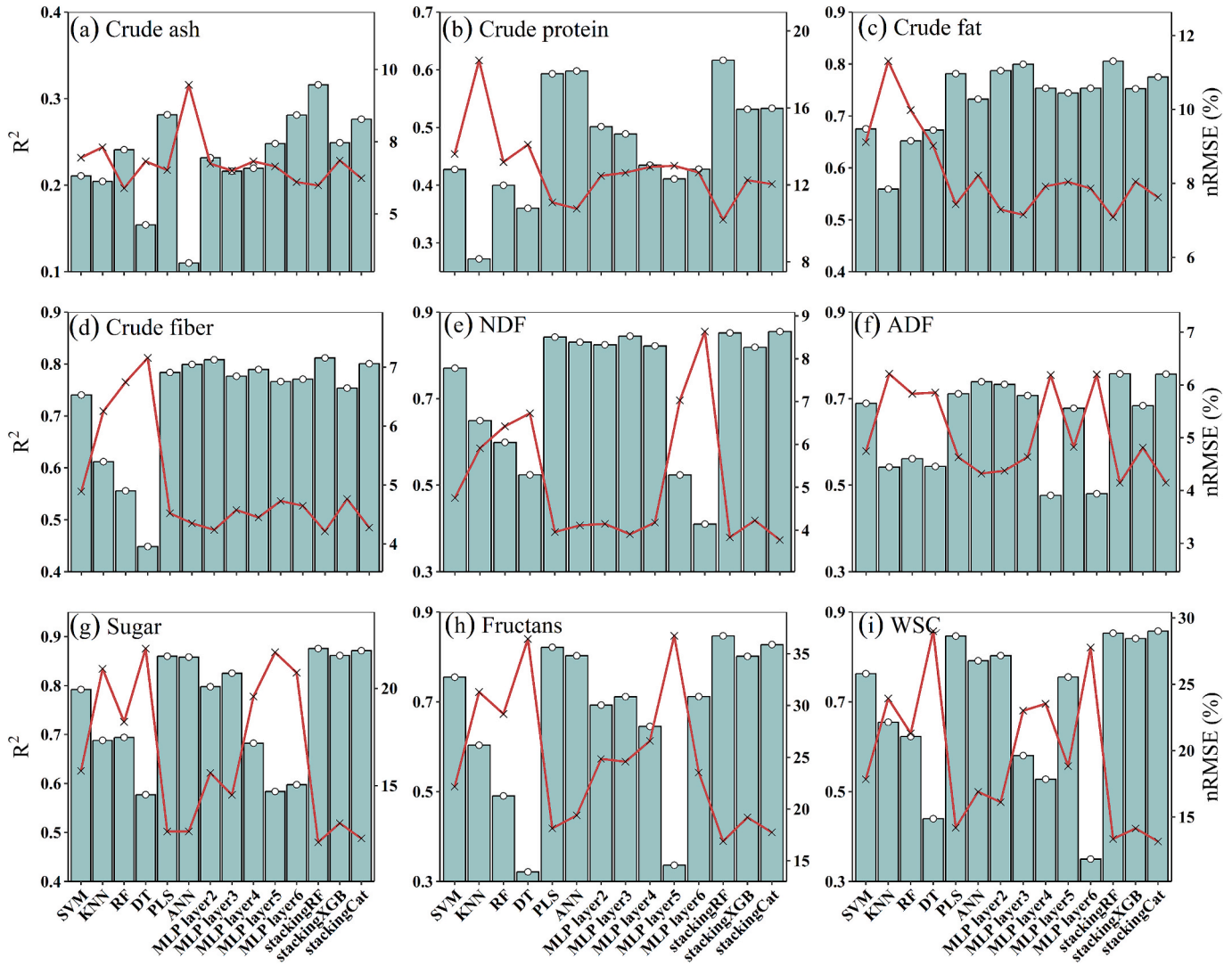


**Fig. 7.** Performance of base machine learning methods and stacking ensembled learning models on grass biomass quality estimation by original reflectance. Bars show correlation coefficients ($R^2$) and lines are nRMSE values. Algorithms in the first layer are: support vector machine (SVM), K-Nearest Neighbor (KNN), random forest (RF), decision tree (DT), partial least square (PLS) regression, shallow artificial neural networks (ANN) and deep multi-layer perceptron neural networks (MLP). Meta learner algorithms in the second layer are Random Forest (stackingRF), eXtreme Gradient Boosting (stackingXGB), and CatBoost (stackingCat).

the standard normal variate of ref, ref1d, and ref2d was better than without using SNV in the quality prediction. The performance was increased when using the 1st derivative and 2nd derivative of the reflectance compared with original reflectance in fiber prediction by PLS (Fig. 9). While the performance was decreased when using the 1st derivative and 2nd derivative of the reflectance compared with without using derivation in fiber prediction by stackingRF (Fig. 9). Stacking ensemble learning with StackingRF significantly outperformed PLS in terms of $R^2$ for most quality parameters, except for crude fiber and ADF, across the different datasets (Fig. 10). In terms of nRMSE, StackingRF also performed significantly better than PLS for all quality parameters except crude ash and sugar.

## 4. Discussion

This study proposed stacking machine learning (stacking-ML) for improved estimation of biomass quality from different hyperspectral datasets- original field hyperspectral reflectance, its 1st and 2nd derivatives, and their SNV counterparts. The stacking-ML combined seven ML models (SVM, KNN, RF, DT, PLS, ANN, and MLP) to improve the accuracy and effectiveness of the estimation. The model's performance was thoroughly evaluated for different biomass quality indices.

### 4.1. Wavelength feature selection and comparison of machine learning models

The dataset underwent preprocessing through PLS feature selection, revealing that selected wavelengths yielded more accurate predictions for grass quality than the entire wavelength spectrum (Fig. 5). The significant wavelengths chosen for each quality indices reflect their relative importance, as determined by the PLS algorithm. Leaf and canopy systems absorb a substantial portion of the visible spectrum (400 to 700 nm) due to the chlorophyll and other leaf pigments that capture most of the incident radiation (Luo et al., 2016). As plants mature, chlorophyll typically declines while fiber components (e.g. NDF, crude fiber) increase, so changes in reflectance in the visible region (500–700 nm) can indirectly track shifts in nutrient composition (Wijesingha et al., 2020). In parallel, reflectance in the NIR–SWIR range (approximately 1350, 1520–1700 nm in this study; Fig. 6) is sensitive to distinct organic components of structural and water–soluble carbohydrates (cellulose and hemicellulose), which play key roles in osmotic adjustment and stress responses, allowing these bands to capture carbohydrate–related

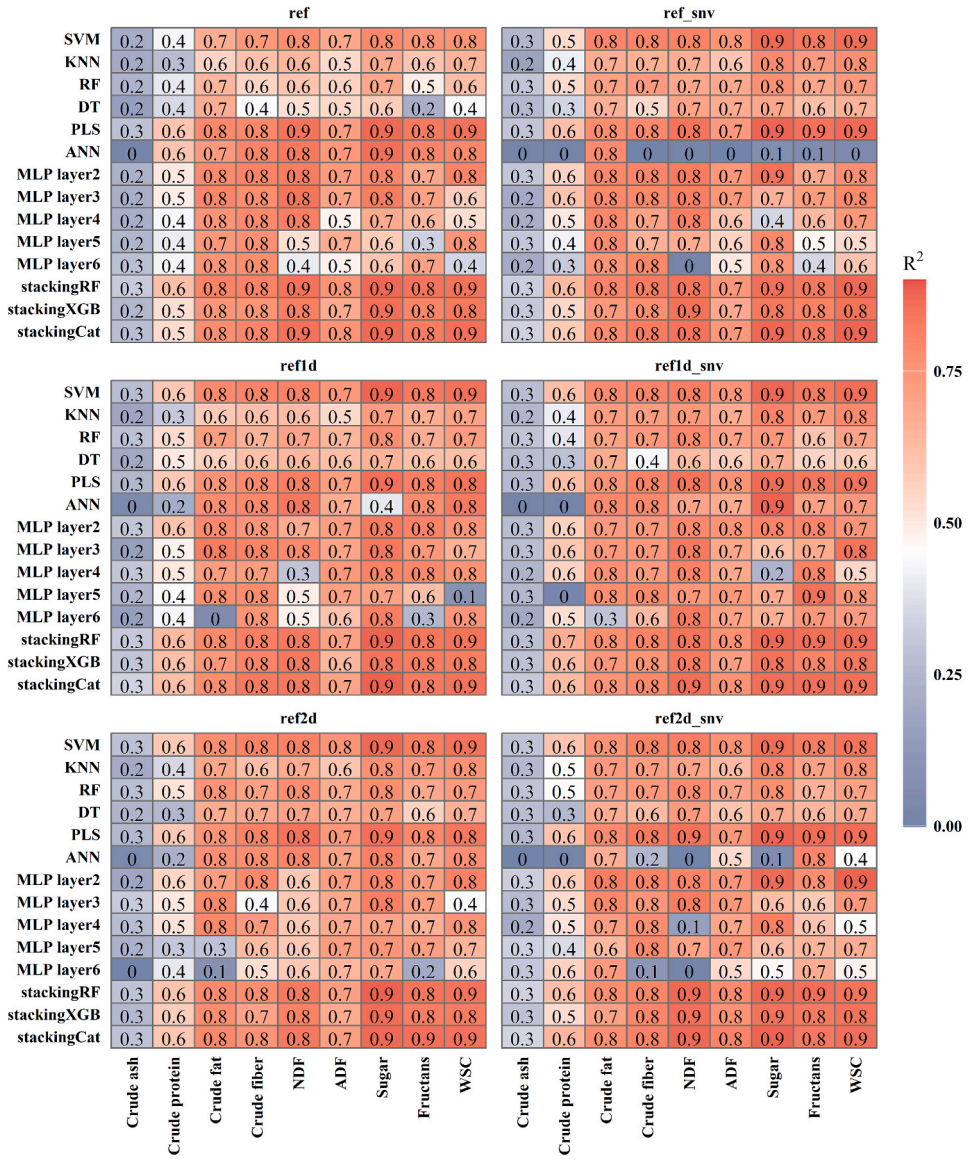


**Fig. 8.** Heatmap of correlation coefficients ($R^2$) for base machine learning methods and stacking ensembled learning models on grass biomass quality estimation by original reflectance (ref), its 1st (ref1d) and 2nd derivatives (ref2d), and their standard normal variate (ref_snv, ref1d_snv and ref2d_snv). Algorithms in the first layer are: support vector machine (SVM), K-Nearest Neighbor (KNN), random forest (RF), decision tree (DT), partial least square (PLS) regression, shallow artificial neural networks (ANN) and deep multi-layer perceptron neural networks (MLP). Meta learner algorithms in the second layer are Random Forest (stackingRF), eXtreme Gradient Boosting (stackingXGB), and CatBoost (stackingCat).

physiological changes (Ertlen et al., 2010; Fonseca et al., 2020). Johnson (2001) noted that nitrogen in the chlorophyll closely correlates with the content of photosynthetic pigments, so indices related to it, such as CP might be indirectly inferred (Fig. 6). The NIR range of 750–1700 nm also plays a critical role as its reflectance is mostly controlled by canopy morphology and vigor. Specific wavelengths within this range are sensitive to molecular vibrations linked to nitrogen compounds, i.e., CP content (Fig. 6; Parrini et al., 2022). The red–edge region (700–780 nm) sits between strong chlorophyll absorption in the red and high internal scattering in the NIR, making it highly sensitive to both biochemical and structural properties of leaves (Mutanga and Skidmore, 2007). Because cell walls, mesophyll structure, water content, and associated biochemicals (minerals, lipids, structural carbohydrates) all influence light scattering and absorption around the red edge, this region can indirectly capture variation in crude fat, fiber, and carbohydrates alongside protein (Fig. 6). Consequently, changes in the steepness and position of the red edge (red–edge shift) provide an integrated spectral indicator of canopy nutritional status and structural development.

The PLS and SVM have proven robust for modeling spectral data related to grass quality while accounting for the effects of water content (Askari et al., 2019). In a study with 235 biomass samples from various grasses and legumes, Sun et al. (2021) reported $R^2$ values of 0.85, 0.78, and 0.66 for NDF, ADF, and CP, respectively, using PLS, and 0.84, 0.78, and 0.71 using SVM. In our study, PLS achieved similar $R^2$ values of 0.85 (NDF), 0.76 (ADF), and 0.64 (CP), while SVM yielded 0.82, 0.76, and 0.62, respectively. With StackingRF, these $R^2$ values increased to 0.87, 0.78, and 0.67 (Fig. 7). Although the improvements are modest, they consistently indicate a tendency for the stacking ensemble approach to enhance prediction accuracy.

Numerous studies have underscored the effectiveness of neural network architectures for deep learning, demonstrating superior accuracy in grass quality estimation compared to SVM (Mutanga and Skidmore, 2004; Oliveira et al., 2023). MLP in this study shows a similar high accuracy in quality estimation; MLP and PLS are the two best-performing base machine learning models (Fig. 8, Table 5). Deep learning models efficiently scale with increasing data sizes, whereas

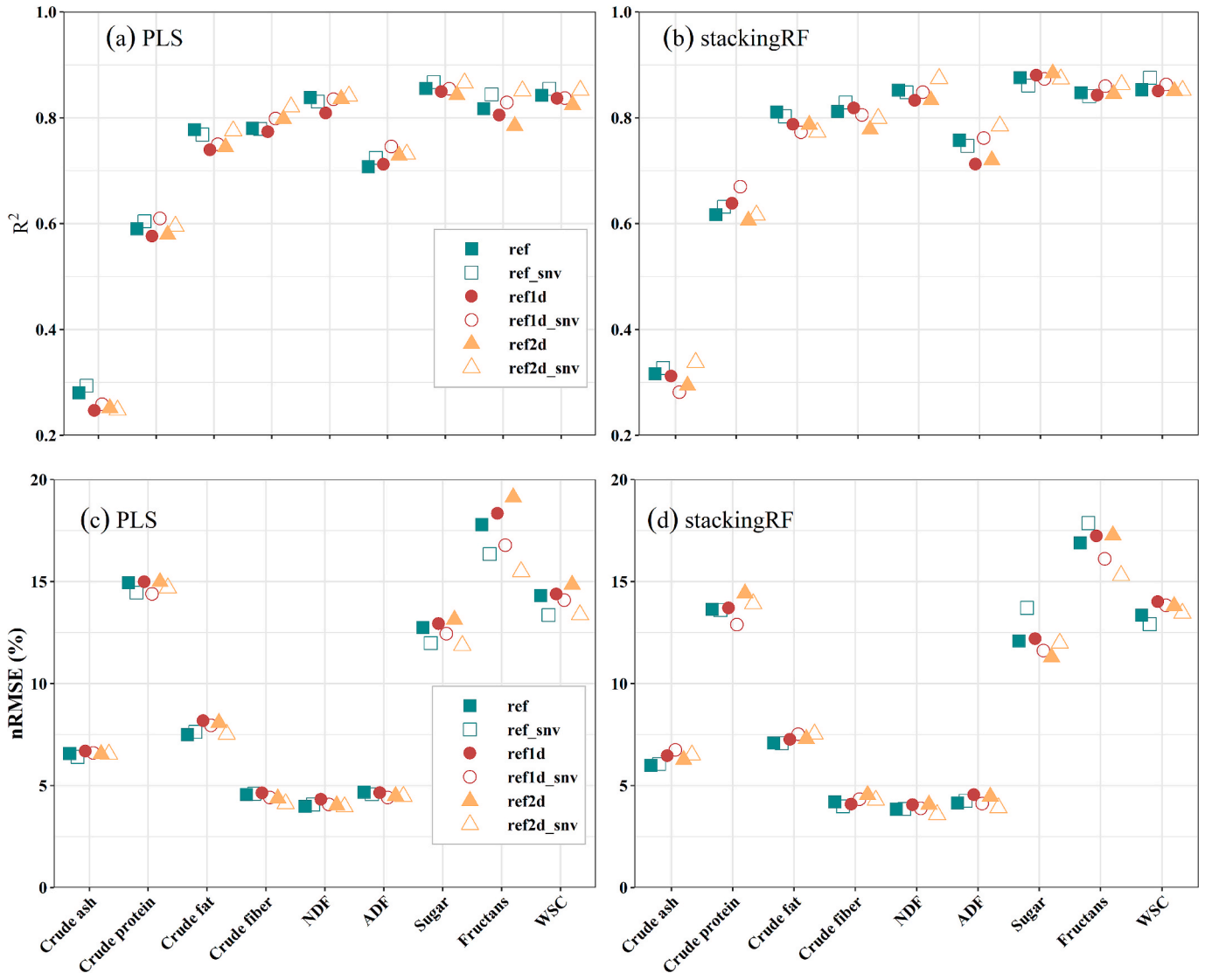


**Fig. 9.** $R^2$ and nRMSE for all quality indices validation sets using the full reflectance for the PLS model. For simplification purposes, we are using the following nomenclature in the legends: ref (reflectance), ref_snv (reflectance standard normal variate), ref1d ($1^{st}$ derivative of reflectance), ref1d_snv ($1^{st}$ derivative of reflectance standard normal variate), ref2d ($2^{nd}$ derivative of reflectance), ref2d_snv ($2^{nd}$ derivative of reflectance standard normal variate).

SVMs often struggle due to their computational complexity, especially in high-dimensional spaces. Neural networks can automatically learn features from raw data through multiple layers, minimizing the need for manual feature extraction. In contrast, SVMs typically require careful feature engineering to achieve optimal performance, a process that can be both time-intensive and less effective (Karila et al., 2021).

Crude ash content, an important indicator of forage nutritional quality, reflects mineral concentrations, making accurate monitoring of this parameter crucial for the quantitative evaluation of forage nutrition (Quirino et al., 2023). Currently, there is limited research on estimating nitrogen-free extract and crude ash content in forage biomass utilizing hyperspectral remote sensing technology (Thomson et al., 2023). Crude ash cannot be associated with a spectral reflectance region (Parrini et al, 2022), probably due to the absence of energy absorption of inorganic substances as minerals. Crude ash is a sum of many minerals (silica, calcium, potassium, magnesium, etc.), and their ratios vary significantly between samples. there was no consistent indication as to how well reflectance will perform when determining mineral ratios (Ikoyi, and Younge, 2020). This may be attributed to the complex composition of nitrogen-free extract and crude ash in forage, leading to disordered spectral absorption characteristics and lower sensitivity of most spectral parameters (Fig. 7; 8; Thomson et al., 2023). For CP, we obtained only moderately good performance ($R^2 \approx 0.63$; Fig. 7) by stackingRF, despite the large range created by contrasting N fertilisation. This is consistent with the fact that CP is a smaller and more rapidly fluctuating component of the grass-clover, strongly affected by day-to-day changes in N supply, weather, and maturity (Thomson et al., 2023). In contrast, fiber and structural carbohydrates, which dominate the biomass (often 40–60 %+ of DM) and have strong, repeated C–H and O–H bonds (Wijesingha et al, 2020), are more robustly expressed in the NIR spectra and therefore better predicted. For this purpose, future studies could test advanced semi-self-learning regressors such as the Deep Neural Networks, which have been shown to notably enhance prediction from the simple RGB image-based models compared to estimators based on hyperspectral imaging (Oliveira et al., 2020; Oliveira et al., 2023).

### *4.2. Applicability of stacking ensemble learning and spectra pre-processing technique*

Singe ML model has a fixed structure and might learn some incorrect features from larger input variables in training, leading to bias. Stacking can overcome this shortcoming through the concept of a *meta*-learner in

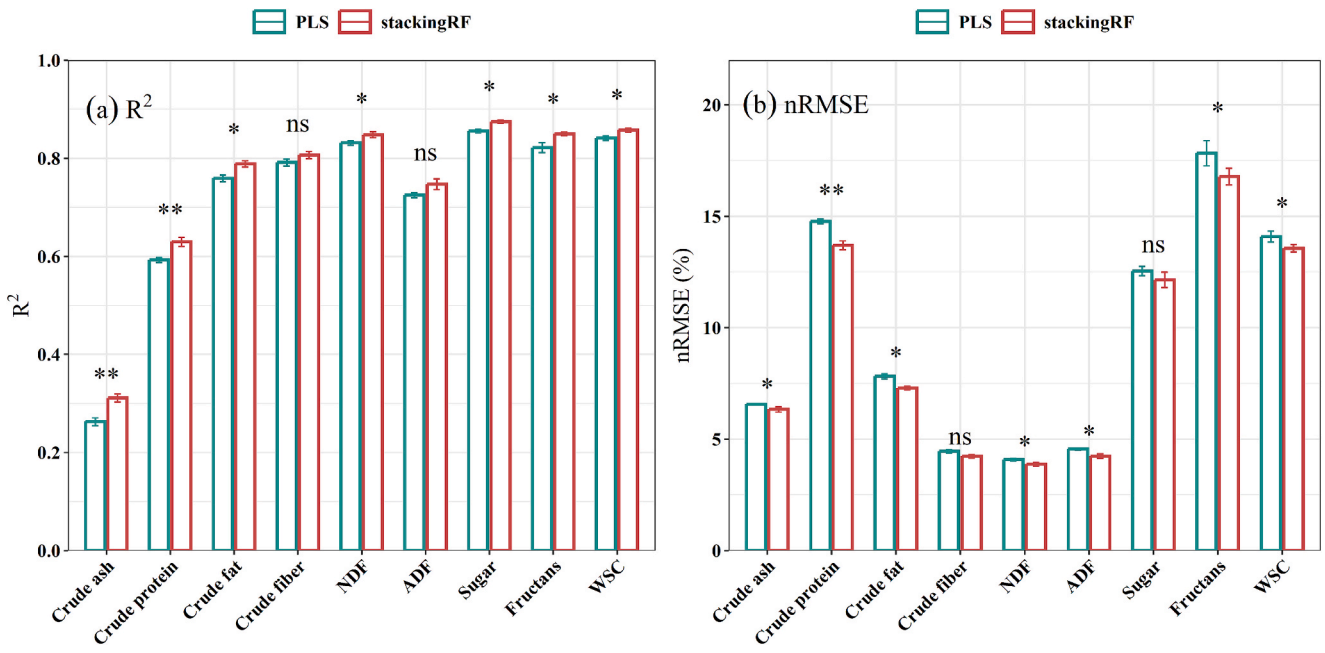


**Fig. 10.** The comparison of $R^2$ and nRMSE of PLS and stackingRF for all the quality indices. The $R^2$ and nRMSE are the mean value of 6 different datasets (ref, ref_snv, red1d, red1d_snv, ref2d, ref2d_snv). The each error bar shows the variance of those 6 datasets. ref (reflectance), ref_snv (reflectance standard normal variate), ref1d ($1^{st}$ derivative of reflectance), ref1d_snv ($1^{st}$ derivative of reflectance standard normal variate), ref2d ($2^{nd}$ derivative of reflectance), ref2d_snv ($2^{nd}$ derivative of reflectance standard normal variate). Note: ** means very significant difference ($P < 0.01$), * means significant difference ($P < 0.05$), ns means no significant difference ($P > 0.05$).

a second layer. Multiple layers of stacking ensemble learning can combine the advantages of a single base model and reduce bias (Zou et al, 2026). Due to the nature of stacked ensembles, it usually produces a more robust predictive performance closer to the actual target than just regular individual models or average ensembles. For example, Du et al. (2024) used the stacking ensemble model in canola leaf chlorophyll prediction and found that it obtained better estimation than single machine learning algorithms ($R^2$ increased from 0.75 to 0.78). In this study, we also found that the stacking ensemble model achieved higher accuracy than PLS and the other individual machine-learning models for predicting grass quality, with statistically significant improvements for all quality indices except crude fiber and ADF (Fig. 10). The effectiveness of estimation by stacking ensemble model was also approved in maize leaf chlorophyll content estimation (Zhai et al. 2023) and soil organic matter estimation (Wu et al., 2023).

It's also necessary to avoid selecting similar or correlated learning algorithms to maximize the prediction accuracy of stacking ensembled learning. We selected four highly performing ML models as base models and used their predictions as input variables through the *meta*-learner. In most cases, the SVM, PLS, ANN, and one of the MLP layers were four well-performed models for quality estimation. The advantages of stacking ensemble learning are best realized when the input feature set and training dataset are large, because the high–dimensional feature space allows different base learners to capture complementary patterns and thus improve overall accuracy. Several studies have shown that ensemble methods, including stacking, tend to outperform single models as data volume and feature richness increase, particularly in complex, high–dimensional regression problems (Shahhosseini et al., 2022).

In this study, for the carbohydrate (sugar, fructans, WSC) and fiber dataset (crude fiber, NDF), the best dataset by the PLS model was the SNV corrected 2nd derivative reflectance (ref2d_snv) with the highest $R^2$ (0.74–0.87) and lowest nRMSE (5–15 %, Figs. 8 and 9). For stackingRF, for NDF, ADF, and WSC, the best dataset was ref2d_snv with the highest $R^2$ (0.78–0.87) and lowest nRMSE (4–15 %). It appears SNV provides a major advantage in most quality estimation in this study (Figs. 8-10). The combination of $2^{nd}$ derivative and SNV as pre-processing techniques has been found to improve the model performance on quality fruits ($R^2$ increase by 4–8 % compared to the standard $2^{nd}$ derivative, Mishra et al., 2020). Thus, the pre-processing technique deserves more attention in future studies as it captures different types of information. The $2^{nd}$ derivative is effective in capturing the peaks related to the chemical components, while the SNV is better at capturing the color information. This complementary information can be used to improve the accuracy of predictions. The combination of SNV and derivative is a much-favored, and often successful, option for pre-treatment of spectra (Passos et al., 2019).

### 4.3. *Strengths and limitations*

This work introduces a comprehensive framework for estimating grass quality using six reflectance datasets (original spectra, first and second derivatives, and their SNV-transformed counterparts). PLS-based feature selection was applied to identify the most relevant wavelengths for machine learning, and 11 base models were trained on these selected features; the best four base learners were then used as inputs to a stacking ensemble model. The results show that PLS feature selection, SNV pre-processing, and stacking ensemble learning all contribute to improved prediction accuracy for grass quality indices.

Botanical composition and growth stage do affect grass-clover canopy spectral reflectance (Thomson et al., 2023). In this study, this variation is present in both the training and test subsets, which were randomly drawn from the full dataset, and all model hyperparameters were tuned using 5-fold cross-validation within the training data. Consequently, the models are fitted and evaluated under the same range of variation in species composition and phenology, so the reported performance already incorporates these sources of variability and remains valid despite them. We worked with very high spectral resolution (10 nm) but relatively coarse spatial resolution at the plot scale in this study, which may not capture fine-scale within-field heterogeneity. As a result, our models are primarily tuned to represent average canopy conditions rather than small-scale variability. Future work should therefore employ stacking ensemble learning with

high-spatial-resolution imaging platforms (e.g., UAV-based hyperspectral sensors; Karila et al., 2021) to better resolve and estimate grass–clover quality at finer scales. Based on the superiority of ML in multi-dimensional data processing, using large sample spatiotemporal data sets to improve the applicability and clarify the spatial distribution pattern of biomass nutrients on a larger spatiotemporal scale is also a major research focus in the future.

## 5. Conclusion

This study proposed a novel method for grass quality estimation by using canopy reflectance and by using the stacking ensemble learning method. We also compared this method by using different datasets from canopy reflectance (ref), canopy reflectance standard normal variate (ref_snv), 1st derivative (ref1d), 1st derivative standard normal variate (ref1d_snv), 2nd derivative (ref2d), and 2nd derivative standard normal variate (ref2d_snv). PLS feature preprocessing can promote estimation accuracy. The grass quality can be well predicted by machine learning and stacking learning, except for crude ash and crude protein; the $R^2$ of other grass-clover quality indices was larger than 0.75. Stacking learning significantly increased the accuracy of quality estimation compared with the best base machine learning methods. ref2d_snv is found to have better performance than other datasets in most cases. Incorporation of these remote sensing techniques and stacking ensemble learning opens the possibilities of monitoring grass with the advantage of fast evaluation of quality under field conditions, and informs efficient decisions in relation to the time of fertilize and harvest.

## CRediT authorship contribution statement

**Shaohui Zhang:** Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Martin Leon Gnyp:** Writing – review & editing, Methodology, Data curation. **Ping Xin:** Writing – review & editing, Visualization, Methodology, Investigation, Formal analysis. **Petra Junklewitz:** Writing – review & editing, Project administration, Funding acquisition. **Jörg Jasper:** Writing – review & editing, Data curation. **Chariton Kalaitzidis:** Writing – review & editing. **Kiril Manevski:** Writing – review & editing, Supervision, Project administration, Funding acquisition.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

This work was partly supported by the project GrassTools ((Innovation Fund Denmark, project number 0224- 00091B). Special acknowledgments are given to the China Scholarship Council for the financial support of the first author (Grant No. 202106300017).

## Data availability

Data will be made available on request.

## References

Askari, M.S., McCarthy, T., Magee, A., Murphy, D.J., 2019. Evaluation of Grass Quality under Different Soil Management Scenarios using Remote Sensing Techniques. Remote Sens. 11 (15), 1835. https://doi.org/10.3390/rs11151835.

Barnes, R.J., Dhanoa, M.S., Lister, S.J., 1989. Standard Normal Variate Transformation and De-Trending of Near-Infrared Diffuse Reflectance Spectra. Appl. Spectrosc. 43, 772–777. https://doi.org/10.1366/0003702894202201.

Batten, G.D., 1998. Plant analysis using near infrared reflectance spectroscopy: the potential and the limitations. Aust. J. Exp. Agric. 38, 697–706. https://doi.org/10.1071/EA97146.

Bec, K., Grabska, J., Bonn, G.K., Popp, M., Huck, C.W., 2020. Principles and applications of vibrational spectroscopic imaging in plant science: a review basic information related to spectra. Front. Plant Sci. 11, 1226. https://doi.org/10.3389/fpls.2020.01226.

Bengtsson, J., Bullock, J.M., Egoh, B., Everson, C., Everson, T., Smith, H.G., Lindborg, R., 2019. Grasslands—More important for ecosystem services than you might think. Ecosphere 10 (2), e02582. https://doi.org/10.1002/ecs2.2582.

Cao, Q., Miao, Y., Feng, G., Gao, X., Li, F., Liu, B., Yue, S., Cheng, S., Ustin, S.L., Khosla, R., 2015. Active canopy sensing of winter wheat nitrogen status: an evaluation of two sensor systems. Comput. Electron. Agric. 112, 54–67. https://doi.org/10.1016/j.compag.2014.08.012.

Cao, Z., Cheng, T., Ma, X., Tian, Y.C., Zhu, Y., Yao, X., Chen, Q., Liu, S., Guo, Z., Zhen, Q., Li, X., 2017. A new three-band spectral index for mitigating the saturation in the estimation of leaf area index in wheat. Int. J. Remote Sens. 38 (13), 3865–3885. https://doi.org/10.1080/01431161.2017.1306141.

Dieterle, P., Engling, P., Horst, H., Paul, C., Robowsky, K.D., Tillmann, P., 2003. Die Anwendung der Nahinfrarotspektroskopie (NIRS) bei der Untersuchung von Futtermitteln und pflanzlichen Produkten. VDLUFA (Verband Deutscher Untersuchungs- und Forschungsanstalten) Standpunkt, Bonn.

Du, R., Lu, J., Xiang, Y., Zhang, F., Chen, J., Tang, Z., Shi, H., Wang, X., Li, W., 2024. Estimation of winter canola growth parameter from UAV multi-angular spectral-texture information using stacking-based ensemble learning model. Comput. Electron. Agric. 222, 109074.

Ertlen, D., Schwartz, D., Trautmann, M., Webster, R., Brunet, D., 2010. Discriminating between organic matter in soil from grass and forest by near-infrared spectroscopy. Eur. J. Soil Sci. 61 (2), 207–216. https://doi.org/10.1111/j.1365-2389.2009.01219.x.

Fernández-Habas, J., Carriere Cañada, M., García Moreno, A.M., Leal-Murillo, J.R., González-Dugo, M.P., Abellanas Oar, B., Gómez-Giráldez, P.J., Fernández-Rebollo, P., 2022. Estimating pasture quality of Mediterranean grasslands using hyperspectral narrow bands from field spectroscopy by Random Forest and PLS regressions. Comput. Electron. Agric. 192, 106614.

Fonseca, C.E.L., Pessoa Filho, M., Braga, G.J., Ramos, A.K.B., Carvalho, M.A., Fernandes, F.D., Karia, C.T., Maciel, G.A., Athayde, N.B., Dessaune, S.N., Thomé, S. P., Garcia, A.C.C., 2020. Near-infrared reflectance spectroscopy as a tool for breeding Andropogon gayanus Kunth for forage quality. IOSR-JAVS 13 (6), 57–66.

Fu, Y., Yang, G., Pu, R., Li, Z., Li, H., Xu, X., Song, X., Yang, X., Zhao, C., 2021. An overview of crop nitrogen status assessment using hyperspectral remote sensing: current status and perspectives. Eur. J. Agron. 124, 126241. https://doi.org/10.1016/j.eja.2021.126241.

Gámez, A.L., Vatter, T., Santesteban, L.G., Araus, J.L., Aranjuelo, I., 2024. On field estimation of quality parameters in alfalfa through hyperspectral spectrometer data. Comput. Electron. Agric. 216, 108463.sssss.

Ikoyi, A., Younge, B., 2020. Influence of forage particle size and residual moisture on near infrared reflectance spectroscopy (NIRS) calibration accuracy for macro-mineral determination. Anim. Feed Sci. Technol. 270, 114674. https://doi.org/10.1016/j.anifeedsci.2020.114674.

Johnson, L.F., 2001. Nitrogen influence on fresh-leaf NIR spectra. Remote Sens. Environ. 78, 314–320.

Karila, K., Alves Oliveira, R., Ek, J., Kaivosoja, J., Koivumäki, N., Korhonen, P., Niemeläinen, O., Nyholm, L., Näsi, R., Pölönen, I., Honkavaara, E., 2021. Estimating Grass Sward Quality and Quantity Parameters using Drone Remote Sensing with Deep Neural Networks. Remote Sens 14 (11), 2692. https://doi.org/10.3390/rs14112692.

Liu, J., Munnaf, M.A., Pätzold, S., Mouazen, A.M., 2026. A dataset-based statistical comparison of gamma-ray and visible-near-infrared sensors for soil texture prediction with ensemble learning. Comput. Electron. Agric. 241, 111326. https://doi.org/10.1016/j.compag.2025.111326.

Luo, J., Ma, R., Feng, H., Li, X., 2016. Estimating the total nitrogen concentration of reed canopy with hyperspectral measurements considering a nonuniform vertical nitrogen distribution. Remote Sens 8 (10), 789.

Manevski, K., Manakos, I., Petropoulos, GP., Kalaitzidis, C., 2011. Discrimination of common Mediterranean plant species using field spectroradiometry. Int J Appl Earth Obs, 13: 922-933. www.doi.org/10.1016/j.jag.2011.07.001.

Manevski, K., Manakos, I., Petropoulos, GP., Kalaitzidis, C., 2012. Spectral Discrimination of Mediterranean Maquis and Phrygana Vegetation: Results From a Case Study in Greece. Ieee J-Stars, 5: 604-616. www.doi.org/10.1109/Jstars.2012.2190044.

Mishra, P., Roger, J.M., Rutledge, D.N., Woltering, E., 2020. SPORT pre-processing can improve near-infrared quality prediction models for fresh fruits and agro-materials. Postharvest Biol. Technol. 168, 111271. https://doi.org/10.1016/j.postharvbio.2020.111271.

Mutanga, O., Skidmore, A., 2004. Integrating imaging spectroscopy and neural networks to map grass quality in the Kruger National Park. South Africa. Remote Sens. Environ 90 (1), 104–115. https://doi.org/10.1016/j.rse.2003.12.004.

Mutanga, O., Skidmore, A.K., 2007. Red edge shift and biochemical content in grass canopies. ISPRS J. Photogramm. Remote Sens. 62 (1), 34–42. https://doi.org/10.1016/j.isprsjprs.2007.02.001.

Naumann, C., Bassler, R., Seibold, R., Barth, C., 1997. Verband Deutscher Landwirtschaftlicher Untersuchungs-und Forschungsanstalten, Methodenbuch Band III, Die chemische Untersuchung von Futtermitteln, 4. Ergänzungslieferung. VDLUFA-Verlag, Darmstadt.

Næs, T., Isaksson, T., Fearn, T., Davies, T., 2002. A User-Friendly Guide to Multivariate Calibration and Classification, NIR Publications: Chichester, Doi: 10.1255/978-1-906715-25-0 .

O'Mara, F.P., 2012. The role of grasslands in food security and climate change. Ann. Bot. 110 (6), 1263–1270. https://doi.org/10.1093/aob/mcs209.

Oliveira, R.A., Näsi, R., Korhonen, P., Mustonen, A., Niemeläinen, O., Koivumäki, N., Hakala, T., Suomalainen, J., Kaivosoja, J., Honkavaara, E., 2023. High-precision estimation of grass quality and quantity using UAS-based VNIR and SWIR hyperspectral cameras and machine learning. Precis. Agric. 25, 186–220.

Oliveira, R.A., Näsi, R., Niemeläinen, O., Nyholm, L., Alhonoja, K., Kaivosoja, J., Jauhiainen, L., Viljanen, N., Nezami, S., Markelin, L., Hakala, T., Honkavaara, E., 2020. Machine learning estimators for the quantity and quality of grass swards used for silage production using drone-based imaging spectrometry and photogrammetry. Remote Sens. Environ. 246, 111830. https://doi.org/10.1016/j.rse.2020.111830.

Parrini, S., Staglianò, N., Bozzi, R., Argenti, G., 2022. Can grassland chemical quality be quantified using transform Near-Infrared spectroscopy? Animals 12 (1), 86. https://doi.org/10.3390/ani12010086.

Passos, D., Rodrigues, D., Cavaco, A.M., Antunes, M.D., Guerra, R., 2019. Non-Destructive Soluble Solids Content Determination for 'Rocha' Pear based on VIS-SWNIR Spectroscopy under 'Real World' sorting Facility Conditions. Sensors 19 (23), 5165. https://doi.org/10.3390/s19235165.

Quirino, D.F., Palma, M.N., Franco, M.O., Detmann, E., 2023. Variations in Methods for Quantification of Crude Ash in Animal Feeds. J. AOAC Int. 106 (1), 6–13. https://doi.org/10.1093/jaoacint/qsac100.

Shahhosseini, M., Hu, G., Pham, H., 2022. Optimizing ensemble weights and hyperparameters of machine learning models for regression problems. Mach. Learn. Appl. 7, 100251. https://doi.org/10.1016/j.mlwa.2022.100251.

Shaver, T., Khosla, R., Westfall, D., 2011. Evaluation of two crop canopy sensors for nitrogen variability determination in irrigated maize. Precis. Agric. 12, 892–904. https://doi.org/10.1007/s11119-011-9229-2.

Sun, S., Zuo, Z., Yue, W., Morel, J., Parsons, D., Liu, J., Peng, J., Cen, H., He, Y., Shi, J., Li, X., Zhou, Z., 2021. Estimation of biomass and nutritive value of grass and clover mixtures by analyzing spectral and crop height data using chemometric methods. Comput. Electron. Agric. 192, 106571. https://doi.org/10.1016/j.compag.2021.106571.

Thomson, A., Jacobs, J., Morse-McNabb, E., 2023. Comparing the predictive ability of Sentinel-2 multispectral imagery and a proximal hyperspectral sensor for the estimation of pasture nutritive characteristics in an intensive rotational grazing system. Comput. Electron. Agric. 214, 108275. https://doi.org/10.1016/j.compag.2023.108275.

Vasseur, F., Cornet, D., Beurier, G., Messier, J., Rouan, L., Bresson, J., Ecarnot, M., Stahl, M., Heumos, S., Ǵerard, M., Reijnen, H., Tillard, P., Lacombe, B., Emanuel, A., Floret, J., Estarague, A., Przybylska, S., Sartori, K., Gillespie, L.M., Violle, C., 2022. A perspective on plant phenomics: Coupling deep learning and near-infrared spectroscopy. Front. Plant Sci, 13, 836488. Doi: 10.3389/fpls.2022.836488.

Viña, A., Gitelson, A.A., Nguy-Robertson, A.L., Peng, Y., 2011. Comparison of different vegetation indices for the remote assessment of green leaf area index of crops. Remote Sens. Environ. 115 (12), 3468–3478. https://doi.org/10.1016/j.rse.2011.08.010.

Wijesingha, J., Astor, T., Schulze-Brüninghoff, D., Wengert, M., Wachendorf, M., 2020. Predicting Forage Quality of Grasslands using UAV-Borne Imaging Spectroscopy. Remote Sens 12 (1). https://doi.org/10.3390/rs12010126.

Wolpert, D.H., 1992. Stacked Generalization. Neural Netw. 5 (2), 241–259.

Wu, M., Dou, S., Lin, N., Jiang, R., Zhu, B., 2023. Estimation and mapping of soil organic matter content using a stacking ensemble learning model based on hyperspectral images. Remote Sens 15 (19), 4713.

Zhai, W., Li, C., Cheng, Q., Ding, F., Chen, Z., 2023. Exploring multisource feature fusion and stacking ensemble learning for accurate estimation of maize chlorophyll content using unmanned aerial vehicle remote sensing. Remote Sens 15 (13), 3454.

Zhang, A., Yan, W., Guo, C., 2018. Inversion model of pasture crude protein content based on hyperspectral image. Transactions of the CSAE, 2018, 34(3): 188-194. Doi: 10.11975/j.issn.1002-6819.2018.03.025.

Zou, Y., Pei, J., Fang, H., Liu, Y., Geng, J., Huang, H., Wang, T., Huang, J., 2026. Organ–level nitrogen diagnosis in rice by integrating hyperspectral spectroscopy with advanced feature selection techniques. Eur. J. Agron. 177, 128109. https://doi.org/10.1016/j.eja.2026.128109.

Zhou, Z., Morel, J., Parsons, D., Kucheryavskiy, S.V., Gustavsson, A., 2019. Estimation of yield and quality of legume and grass mixtures using partial least squares and support vector machine analysis of spectral data. Comput. Electron. Agric. 162, 246–253. https://doi.org/10.1016/j.compag.2019.03.038.