Estimation of forage yield and quality using UAV multispectral imagery

Sarah Carola Vecchietti

Independent project in Biology G2E 15 credits
Swedish University of Agricultural Sciences, SLU
Department Crop Production Ecology
Agriculture Programme – Soil/Plant
Uppsala 2025

Estimation of forage yield and quality using UAV multispectral imagery

Utvärdering av avkastning och kvalitet på foder med hjälp av UAV-multispektral avbildning

Sarah Carola Vecchietti

Supervisor:	**Julianne Oliveira, Swedish University of Agricultural Sciences, Department of Crop Production Ecology**
Assistant supervisor:	Sanna Bergqvist, Swedish University of Agricultural Sciences, Department of Crop Production Ecology
Examiner:	David Parsons, Swedish University of Agricultural Sciences, Department of Crop Production Ecology

Credits:	15 credits
Level:	First cycle, G2E
Course title:	Independent project in Biology
Course code:	EX0894
Programme/education:	Agriculture Program – Soil/Plant
Course coordinating dept:	Department of Aquatic Science and Assessment
Place of publication:	Uppsala
Year of publication:	2025
Keywords:	Multispectral imaging, UAV, forage quality, remote sensing

Swedish University of Agricultural Sciences
Faculty of Natural Resources and Agricultural Sciences
Department of Crop Production Ecology

Abstract

The quality of forage is a critical factor influencing animal health, productivity, and the nutritional value of derived products. Accurate assessment of forage quality and yield is essential for optimizing livestock nutrition, improving productivity, and promoting animal welfare. While laboratory analyses provide precise nutritional parameters, they are often labor intensive and time consuming. Remote sensing technology offers a promising alternative by enabling rapid, large scale monitoring of forage characteristics. Widely regarded as a key tool in precision agriculture, remote sensing facilitates efficient field monitoring and data-driven decision-making. This study investigates the potential of multispectral imaging technology to estimate both the quantity and quality of forages. Over a two-year period, data on grass-legume mixtures were collected in Northern Sweden using an Unmanned Aerial Vehicle (UAV) equipped with RGB and multispectral cameras. Captured images underwent radiometric corrections and were processed into orthoimages to extract reflectance data. To predict the key forage quality indicators Dry Matter (DM) yield, Crude Protein (CP), Neutral Detergent Fiber (aNDFom), Organic Matter Digestibility (OMD), and Metabolizable Energy (ME), four statistical models were tested: Multiple Linear Regression (ML), Partial Least Squares (PLS and caretPLS), and Support Vector Machine (SVM). Model performance was evaluated using Root Mean Square Error (RMSE), Relative RMSE (RRMSE), and Coefficient of Determination (R^2). Among the models, PLS demonstrated a superior performance in predicting certain parameters in the mixed grass and clover dataset. Specifically, PLS achieved an R^2 of 0.71 for ME and 0.75 for DM yield. ML exhibited the lowest RMSE and RRMSE values and the highest R^2 values across all parameters. The ML results were likely effected by overfitting as evidence of multicollinearity was observed, suggesting potential redundancy among predictor variables. CP and aNDF showed a lower reliability due to the missing data. These findings underscore the potential of forage quality prediction using UAVs while emphasizing the need for feature selection to address multicollinearity issues.

Keywords: Multispectral imaging, UAV, forage quality, remote sensing

Table of Contents

List of tables

List of figures

Abbreviations

Abbreviation	Description
aNDF	Amylase-treated NDF
CP	Crude Protein
DM	Dry Matter yield
DSM	Digital Surface Model
DTM	Digital Terrain Model
GCP	Ground Control Point
GIS	Geographical Information System
MSI	Multispectral Imaging
ME	Metabolizable Energy
NDF	Neutral Detergent Fiber
NDVI	Normalized Difference Vegetation Index
NIR	Near-infrared
OMD	Organic Matter Digestibility
RGB	Red, Green Blue, primary colors
UAV	Unmanned Aerial Vehicle
VIF	Variance Inflation Factor.

1. Introduction

1.1 Study context

To meet the growing demand for food and fiber from a global population projected to exceed 9.1 billion by 2050, the 21st century must significantly increase production per unit of land while ensuring the protection of soil health and the conservation of natural resources (Tilman et al. 2011).

In this context, forage crops play a crucial role in livestock systems, which are essential for converting non-edible crops like grass into energy forms that humans can utilize, such as diary and meat (Capstaff & Miller 2018). In Northern Sweden, forage grasslands are among the most important crops, accounting for more than 80% of the total agricultural land use (Jordbruksmarkens användning 2023. Slutlig statistik 2023). Due to climatic conditions and the short growing season in large parts of Sweden, farmers in these regions primarily grow crops for animal feed, such as forage for livestock as well as coarse grains (Facts about Swedish agriculture 2009). However, the long, bright summer days allow herbaceous plants to accumulate carbohydrates from photosynthesis for almost 24 hours a day, while the relatively low temperatures at the beginning of summer reduce lignification, promoting forage with high nutritional potential (Krizsan et al. 2021).

Forage quality directly influences animal nutrition, critical to both animal performance and producer profitability. However, the nutritional composition of forages can vary significantly, introducing uncertainties and inconsistencies in nutrient delivery to animals. Such variability can, in turn, affect their health, productivity, and overall welfare (St-Pierre & Weiss, 2015). Substantial differences in forage quality can arise among various forage crops, influenced by a range of factors. Key determinants include the stage of maturity at harvest and, for preserved forages, the methods of harvesting and storage. Additionally, secondary factors, such as soil fertility, fertilization practices, climatic conditions during the growth period, and the specific crop variety, also play a significant role (Oregon State University 2019). A comprehensive understanding and meticulous management of the nutritional variability in forages are, therefore, essential for improving dairy farm productivity while simultaneously ensuring enhanced animal welfare (Bullet et al., 2007). Forage production is also significantly influenced by the cutting intervals. Shorter cutting intervals typically enhance forage quality but reduce the yield per cut, while longer intervals lead to higher yields per cut but lower forage quality.

Adjusting cutting intervals is a key management strategy in forage production, as it directly affects both the quantity and nutritional value of the forage available for livestock (Krizsan et al. 2021). In Sweden, one commonly used method is Vallprognos.se, a service operated by the Vallprognos project. This tool relies on temperature totals and crop samples, which are analyzed in a laboratory, to determine the optimal timing for the first harvest. Weather stations across the country provide the necessary data for these predictions (Vallprognos.se, 2025). Although Vallprognos is highly effective for estimating the timing of the first cut, it does not offer guidance for subsequent harvests. This limitation had led to

exploration of remote sensing as a potential solution, as it enables frequent monitoring, which is important for a dynamic system like forage crops.

Remote sensing is generally considered one of the most important technologies for precision agriculture, as it uses specialized sensors to detect certain physical or chemical properties of an object, area, or phenomenon through data collected by a system that is not in direct contact with the observed subject (Plant 2018). It relies on, for example, satellite imagery or unmanned aerial vehicles (UAVs) to monitor vegetation properties during specific growth stages, weed mapping and management, irrigation management, crop spraying, and disease detection (Tsouros et al. 2019a). Data collection through remote sensing is a non-destructive method. Furthermore, the data can be stored and reused multiple times without reducing quality (Pranga et al. 2021).Remote sensing in precision agriculture offers several key advantages, including the ability to obtain measurements across multiple locations and times, the rapid data collection and processing, the relatively low cost of various types of data, and the ease of obtaining information from areas that are typically difficult to access on the ground (Wachendorf et al. 2018). This provides farmers with a comprehensive view of their crops in a short time, making it possible to take timely action in case of issues (Shy et al. 2016). One of the most significant technologies associated with UAVs in agriculture is the use of drones equipped with multispectral cameras, which capture images at various wavelengths to collect crucial data on plant health and growth.

This enables farmers to respond promptly to issues such as water stress, nutrient deficiencies, or pest infestations, while also providing the ability to estimate crop yields (Tsouros et al. 2019b).

1.2 Research objectives

The objective of this thesis is to evaluate the use of Multispectral Imaging (MSI) sensors from an unmanned aerial vehicle (UAV) for estimating forage yield and quality of ley fields in Northern Sweden. By integrating advanced imaging techniques and laboratory analysis of physical samples, this study aims to improve the efficiency and accuracy of field monitoring methods. The ultimate goal is to provide farmers with a reliable alternative to optimize crop management, enhance forage production, and support sustainable agricultural practices.

2. Theoretical background

2.1 Electromagnetic radiation and multispectral imaging

Sunlight, a form of electromagnetic (EM) radiation emitted by the Sun, propagates as waves with oscillating electric and magnetic fields at the speed of light. Upon reaching the Earth's surface, sunlight is either reflected or absorbed, converting into thermal or chemical energy (Huang et al. 2019).

EM radiation spans from gamma rays (wavelength <1 pm) to radio waves (>1 cm), with visible light (400–700 nm) between them. The visible spectrum ranges from violet (380–450 nm) to red (625–750 nm) and includes the colors blue (450–485 nm), green (500–565 nm), yellow (565–590 nm), and orange (590–625 nm). As shown in Figure 1, healthy vegetation absorbs blue and red light for photosynthesis while reflecting near-infrared (NIR) radiation (700–1100 nm), providing valuable data for assessing crop health. Stressed plants, due to factors like water scarcity or pathogens, show altered ratios of absorbed visible light to reflected NIR radiation (Li et al. 2014).

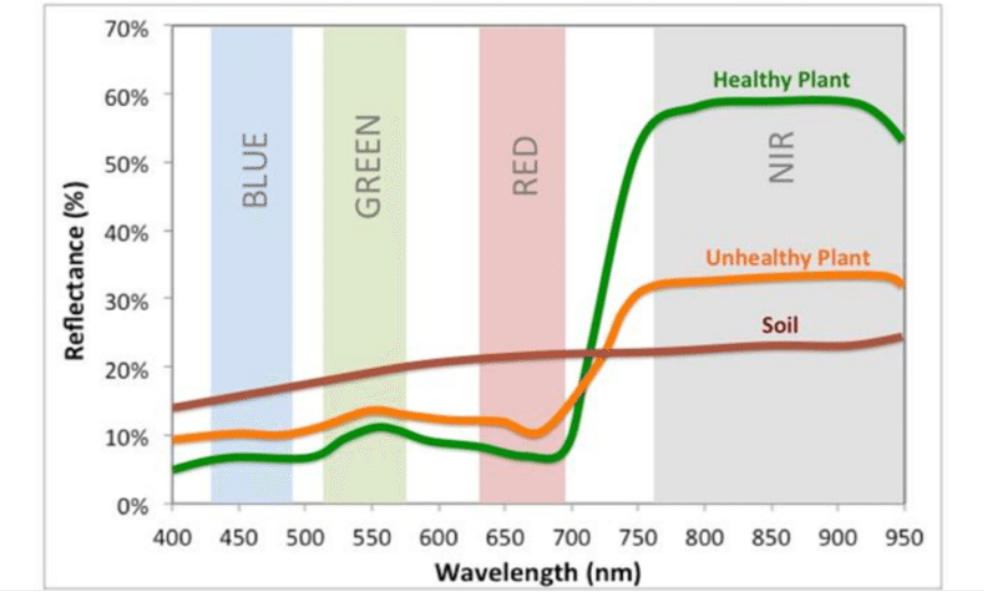

Figure 1. *Reflectance of healthy vegetation (Lum et al. 2016).*

MSI is an advanced technique for capturing images across specific spectral bands, which enable the inference of physical and chemical characteristics of

observed objects. In agriculture, it is used to study crop interactions with electromagnetic radiation. Figure 2 shows the multispectral bands used in this study compared to the electromagnetic spectrum.

The collected data are processed to generate images in multiple spectral bands, illustrating the proportion of light absorbed or reflected at different wavelengths, including the visible spectrum (Red, Green, Blue), near-infrared (NIR), and, in some cases, mid-infrared. MSI facilitate the calculation of vegetation indices across various spectral bands, highlighting key vegetation attributes such as chlorophyll content and biomass. The resulting outputs are monochromatic grayscale images, each corresponding to a specific spectral band (Multispectral imaging 2024).

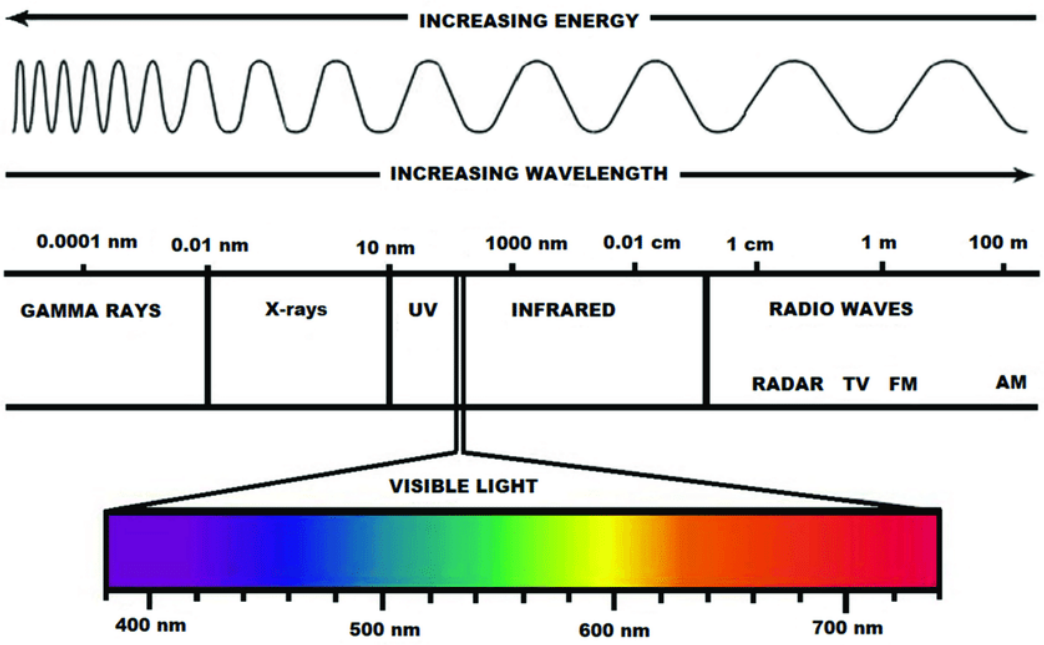

Figure 2. Electromagnetic spectrum (Ailioaie & Litscher 2020).

2.2 Vegetation indices

Vegetation indices are calculated using data collected from drones, satellites, or other remote sensing technologies to evaluate biomass, health status, and vegetation quality. These indices are based on light reflectance in different spectral bands, such as visible, NIR, and mid-infrared. Among the most commonly used indices is the Normalized Difference Vegetation Index (NDVI). This is calculated using the reflectance of electromagnetic radiation in the red and NIR bands. The NDVI value ranges between -1 and +1. The greater the reflected NIR and the more red light absorbed, the higher the chlorophyll content in the plant. A higher NDVI value indicates healthier and more photosynthetically active crops (Huang et al. 2021). The NDVI is based on red and NIR reflectance according to Equation 1:

$$NDVI = \frac{NIR - Red}{NIR + Red}$$

Equation 1

3. Method and materials

3.1 Study area

For this thesis, a total of 14 missions of drone flights were analyzed across two study sites located in Umeå, as shown in Figure 3 and detailed in Table 1. The first site, Röbäcksdalen (RBD) (63.809°N, 20.231°W), is a field research station jointly managed by the Department of Crop Production Ecology and a research barn housing dairy cows, which is affiliated to the Faculty of Veterinary Medicine. This site is subdivided into several areas, with fields F17 and F35 selected for the study. Nine missions of flights were conducted at this site between the summer of 2021 and 2022. Agricultural production at RBD includes either timothy (*Phleum pratense*) and red clover (*Trifolium pratense)* or timothy, meadow fescue (*Lolium pratense*) and red clover.

The second study site is located in Djäkneböle (DKB) (63.773°N, 20.049°W),where the plant species composition was not recorded, but consisted mainly of timothy and red clover.

These missions were part of the Cybergrass 1 project (https://www.slu.se/en/departments/crop-production-ecology/research/crop-production-specialized-in-forages/finished-projects/cybergrass/), funded by the Interreg Botnia-Atlantica.

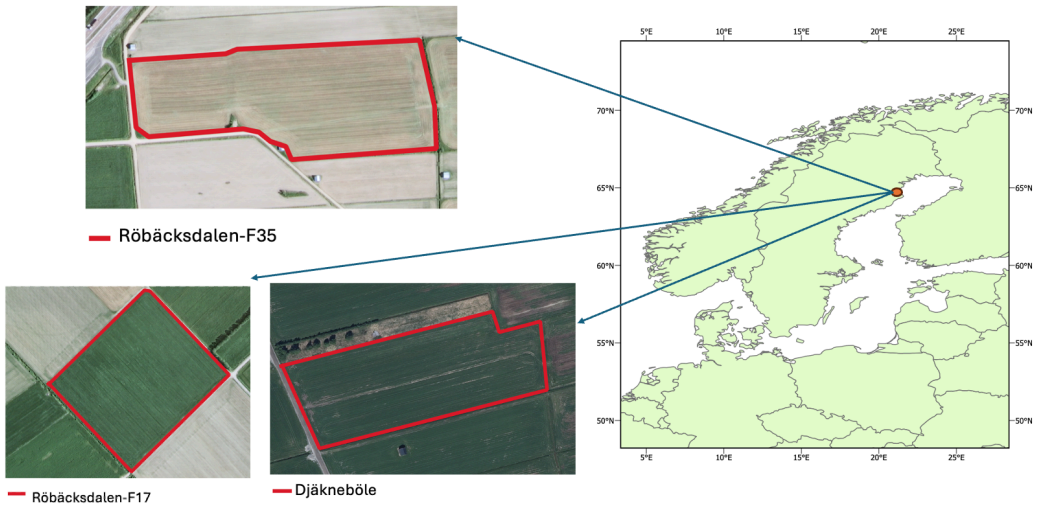

Figure 3. Spatial overview of study fields in Umeå [Läntmäteriet Orthophoto RGB 0.25/0.50 m. Accessed through SLU (Geodata Extraction Tool 2025) and generated with QGIS). Map with world administrative boundaries generated with ArcGIS (World Administrative Boundaries - Countries and Territories 2025).

Table 1. Overview of the missions with date and site locations.

Mission	Date	Flight id	Site
1	2021-06-04	20210604_RBD_CGS_P4MS	RBD
2	2021-06-08	20210608_RBD_CGS_P4MS	RBD
3	2021-07-09	20210709_RBD_CG_P4MS	RBD
4	2021-07-12	20210712_RBD_CG_P4MS	RBD
5	2021-07-15	20210715_RBD_CG_P4MS	RBD
6	2021-09-08	20210908_DKB_CG_P4MS	DKB
7	2022-06-07	20220607_RBD_CG_P4MS	RBD
8	2022-06-14	20220614_DKB_CG_P4MS	DKB
9	2022-07-25	20220725_DKB_CG_P4MS	DKB
10	2022-07-28	20220728_DKB_CG_P4MS	DKB
11	2022-08-23	20220823_RBD_CG_P4MS	RBD
12	2022-08-26	20220826_RBD_CG_P4MS	RBD
13	2022-08-29	20220829_RBD_CG_P4MS	RBD
14	2022-09-01	20220901_DKB_CG_P4MS	DKB

In each field, samples were collected using quadrats (50×50 cm) before and close to each of the three distinct periods during the vegetation season: Period 1, Period 2, and Period 3. The positions of these quadrats were recorded using a GPS system. Vegetation within each quadrat was cut and transported to the laboratory where it was manually separated into grass and legumes. The separated samples were weighed and subsequently dried in an oven at 60℃. After drying, the samples were weighed again and dry matter yield (DM Yield, kg DM/ha) was calculated. Samples were sent to SLUs forage quality lab at Ultuna for forage quality assessment.

These samples were analyzed to estimate various quality-related variables, including, Organic Matter Digestibility (OMD, %), Metabolizable Energy (ME, MJ/kg DM), Amylase-treated NDF (aNDF, g/kg DM), and crude protein (CP, g/kg DM). Table 2 provides a summary of the sampling events conducted across various sites and cut periods during the study.

Table 2. Summary of sampling events across sites and cut periods.

Site	Date	Tot. quadrats	Cut period
RBD	2021-06-04	3	1
RBD	2021-06-08	1	1
RBD	2021-06-09	1	1
RBD	2021-06-10	1	1
RBD	2021-07-09	6	2
RBD	2021-07-12	6	2
RBD	2021-07-15	6	2
DKB	2021-09-08	6	3
DKB	2022-06-07	6	1

RBD	2022-06-14	6	1
RBD	2022-07-25	6	2
RBD	2022-07-28	6	2
DKB	2022-08-23	6	3
RBD	2022-08-26	6	3
RBD	2022-08-29	6	3
DKB	2022-09-01	6	3

3.2 Data acquisition

Aerial imagery was captured using a DJI Phantom 4 Multispectral UAV, a drone specifically designed for agricultural and environmental applications. This UAV is equipped with both a high-resolution RGB camera and multispectral sensors, capable of capturing images across five spectral bands: blue, green, red, red-edge, and near-infrared (NIR) (*DJI Phantom 4 Multispectral* 2025), with the specific wavelengths detailed in Table 3.

Table 3. Band number, name, and wavelengths of the DJI Phantom 4 Multispectral UAV.

Band number	Band names	Center wavelength (nm)
1	Blue	475
2	Green	560
3	Red	665
4	Red-edge	740
5	NIR	790

Ground Control Points (GCPs) were strategically placed across the area of interest, typically positioned at each corner and in the center of the field. These GCPs were essential for ensuring accurate georeferencing of images captured at different times and for aligning them during the orthomosaic generation process. The coordinates of the GCPs were recorded using a Trimble GPS system. Radiometric calibration targets, known as Calibrated Reflectance Panels (CRPs), were used to calibrate the images based on daily solar radiation levels. These targets allowed for the acquisition of absolute reflectance values, which were then used in the second stage of the process to generate reflectance maps and indices (Pix4D Documentation 2024). Calibration tiles with reflectance values of 50%, 46%, and 84%, as well as one with 20%, were placed for calibration. Flights on 2022-08-26 and 2022-08-29 were conducted at an altitude of 80 meters. Additional image acquisition missions were carried out at 50 and 100 meters, capturing images from the visible spectrum in JPEG and TIFF formats.

3.3 Imaging processing

The next phase involved processing multispectral imaging through a series of steps to extract useful information from the acquired data. The Pix4D Desktop software (Pix4D, SA Lausanne, Switzerland) was utilized for image processing and orthomosaic creation. The geodetic reference system employed was WGS 84 (World Geodetic System 1984), with a projected coordinate system in the UTM zone 34N (epsg 32634). A full-resolution GeoTIFF file was generated, combining the individual tiles. The image processing workflow in Pix4D consisted of three main stages: point cloud and mesh generation, creation of the digital surface model (DSM), and orthomosaic and index generation.

Between 5 and 10 Ground Control Points (GCPs) were incorporated into the project after the initial processing phase using the Ray Cloud Editor. The Ray Cloud Editor enables a three-dimensional representation of the drone-captured images. It enables the visualization of a cloud of rays extending from the camera to the registered point, ensuring higher precision in the georeferencing process.

Five reflectance maps were generated by processing the collected data, which represent the reflectivity of surfaces at specific spectral bands. For this type of map, no color balancing was applied, and the weight of each pixel in the original images is calculated to accurately represent the reflectance of the analyzed object. Each pixel in the map represents the proportion of light reflected by objects in a specific spectral band, providing detailed information about the surface's physical or chemical properties. In addition, NDVI (Normalized Difference Vegetation Index) maps were generated using Pix4D software. These maps are derived from the reflectance data and are valuable for assessing vegetation health by highlighting areas of strong photosynthetic activity. An orthomosaic was also generated: a single image that maintains consistent scale and geometry. In this case, the software applies color balancing to ensure that the images blend better together, producing a visually more appealing result. Both the reflectance map and orthomosaic are derived from the DSM (Digital Surface Model) (*Pix4D Documentation* 2024). Additionally, a quality report was produced to provide a detailed assessment of the photogrammetric accuracy and the overall data processing quality.

3.3.1 Data extraction

The analysis of information from the outputs generated by Pix4D was subsequently conducted using a geoinformation software Quantum Geographic Information System (QGIS) to extract statistical data such as median, mean, majority and standard deviation for each sample location. A circular area with a diameter of 50 cm was used for data processing in QGIS using the Zonal Statistics tool. These sample locations were represented as quadrats of 50x50 cm overlaid on the reflectance map. By intersecting the quadrats with the reflectance map, QGIS computed the statistical summaries of reflectance values within each quadrat, providing detailed insights into the surface properties at these specific

locations. These values were then exported as CSV files and converted to Excel format for further analysis. Figure 4 shows the workflow for data processing and extractions.

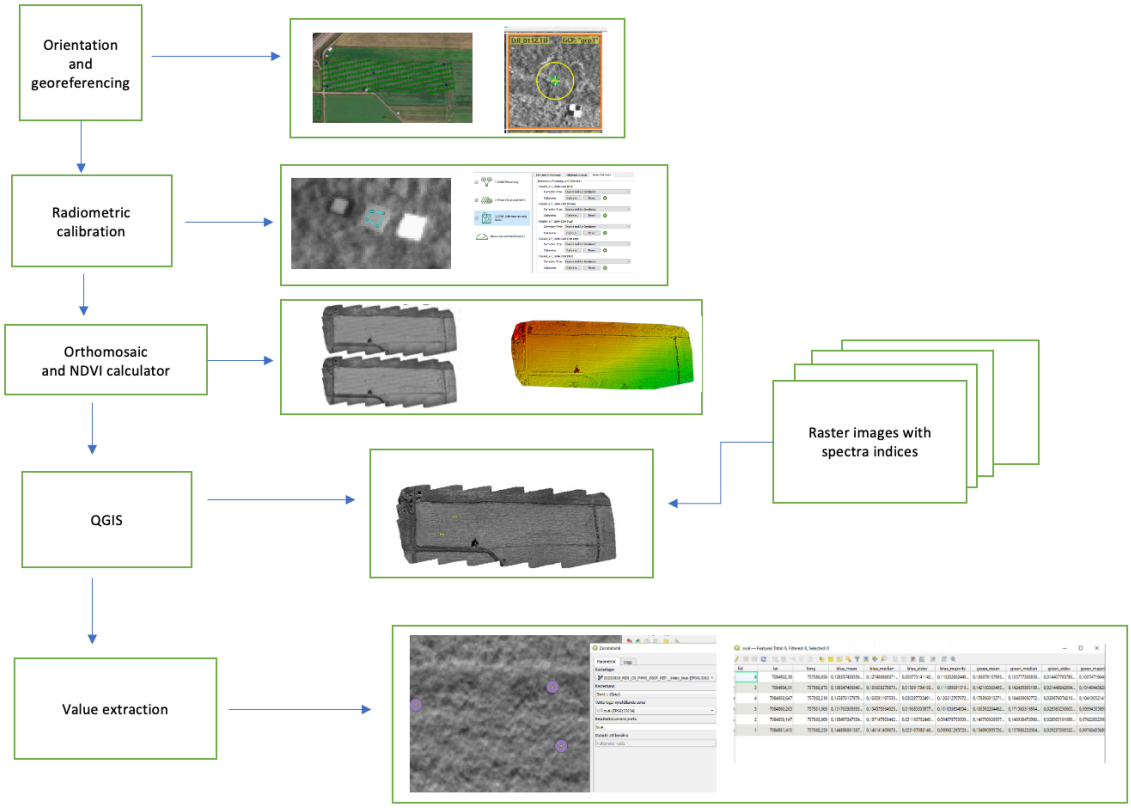

Figure 4. Workflow for data processing in Pix4D and QGIS.

3.4 Data analysis

R programming language was used for processing and analyzing data. A dedicated R script was developed to estimate forage quantity and quality utilizing spectral data collected via remote sensing and laboratory measurements. The analysis involved developing predictive models for different outcomes based on the observed data and evaluating their predictive performance.

The predictive spectra variables consisted of the median, standard deviation, and majority of the following selected bands: blue, green, red, red-edge, NIR.

3.4.1 Predictive models

For each lab measurement, four alternative prediction models were considered:

1. Multivariate linear (ML) models, where the outcome variable is modelled as a linear combination of the spectra variables.
2. Partial Least Squares (PLS) model as implemented in the R package pls, where the correlated spectra variables are first reduced in a set of principal components, later used for predicting outcomes.
3. Similar PLS models implemented in the caret R package (caretPLS), which offers a better tuning of hyperparameters to avoid overfitting.
4. The alternative linear Support Vector Machine implemented in the caret R package (caretSVM), where a hyperparameter controls the trade-off between model complexity and misclassification error.

A 10-fold repeated cross-validation was implemented (except for the ML models) to train and evaluate the predictive models to properly characterize their predictive performances considering possible overfitting.

To assess the presence of multicollinearity among the independent variables, the Variance Inflation Factor (VIF) was computed using the *vif* function in the *car* R package.

The VIF measures how much the variance of a regression coefficient is inflated due to multicollinearity. Generally, a VIF value above 10 suggests a high level of multicollinearity, while values between 1 and 5 indicate moderate correlation but not severe multicollinearity (Ngabire et al. 2022).

3.4.2 Evaluating model performance

To measure how well the models predicted the outcome variables, the following measures were evaluated:

- Root Mean Square Error (RMSE): the square root of the mean of the squared differences between predicted and observed values.
- Relative Root Mean Squared Error (RRMSE): the RMSE normalized by the mean of the observed values.
- The coefficient of determination R^2: it indicates the proportion of the variance in the observed values that is explained by the model, ranging from 0 (no predictive value) to 1 (perfect prediction).

Additionally, the selected predictive measures were complemented with descriptive scatter plots. For each outcome variable and predictive model, the predicted versus the observed was plotted. Scatter plots for superior predictive models are expected to exhibit points scattered around the diagonal line. The diagonal line (y = x) serves as a reference to illustrate the strength and the direction of the correlation, which can be either positive (increasing) or negative (decreasing). Data points positioned near the line, indicates a strong relationship between the two variables.

3.5 Missing values and error exclusion

Two common issues in prediction of field data relate to the presence of missing measurement as well as incorrect or corrupted measurements.

For missing values, predictive models were trained and evaluated only on the subset of complete observation.

For aNDF and CP, only the samples containing mixed grass and legume were used, which reduced the number of observations compared to the other estimations and may have affected accuracy of the results.

For error identification, supervised knowledge of experts in the field was used to identify selected measurements, which demonstrated failure in the spectra measurements. These identified errors were excluded from the analysis in order to ensure the robustness and accuracy of predicted results.

4. Results

This section presents the results for the different tasks described in the methods section. First, a case study demonstrates the process of reconstructing the spectral variables for the drone data during an exemplary mission. Then a descriptive analysis describes the forage yield and quality and presents the spectral variables used in the prediction models. Finally, the predictive performance of the constructed prediction models is evaluated and contrasted.

4.1 Case Study

The mission identified as 20220829_RBD_CG_P4MS was conducted at the RBD site to acquire data for the study. The drone flight lasted approximately 30 minutes. During the mission, operations were temporarily paused when a cloud obstructed the sun, ensuring consistency in radiation conditions and minimizing potential variability in the dataset. The weather was partially cloudy with medium wind. Three spectral targets, with reflectance values of 50%, 84%, and 46%, respectively, were placed on the ground to facilitate calibration and validation of the data. As shown in Figure 5, five Ground Control Points (GCPs) were strategically distributed across the study area. These points were georeferenced and used to improve the spatial accuracy of the outputs. As reported in Figure 6, Point Cloud visualization of field was generated using Pix4Dmap.

Calibration was performed using the spectral target with a 46% reflectance value. The processing software generated raster index maps for the five bands used during the mission and the NDVI index. These raster maps were subsequently imported into QGIS for further analysis. Specifically, Zonal Statistics were calculated to evaluate the distribution and characteristics of the indices in specific zones of interest (Figure 7).

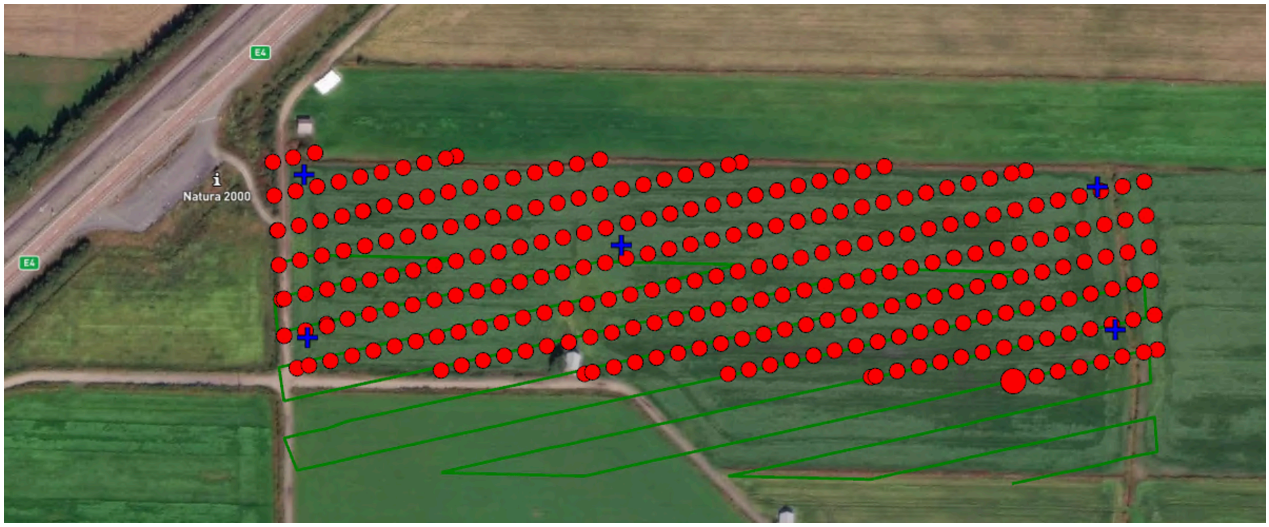

Figure 5. Field mapping using Pix4Dmap, with GCPs indicated in blue.

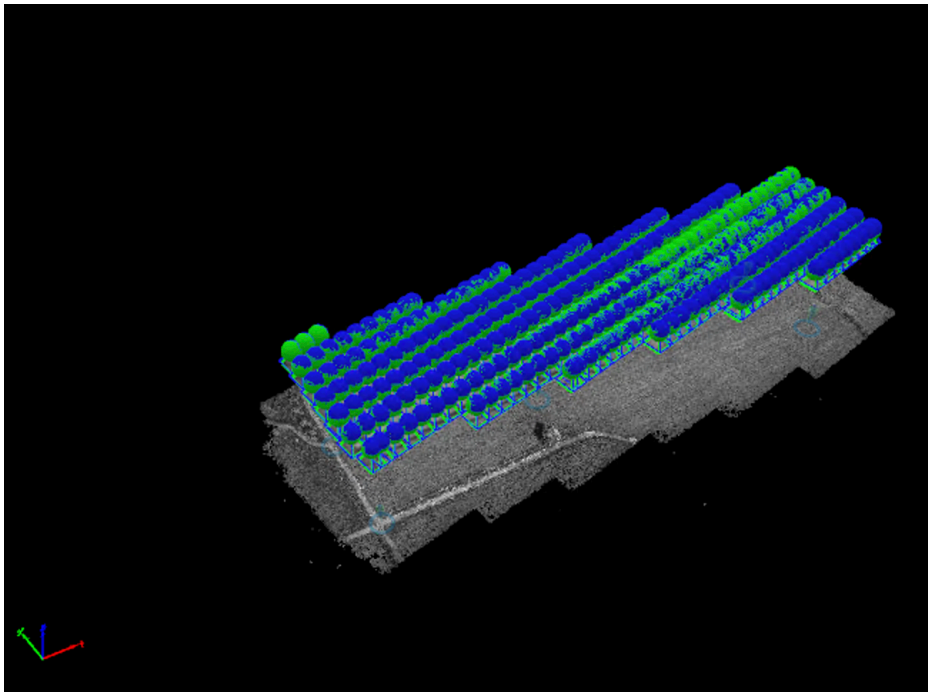

Figure 6. 3D Point cloud visualization of field vegetation generated using Pix4Dmap.

Figure 7. NIR reflectance raster overlaid with polygon boundaries in QGIS for zonal statistical analysis.

4.2 Data selection and descriptive summary of the forage quality and yield

All the multispectral analyses were conducted according to the schedule. Due to a problem with the reflectance panel, three missions were excluded from the statistical analysis (2021-07-09, 2021-07-12, 2021-07-15). The mission from 2021-09-08 was excluded as there was no reflectance panel used during this flight. Moreover, the mission 2022-06-14 was also excluded due to missing values.

The results of the descriptive analysis for the spectral parameters are presented in Table 4. The average reflectance values for the red band were 0.04 for RBD in 2022 and 0.05 in 2021, while DKB had a lower reflectance value of 0.03. The green band showed similar values for RBD in both 2021 and 2022, but DKB had a notably lower value of 0.073.

For the NIR band, RBD in 2021 had the highest average value of 0.816, while RBD in 2022 and DKB exhibited similar reflectance values, also averaging around 0.5. In this experiment, no remarkable differences were observed in the NDVI values across sites and years. The median NDVI values averaged around 0.8, with a maximum of 0.902 and 0.892 for RBD, respectively in 2021 and 2022, and 0.914 for DKB.

The measured yields and quality parameters followed an expected distribution (Table 5). The DM yield showed significant differences between years for RBD and DKB. In 2021, the average DM yield for RBD was 2986 kg DM/ha, decreasing significantly in 2022 to 1690 kg DM/ha. For DKB, the average DM yield was approximately 1900 kg DM/ha. Similar trends were observed in the minimum and maximum values for both fields, indicating that the yield was not directly correlated with NDVI.

For OMD, there were no substantial differences, with an average value of around 75%. The minimum and maximum values were also consistent across samples. In contrast, CP showed greater variability. The average CP for RBD increased from 172 g/kg DM in 2021 to 212 g/kg DM in 2022, while DKB had a lower average CP of 158 g/kg DM.

aNDF showed differences between years for RBD and DKB as the average value for RBD was 469 g/kg DM and decreased in 2022 to 450. For DKB the average value was approximately 436 g/kg DM.

Table 4. Descriptive analysis for the spectral parameters divided per site and year.

Site	Year	mean	sd	min	p25	median	p75	max
NDVI								
RBD	2021	0.872	0.025	0.804	0.862	0.875	0.890	0.902
RBD	2022	0.862	0.018	0.834	0.851	0.856	0.869	0.892
DKB	2022	0.880	0.026	0.843	0.853	0.878	0.902	0.914
NIR								
RBD	2021	0.816	0.094	0.670	0.760	0.786	0.893	0.951
RBD	2022	0.553	0.113	0.453	0.471	0.491	0.663	0.818
DKB	2022	0.525	0.081	0.401	0.471	0.499	0.555	0.697
Green								
RBD	2021	0.138	0.035	0.097	0.105	0.136	0.166	0.233
RBD	2022	0.142	0.054	0.077	0.097	0.128	0.183	0.233
DKB	2022	0.073	0.012	0.058	0.061	0.075	0.080	0.094
Red								
RBD	2021	0.056	0.011	0.040	0.049	0.056	0.062	0.088
RBD	2022	0.040	0.006	0.032	0.037	0.039	0.042	0.051
DKB	2022	0.033	0.010	0.022	0.024	0.033	0.039	0.052
Red edge								
RBD	2021	0.474	0.059	0.400	0.416	0.474	0.551	0.556
RBD	2022	0.328	0.104	0.204	0.219	0.308	0.422	0.514
DKB	2022	0.331	0.045	0.281	0.292	0.325	0.351	0.428
Blue								
RBD	2021	0.068	0.032	0.030	0.040	0.057	0.088	0.143
RBD	2022	0.054	0.014	0.033	0.039	0.057	0.063	0.075
DKB	2022	0.045	0.015	0.031	0.032	0.038	0.061	0.065

Table 5. Descriptive analysis for the measured yields and quality parameters.

Site	Year	mean	sd	min	p25	median	p75	max
Crude protein mixed grass and clover								
RBD	2021	171.9	12.6	153.7	163.2	171.2	176.5	199.7
RBD	2022	212.3	19.4	193.3	200.3	208.9	220.9	238.1
DKB	2022	157.7	12.5	140.4	149.1	158.7	168.1	171.3
DM Yield (kg DM/ha)								
RBD	2021	2986	558.8	1686	2666	3048	3394	3936
RBD	2022	1690	535.5	696.6	1310	1751	2007	2638
DKB	2022	1899	470.7	925.7	1664	1997	2155	2673
ME (MJ/kg DM)								
RBD	2021	10.7	0.4	10.1	10.3	10.7	11.1	11.7
RBD	2022	11.2	0.7	9.9	10.6	11.3	11.8	12.3
DKB	2022	10.9	0.6	9.8	10.6	10.8	11.4	11.7
OMD (%)								
RBD	2021	74.9	2.6	71.2	72.6	74.3	76.9	80.3
RBD	2022	77.3	3.5	70.7	74.8	78.7	79.9	81.2
DKB	2022	74.9	3.0	68.2	74.5	76.1	76.8	78.1
aNDF mixed grass and clover								
RBD	2021	468.6	34.9	421.1	437.0	466.1	496.8	523.7
RBD	2022	449.5	37.4	395.0	440.6	462.8	471.7	477.2
DKB	2022	435.7	15.5	413.8	427.3	435.7	443.9	458.1

In the descriptive matrix plots the association between selected outcomes, i.e. lab measurements, and the median values of the predictor spectra variables are provided for DM yield (Figures 8), ME (Figure 9), OMD (Figure 10).

The DM yield was centered around 2000 kg DM/ha, with a range from 1000 to 3000 kg DM/ha. The scatter plots on the outer diagonal illustrate the associations between the median of spectral variables and the outcomes, as well as with one another. A positive association is evident between all spectral variables and DM yield, with correlations ranging from 0.2 to 0.7.

ME centered around 11 MJ/kg DM, and ranges from 9.8 to 12.3. A negative association is identified between most spectral variables (with the exception of the blue spectral value) and ME, with correlations ranging from -0.187 to -0.700. Conversely, the blue spectral value exhibits a positive non statistically significant correlation of 0.156 with ME. Regarding OMD, the distribution is centered around 76%, with values ranging from 68.2% to 81.2%. The blue spectral value shows a positive correlation of 0.370 with OMD, while the other spectral values exhibit negative correlations, ranging from -0.613 to -0.096.

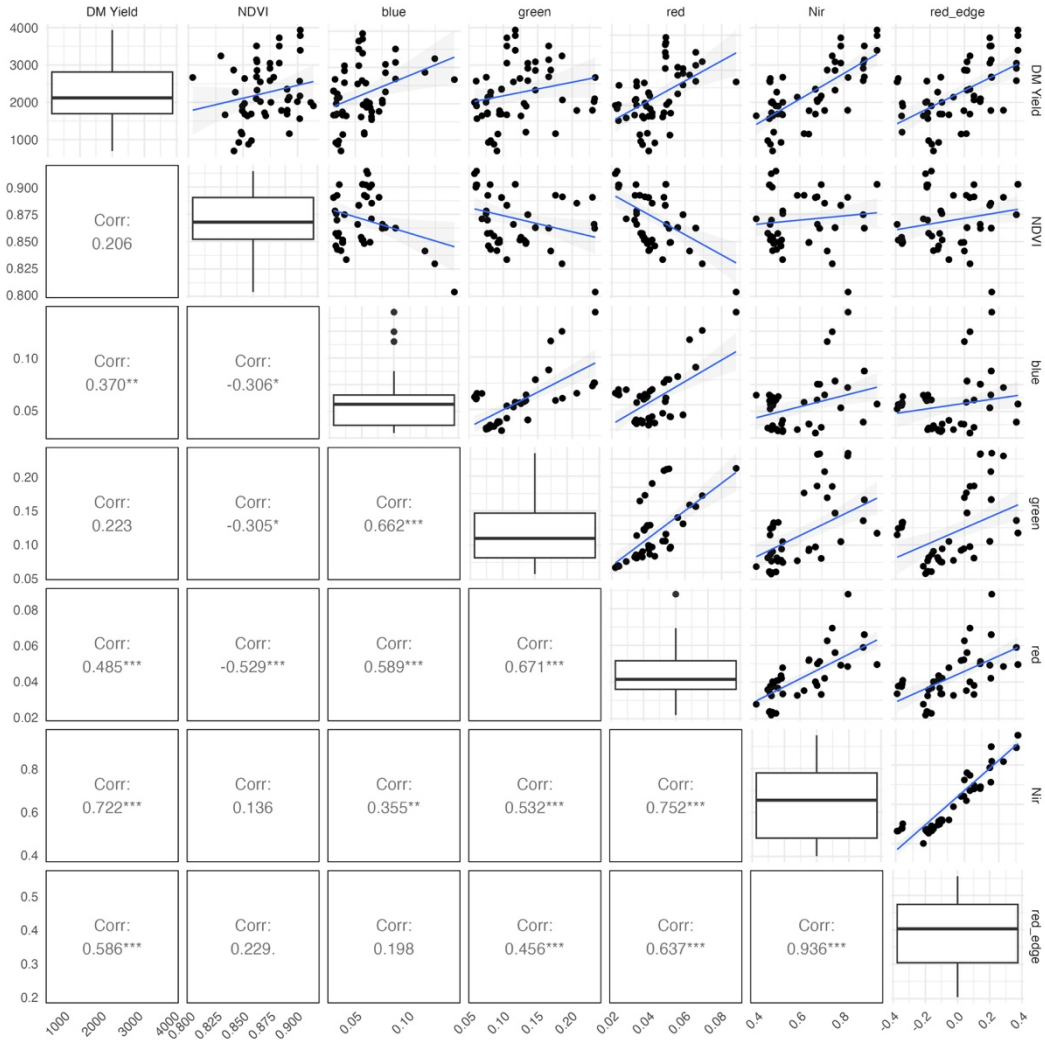

*Figure 8. Correlation matrix and scatterplots depicting the relationships between dry matter yield (DM Yield) and several spectral indices and reflectance values, including NDVI, blue, green, red, NIR (near-infrared), and red-edge bands. * indicates statistically significant differences.*

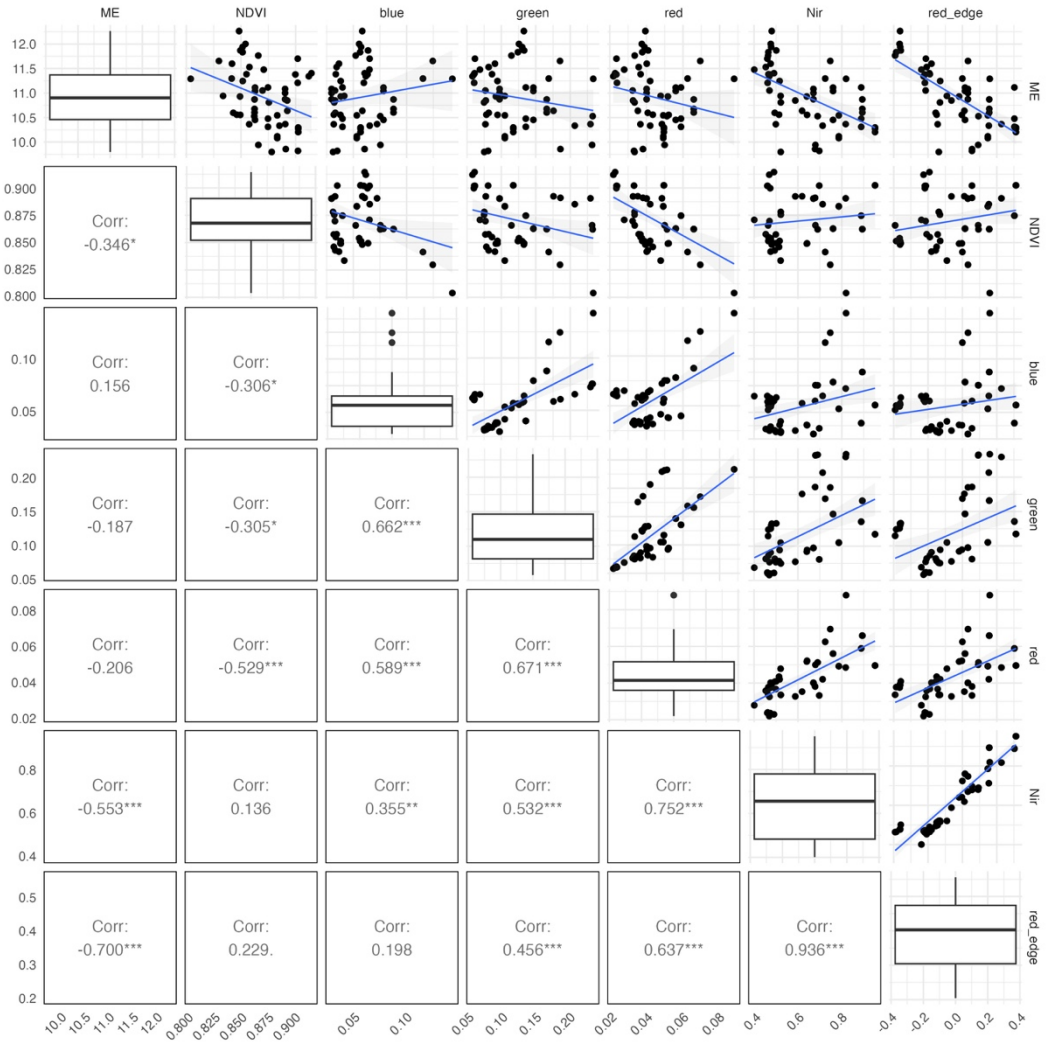

Figure 9. A correlation matrix and scatterplots depicting the relationships between Metabolizable energy (ME) and several spectral indices and reflectance values, including NDVI, blue, green, red, NIR (near-infrared), and red-edge bands. indicates statistically significant differences.*

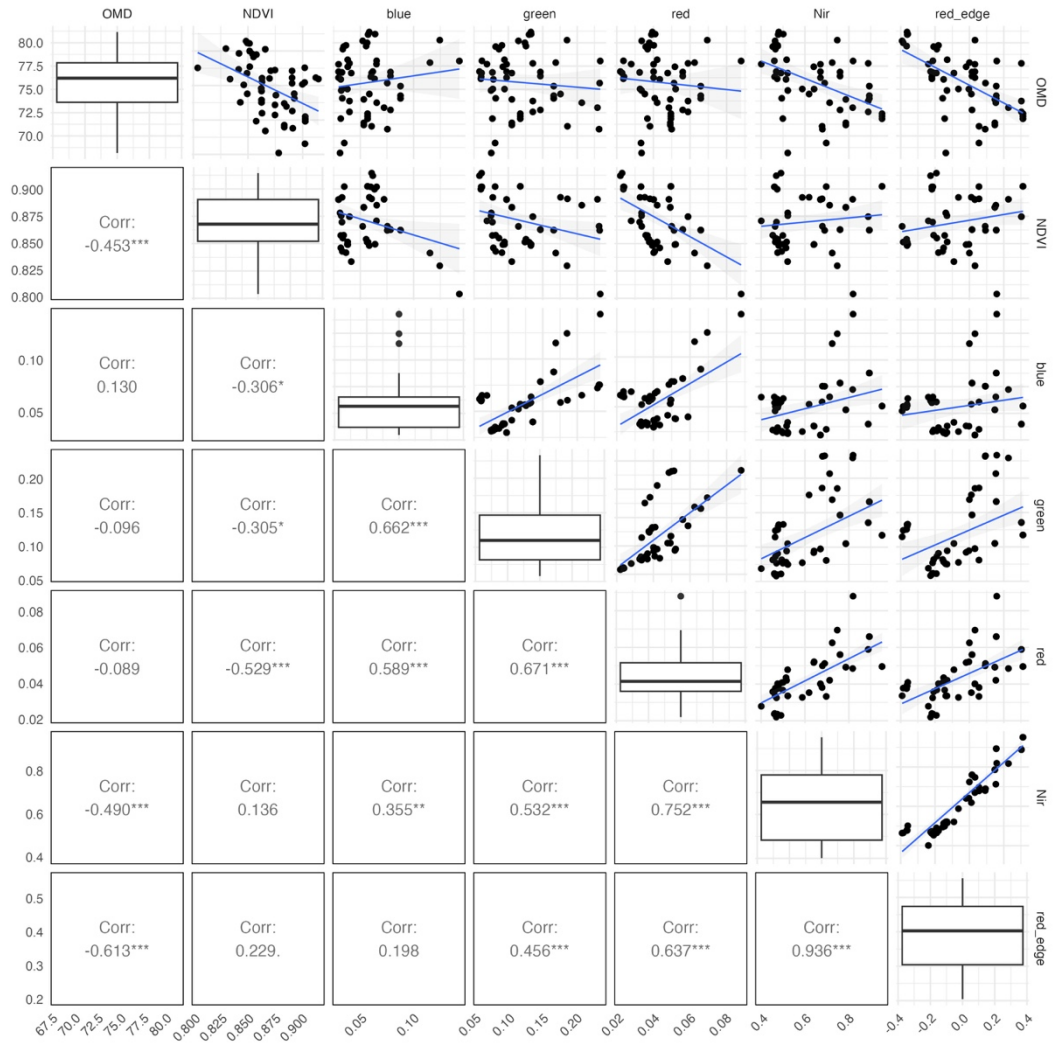

*Figure 10. A correlation matrix and scatterplots depicting the relationships between OMD and several spectral indices and reflectance values, including NDVI, blue, green, red, NIR (near-infrared), and red-edge bands. * indicates statistically significant differences.*

4.3 Prediction models

The predictive performances of the described models are presented in Table 6. For all laboratory measurements, the RMSE and RRMSE indicate a lower prediction error for the ML models compared to the alternative ones. The magnitude of the RMSE depends on the scale of the laboratory measurements considered. In contrast, the RRMSE allows for the comparison of prediction errors across outcomes measured on different scales. A RRMSE ranging from 0.05 to 0.10 was consistently observed, with the exception of DM Yield, where higher values (from 0.16 to 0.22) indicate a greater relative prediction error compared to the other laboratory measurements.

The coefficient of determination (R^2) in Table 6 complements the described prediction errors. Specifically, the R^2 values of ML model is very high, with 0.8

for aNDF in mixed grass and clover and 0.9 for CP, and around 0.7 for predicting OMD (%), ME (MJ/kg DM), and DM Yield (kg DM/ha). These results suggest that the model explains a significant portion of the relationship between the dependent and independent variables. However, the analysis also revealed extremely high VIF values as shown in Table 7, which strongly indicate the presence of multicollinearity among the independent variables. Multicollinearity can lead to unreliable coefficient estimates, making their interpretation problematic. In this case, the high VIF values suggest that some independent variables are highly correlated, which increase the risk for overfitting. The VIF analysis suggests that there is multicollinearity among the spectral bands, confirmed by the high values for the variance inflation factor, which were all above 10.

The R^2 for the alternative models was generally lower, with values around 0.4 for CP mixed grass and clover, between 0.4 and 0.88 for aNDF in mixed grass and clover, around 0.6 for OMD (%), and 0.7 for ME (MJ/kg DM) and DM Yield (kg DM/ha). Because of that, alternative models such as PLS might be preferable, as their collinear predictors are reduced to a lower number of uncorrelated principal components.

Finally, the agreement between observed and predicted laboratory measurements was visually inspected in the scatter plots in Figures 11-14. Overall, a good agreement between observed and predicted laboratory measurements was observed for the ML models, where most points scatter around the diagonal line (y=x), whereas the alternative prediction models show greater variability and weaker consistency. The scatter plots indicate varying levels of agreement between observed and predicted values, with the models generally showing better performances for certain parameters (e.g., aNDF and OMD) than others (e.g., CP).

Table 6. Performance of predictive models.

Model	RMSE	RRMSE	R2
Crude protein mixed grass and clover (g/kg DM)			
ML	6.22	0.04	0.905
PLS	16.01	0.09	0.370
caretPLS	16.45	0.09	0.463
caretSVM	17.74	0.10	0.481
aNDF mixed grass and clover (g/kg DM)			
ML	10.83	0.02	0.897
PLS	11.48	0.02	0.884
caretPLS	27.03	0.06	0.483
caretSVM	21.79	0.05	0.661
OMD (%)			
ML	1.62	0.02	0.734
PLS	1.85	0.02	0.656
caretPLS	2.16	0.03	0.549
caretSVM	2.07	0.03	0.605
ME (MJ/kg DM)			
ML	0.28	0.03	0.797
PLS	0.33	0.03	0.711
caretPLS	0.35	0.03	0.724
caretSVM	0.36	0.03	0.675
DM Yield (kg DM/ha)			
ML	361.0	0.16	0.787
PLS	390.6	0.17	0.751
caretPLS	457.2	0.21	0.690
caretSVM	483.1	0.22	0.677

Table 7. Variance Inflation Factor for ML model.

Variable	Crude protein mixed grass and clover (g/kg DM)	aNDF mixed grass and clover (g/kg DM)	OMD (%)	ME (MJ/kg DM)	DM Yield (kg DM/ha)
Majority_NDVI	448	448	13	13	13
SD_NDVI	337	337	10	10	10
Median_NDVI	1903	1903	51	51	51
Majority_Blue	2164	2164	391	391	391
SD_blue	5852	5852	57	57	57
Median_blue	3832	3832	535	535	535
Majority green	3232	3232	254	254	254
SD_green	20525	20525	143	143	143
Median_green	16520	16520	400	400	400
Majority_red	1877	1877	116	116	116
SD_red	1697	1697	98	98	98
Median_red	1691	1691	320	320	320
Majority_Nir	40065	40065	1054	1054	1054
SD_Nir	370	370	56	56	56
Median_Nir	56965	56965	834	834	834
Majority_red_edge	20130	20130	1055	1055	1055
SD_red_edge	624	624	70	70	70
Median_red_edge	25857	25857	904	904	904

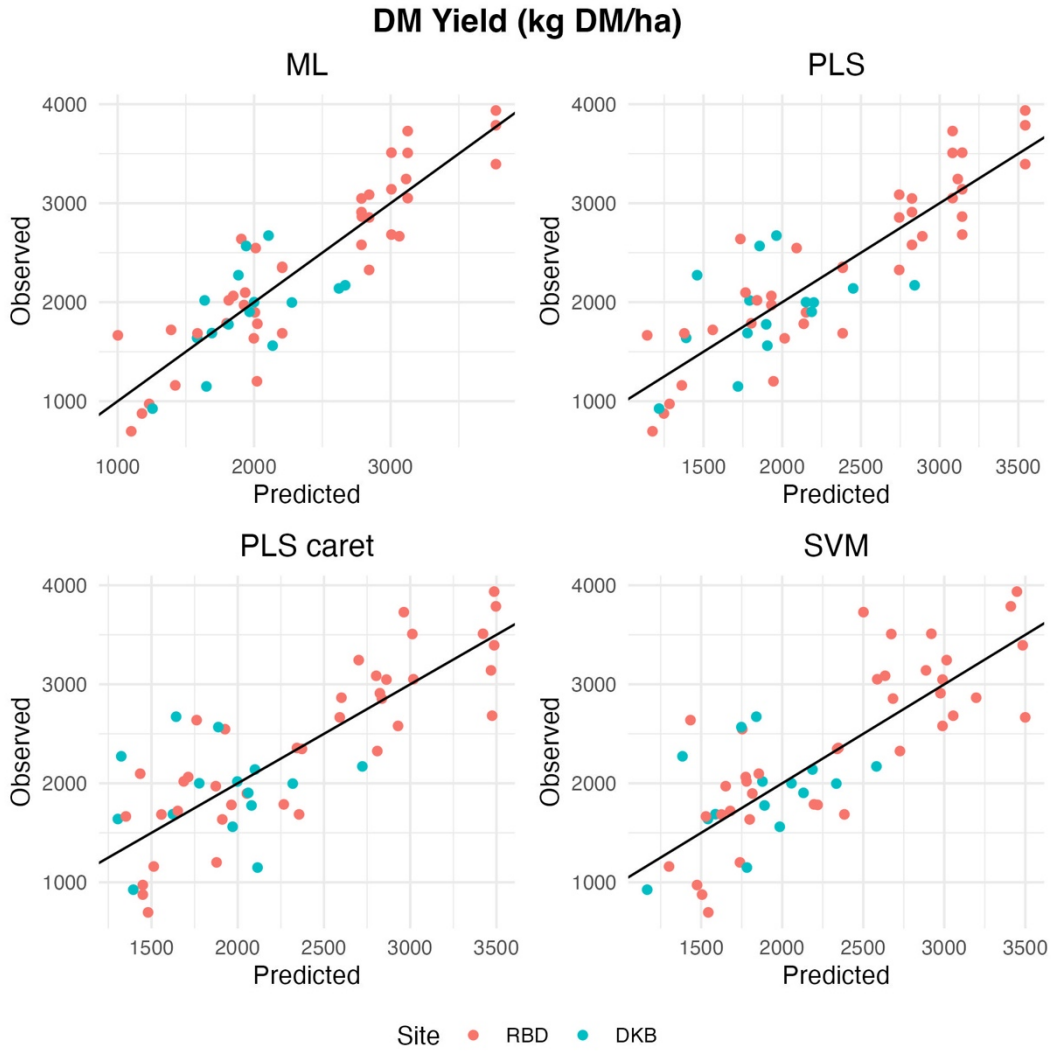

Figure 11. Prediction plot showing the scatterplot between predicted and observed DM Yield (kg DM/ha) from fitted prediction models, colored by site. The solid black line represents the diagonal (y = x).

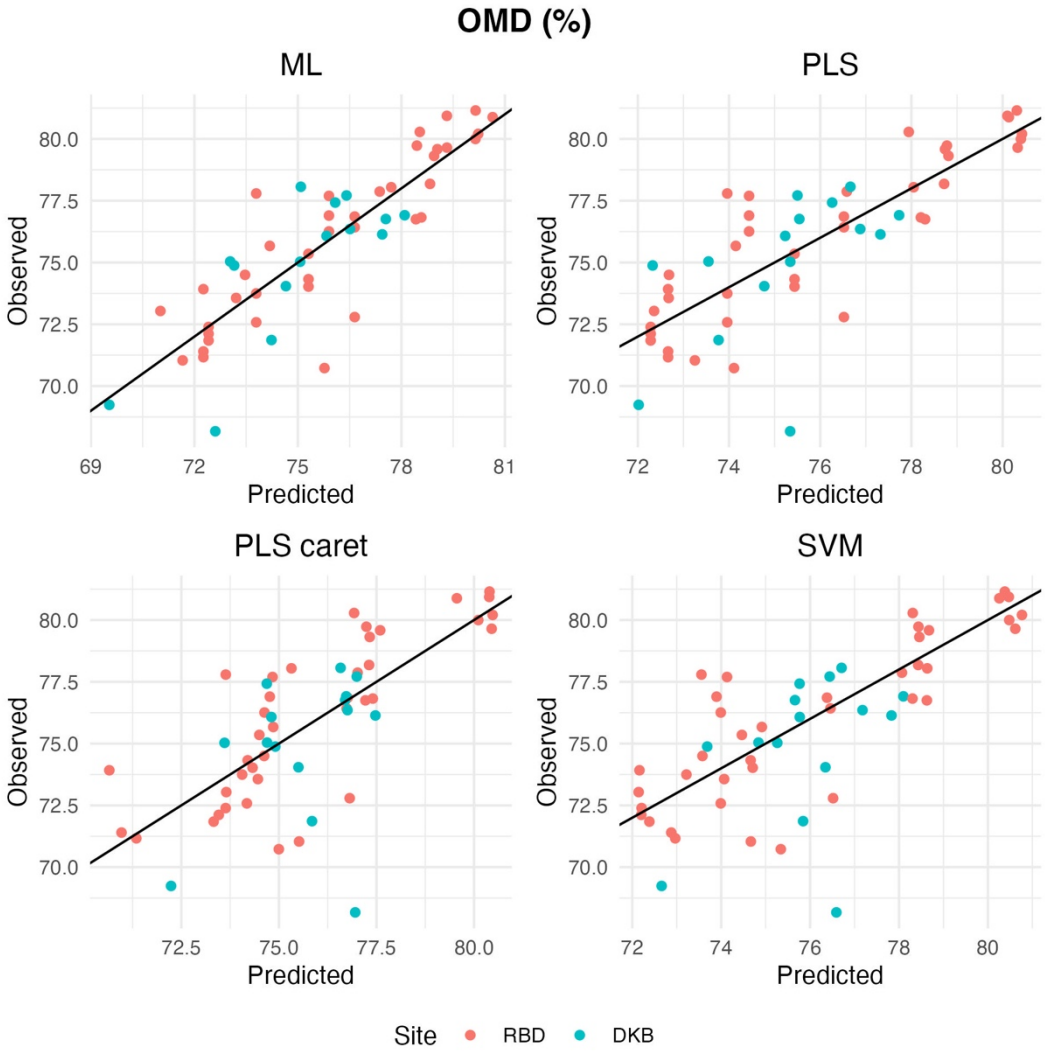

Figure 12. Prediction plot showing the scatterplot between predicted and observed OMD% from fitted prediction models, colored by site. The solid black line represents the diagonal (y = x).

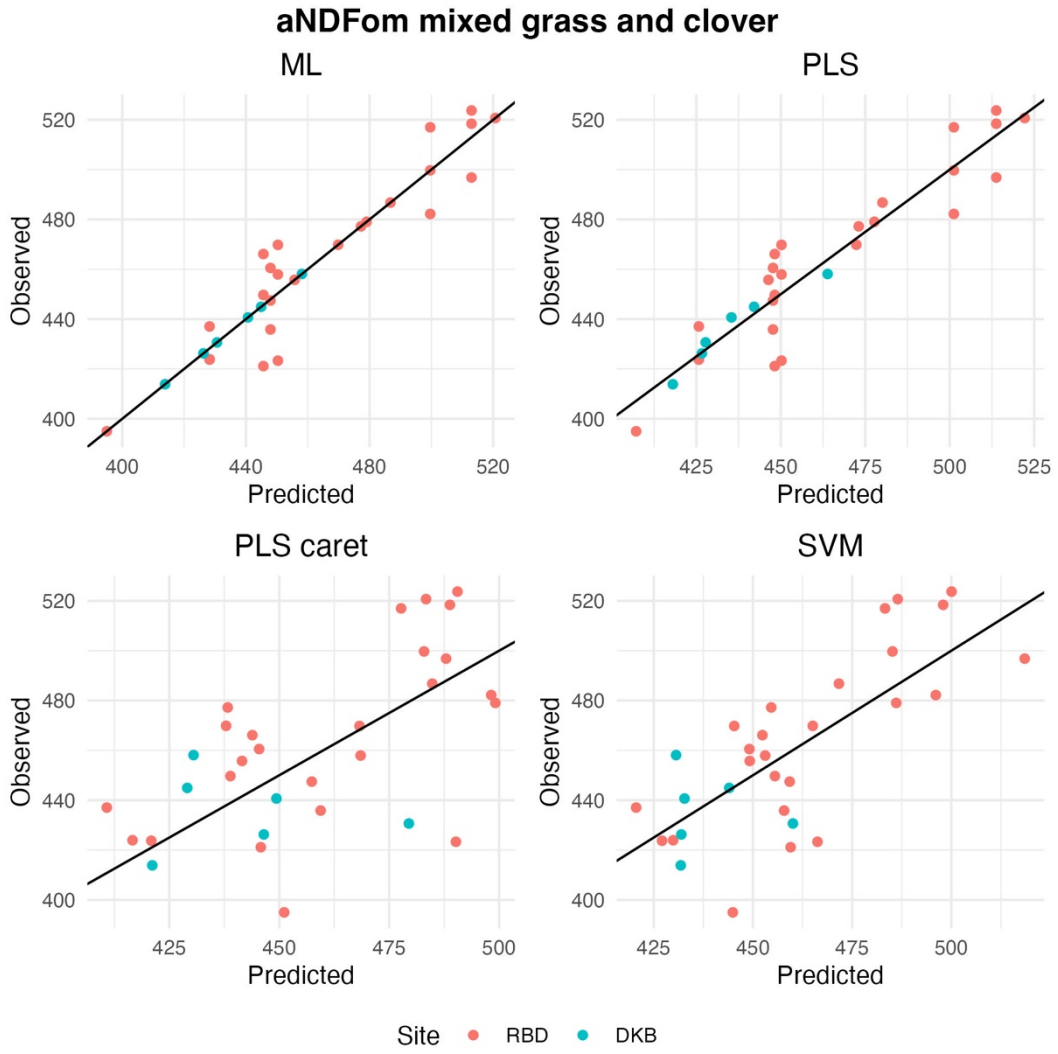

Figure 13. Prediction plot showing the scatterplot between predicted and observed aNDFom mixed grass and clover from fitted prediction models, colored by site. The solid black line represents the diagonal (y = x).

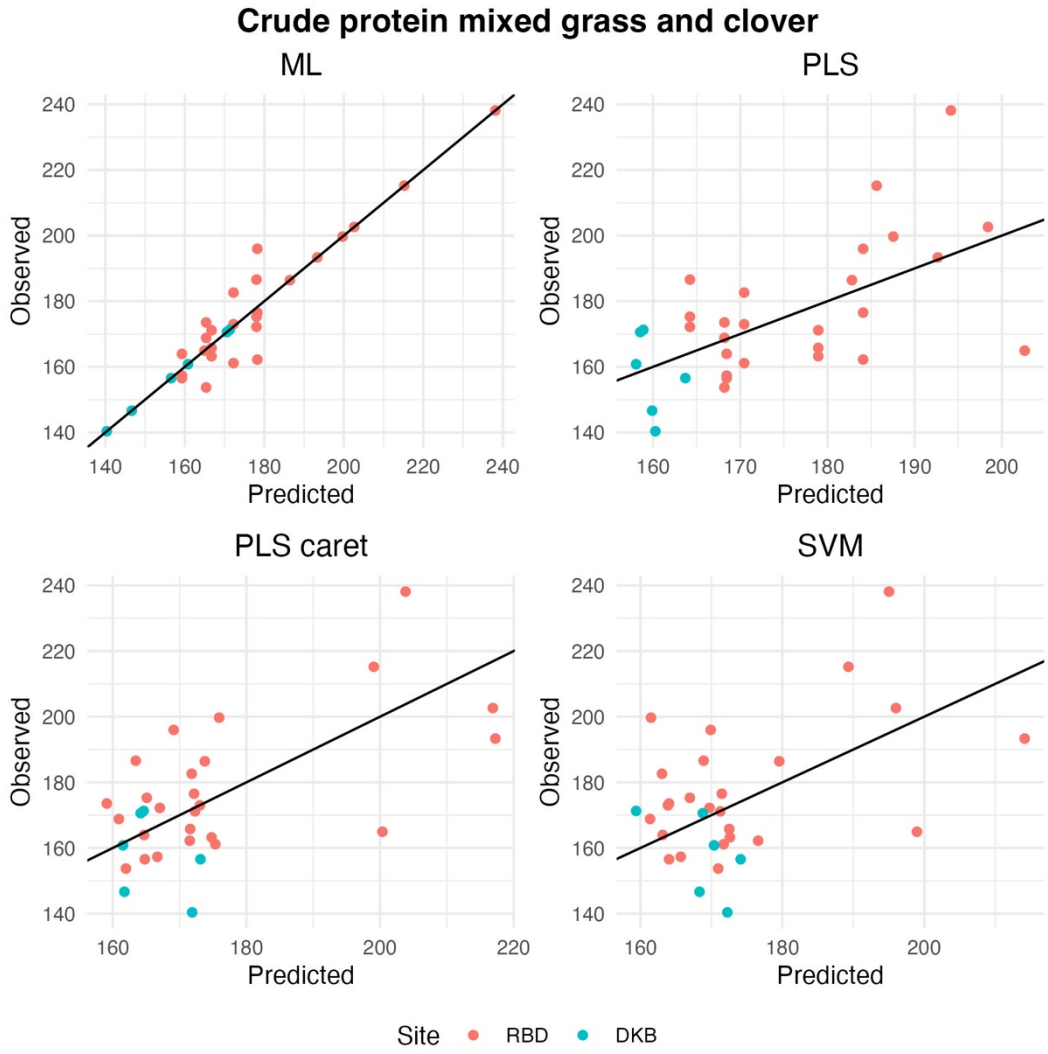

Figure 14. Prediction plot showing the scatterplot between predicted and observed crude protein mixed grass and clover from fitted prediction models, colored by site. The solid black line represents the diagonal (y = x).

5. Discussion

5.1 Evaluating the grassland vegetation using NDVI and reflectance bands across different sites and years

In this study, the NDVI values were generally high across all sites and years, indicating healthy vegetation. The NDVI, which is based on red and NIR reflectance, are associated with the photosynthesis process (Geipel et al. 2014). This vegetation index is usually related to biomass and chlorophyll content as it helps determine the nitrogen content of crops (Shu et al. 2024). The highest mean NDVI value of 0.872 was observed in RBD 2021, with a relatively narrow standard deviation (0.025) indicating uniform vegetation health across the site. In 2022, a slight decrease in mean NDVI (0.862) compared to 2021, along with a narrower range (min: 0.834, max: 0.892), may indicate the different stages of growth. The mean NDVI value of 0.880 in DKB 2022 is slightly higher than RBD in 2022 but lower than RBD in 2021. The maximum value (0.914) suggests better peak vegetation health, but the higher standard deviation (0.026) indicates greater variability in vegetation across this site.

NIR values showed notable variations between sites and years, which correlate with the health and density of the vegetation. The highest mean NIR value (0.816) was recorded for RBD 2021, with a broad range from 0.670 to 0.951. This reflects dense and healthy vegetation. A significant drop in mean NIR (0.553) in 2022 may suggest lower biomass or reduced water content. The minimum value (0.453) further supports this observation. NIR values for DKB were similar to RBD 2022, with a mean of 0.525 and a narrower range (0.401–0.697). Red reflectance, which is absorbed by chlorophyll, showed variations across years and sites. A higher mean red reflectance (0.056) was recorded in RBD 2021, suggesting slightly less chlorophyll absorption compared to 2022. While in 2022, the lowest mean red reflectance (0.040) indicates vegetation with higher chlorophyll content. Similar to RBD 2022, the red reflectance value (0.033) in DKB suggests high photosynthetic activity. Green reflectance values were generally low, consistent with healthy vegetation that absorbs green light for photosynthesis.

5.2 Site and year variability in forage quality: correlations with spectral indices and nutritional parameters

The dataset highlights differences across sites (RBD and DKB) and years (2021 and 2022) in key parameters related to forage quality, including CP, DM yield, ME, OMD and aNDF. CP is a key determinant of forage quality, with its variation typically being positively correlated with precipitation and negatively correlated with annual temperature (Lugassi et al. 2019; Irisarri et al. 2022). Accurate detection of changes in the herbage mass of CP would enable more efficient pasture management, including optimal timing and amount of fertilizer application. NDVI is considered the most appropriate index for assessing CP (Lim et al. 2015) as it demonstrates a stronger correlation with the ratio-based index using red and NIR bands (Watanabe et al. 2014). Furthermore, several studies have reported a correlation between chlorophyll content in plant tissue and nitrogen concentration (Rodriguez & Miller 2000). Nitrogen is a critical component of amino acids, essential for protein synthesis, leading to higher CP content. For this reason, NDVI, which reflects chlorophyll content (as healthy plants with higher chlorophyll absorb red light and reflet NIR), can serve as a reliable indicator of CP. However, the results did not show any significant findings (Figure 15 in the appendix) where a stronger correlation with the green band is observed. Crude protein content varied significantly between years and sites. At RBD, there was a notable increase in crude protein from 171.9 g/kg DM in 2021 to 212.3 g/kg DM in 2022, reflecting improved forage quality potentially influenced by environmental factors or management practices, such as fertilizer application. In contrast, the mean crude protein at DKB (157.7) was the lowest among all groups, indicating reduced forage nutritional quality compared to RBD in the same year. During sample collection, it was observed that the fields appeared to exhibit signs of nitrogen deficiency. The variation in protein content could be attributed to the balance of plant species present, as clover is naturally richer in protein than grasses. Accurately estimating crude protein content in the crop remains fundamental for effective forage monitoring.

DM yield is a key indicator of overall productivity as it measures the total mass of plant material (excluding water content), providing insight into the amount of biomass produced. NDVI is also correlated to biomass as it increases with greater vegetation cover and leaf area (Goswami et al. 2015; Barboza et al. 2023). However, exploring different band combination can reveal variations in the relationship. The weaker correlation observed in this study can be explained by saturation during the peak season as reported in other studies (Zhu & Liu 2015) (Santin-Janin et al. 2009). At higher biomass levels, the relationship between biomass and NDVI may plateau due to increased canopy density, where additional biomass does not result in a significant increase in NDVI. This suggests that NDVI might not be strongly correlated with DM yield in this specific dataset. While higher NDVI values generally correspond to higher biomass and potentially greater DM yield, Table 5 reveals that RBD exhibited significant differences in DM yield between 2021 and 2022, despite relatively small changes in NDVI values for the same site. This variability in DM yield, particularly

between years in RBD, can be attributed to the sampling strategy. Samples were deliberately collected two to three times to harvest rather than at the peak growth stage, ensuring variability in yield and quality. This approach supports the development of more robust and accurate predictive models. However, there appears to be a correlation between DM yield and the spectral indices of NIR, red edge, and red, as they show a higher correlation value as reported in Figure 8. As noted in previous studies, the relationships between vegetation indices, such as NDVI, and biophysical parameters, like biomass, are often non-linear, which can restrict the predictive effectiveness of linear models.

OMD and ME are interconnected, as they all measure the digestible portion of the feed that can be effectively utilized by the animal. OMD reflects the proportion of forage organic matter that can be digested by the animal. It is strongly influenced by plant maturity, as digestibility decreases with age due to the accumulation of structural carbohydrates. In general, immature plant tissue are more digestible than mature, stem-dominated material (Ball et al. 2001). The plant species also plays a significant role in digestibility. For example, clovers tend to maintain relatively higher digestibility compared to grasses as they mature, as they preserve a higher leaf to stem ratio (Forejtová et al. 2005). In case of timothy grass, when the temperature sum reaches 250 degrees-days, the OMD is estimated to be around 76% which is considered a suitable lower limit for high quality for dairy cows (*Hjälp* 2024). In this study, OMD values ranged from 70% to 81%. The relationship with spectral indices shows a strong negative correlation with red edge band (-0.613) and NIR band (-0.490).

ME estimates the portion of energy content in feed that is absorbed through the digestive tract and utilized for metabolic process in ruminant animal, excluding energy lost in feces, urine and gases. Values of ME were relatively stable across sites and years, with minor differences. At RBD, an increase in ME from 10.7 (MJ/kg DM) in 2021 to 11.2 (MJ/kg DM) in 2022 reflects slightly improved energy content in forage. Meanwhile, the ME at DKB (10.9) was comparable to RBD in 2021, indicating similar energy availability in forage. The relationships between ME and spectral indices shows a strong negative correlation with red edge band (-0.7).

The aNDF represents the fibrous component of forage, which varies based on species, maturity and growing conditions. It includes the slowly digested hemicellulose and cellulose, as well as indigestible lignin within the plant. Adequate aNDF in diet is essential for dairy cows to maintain proper rumen function and optimize milk production (Oba & Allen 1999). Slight changes in aNDF values contributed to the training of the predictive models. A decrease from 468.6 at RBD in 2021 to 450 in 2022 suggests that lower aNDF indicates reduced fiber content and greater digestibility. The lowest aNDF value (435.7) was observed at DKB in 2022, highlighting improved digestibility compared to RBD in 2021. All the reflectance values showed a positive correlation with aNDF. In particular, NIR and red-edge wavebands exhibited significant correlation with this forage quality variable, as shown in Figure 16 in the appendix.

5.3 Evaluating the performance of predictive models for forage quality parameters

The evaluation of multiple models for predicting forage quality parameters, such as CP, aNDF, OMD, ME, and DM yield, revealed varying levels of accuracy and reliability. In this study, VIF was used to assess multicollinearity among spectral indices within ML prediction model and its impact on forage quality estimation. The findings reveal that multiple spectral indices elevated VIF values, indicating considerable overlap and redundancy within the dataset. This multicollinearity poses challenges in predictive modelling, as highly correlated variables can inflate variance and affect result interpretation. The ML model consistently outperformed other models across most parameters, as evidenced by lower RMSE, lower RRMSE, and higher R^2 values. However, it also experienced significant multicollinearity issues due to strong correlations among spectral bands, potentially undermining model stability and interpretability. Exceptionally high VIF values further confirmed this issue, indicating that some predictors were excessively correlated, which could lead to unreliable results and overfitting.

In contrast, PLS, caretPLS, and SVM models were less affected by multicollinearity, as they transform correlated variables into uncorrelated components. This approach effectively mitigates redundancy, enhances model robustness, and improves predictive performance.

For CP prediction, the missing data likely disrupted the relationship between the variables, reducing the model ability to make accurate prediction. The ML model demonstrated the highest accuracy ($R^2 = 0.905$), with a significantly lower RMSE (6.22) and RRMSE (0.04) compared to other models. However, the presence of missing values in the dataset significantly affected the predictive performance of ML. This limitation is evident in Table 8, which highlights issues of multicollinearity and predictor ability. The PLS, PLS caret, and SVM models showed lower R^2 values (ranging from 0.370 to 0.481), indicating a limited ability to accurately predict crude protein.

Similarly, for aNDF, ML achieved superior results ($R^2 = 0.897$), with the lowest RMSE (10.83) and RRMSE (0.02). However, these results may not be entirely reliable due to evidence of multicollinearity, which raises concerns about the model's interpretability. While the PLS model achieved a comparable R^2 value (0.884), its RMSE (11.48) and RRMSE (0.02) suggest slightly lower performance. The PLS caret and SVM caret models demonstrated weaker predictive capabilities, with R^2 values of 0.483 and 0.661, respectively. Missing values have reduced the predictor ability.

For OMD, ML achieved moderate performance with an R^2 of 0.734, while the PLS, PLS caret, and SVM models showed slightly lower accuracy, with R^2 values ranging from 0.549 to 0.656. Although ML performed better, the relatively lower R^2 for all models compared to other parameters suggests that OMD may be more challenging to predict accurately using these models. The prediction of ME showed similar trends, with ML outperforming the other models ($R^2 = 0.797$, RMSE = 0.28). The PLS and caretPLS models yielded R^2 values of 0.711 and 0.724, respectively, indicating relatively good performance.

The analysis shows that the constructed models exhibited good performance in predicting quality parameters using UAV-collected multispectral data. PLS and

caretPLS demonstrated superior predictive ability with predictive R^2 ranging between 55% to 70% for ME, DM yield and OMD. SVM exhibited variation in performance, with values ranging from 48% to 67%. These models performed better with some variables than others, possibly because not all variables are strongly correlated with the wavelengths or because a larger number of observations may be needed.

The results demonstrate that the ML model consistently provides superior performance is particularly evident for crude protein, and aNDF, where it achieves both high R^2 values and low error metrics. However, despite these strong results, ML model was affected by certain limitations, particularly multicollinearity within the spectral data and the missing data in the dataset. These issues likely reduced the model's reliability and prediction accuracy, and further refinement is needed to address data issues.

While the results indicated good parameter estimation, the validation was internal, as the data used came from the same fields and methods. Although cross-validation can help reduce overfitting, it has its limitations. To enhance the model's reliability, external validation should be performed by testing the models with data from new fields and missions. This would allow for assessing their performance in different conditions, ultimately confirming whether the models can be successfully applied beyond the original study area.

Many studies have shown that using multiple vegetation indices derived from a combination of different spectral bands can improve nitrogen estimation, helping to understand stress variability and its potential agronomic benefits.

A notable strength of this study lies in the utilization of a multiyear dataset collected from diverse field conditions. This approach enables the development of models that account for a wide range of variables, including temperature fluctuations, varying weather patterns, and differences in soil properties, all of which significantly influence crop health. Furthermore, the adoption of a cost-effective multispectral imaging sensor mounted on UAVs to estimate forage agronomic traits presents a clear advantage over traditional near-ground measurement methods, which are frequently invasive and destructive.

6. Conclusions

- The use of UAV-mounted multispectral sensors for estimating forage quality may present advantages over traditional methods. However, data issues such as missing values and multicollinearity within the spectral bands affected the model's performance, suggesting that further refinement is needed for reliable predictions.
- Multilinear regression (ML) models showed the highest R^2 and lowest RMSE values all forage quality characteristics. However, high values for the variance inflation factor (VIF > 10) suggested multicollinearity between predictor variables and, therefore, occurrence of overfitting, which limits their applications on other datasets.
- Alternative modelling approaches such as PLS showed relatively good performance for predicting aNDF (R^2 = 0.88, RMSE = 11.48 g/kg DM), OMD (R^2 = 0.66, RMSE = 1.85%), ME (R^2 = 0.71, RMSE = 0.33 MJ/kg DM) and DM yield (R^2 = 0.75, RMSE = 390.56 kg DM/ha).
- PLS and SVM models showed limited ability to accurately CP (R^2 ranging from 0.37 to 0.48, and RMSE ranging from 16.01 to 17.74 g/kg DM).

General perspectives and recommendations:

- UAV data acquisition and analysis offers significant potential for near real-time decision-making in agriculture, enabling targeted interventions within days and reducing management costs. Moreover, by assessing the quality of the forage, the model can ensure that the forage meets the nutritional needs of livestock, potentially improving animal health and productivity. Higher quality forage may also reduce the need for supplementary feeding, which can lower costs and improve the efficiency of the operation.
- Future studies should adopt more accurate protocols to ensure the collection of complete and reliable datasets. For instance, placing reflectance panels on the ground during data acquisition would improve calibration and prevent data loss.
- A deeper analysis of forage quality parameters across different harvest times is recommended to gain further insights, as this aspect was limited in the current study due to time constraints.

7. Reflections

During this study, I had the opportunity to learn how to process images with PIX4D and gain a deeper understanding of the most common applications of UAVs in Precision Agriculture. I learned to acquire data across multiple wavelengths captured by a drone to generate high-resolution orthomosaic images and digital surface models, ensuring that pixels are perfectly aligned with their actual geographic positions by using GCPs and triangulation to guarantee spatial accuracy.

Using QGIS, I developed the ability to compute various statistics based on the values of raster cells within zones defined by another raster or vector dataset. Additionally, I learned to use the R programming language to create scripts for regression analysis aimed at predicting agronomic variables, employing cross-validation to train my model effectively.

References

Ailioaie, L. & Litscher, G. (2020). Molecular and Cellular Mechanisms of Arthritis in Children and Adults: New Perspectives on Applied Photobiomodulation. *International journal of molecular sciences*, 21. https://doi.org/10.3390/ijms21186565

Ball, D., Collins, M., Lacefield, G., Martin, N., Mertens, D., Olson, K., Putnam, D., Undersander, D. & Wolf, M. (2001). Understanding forage quality.

Barboza, T.O.C., Ardigueri, M., Souza, G.F.C., Ferraz, M.A.J., Gaudencio, J.R.F. & Santos, A.F. dos (2023). Performance of Vegetation Indices to Estimate Green Biomass Accumulation in Common Bean. *AgriEngineering*, 5 (2), 840–854. https://doi.org/10.3390/agriengineering5020052

Capstaff, N.M. & Miller, A.J. (2018). Improving the Yield and Nutritional Quality of Forage Crops. *Frontiers in Plant Science*, 9. https://doi.org/10.3389/fpls.2018.00535

DJI Phantom 4 Multispectral (2025). *Din DJI butik i Sverige- DJI STOCKHOLM*. https://djistockholm.se/product/phantom/phantom-4-multispectral/ [2024-12-31]

Facts about Swedish agriculture (2009). https://webbutiken.jordbruksverket.se/sv/artiklar/facts-about-swedish-agriculture.html [2025-02-04]

Forejtová, J., Lád, F., Třináctý, J., Richter, M., Gruber, L., Doležal, P., Homolka, P. & Pavelek, L. (2005). Comparison of organic matter digestibility determined by in vivo and in vitro methods. *Czech Journal of Animal Science*, 50 (2), 47–53. https://doi.org/10.17221/3994-CJAS

Geipel, J., Link, J. & Claupein, W. (2014). Combined Spectral and Spatial Modeling of Corn Yield Based on Aerial Images and Crop Surface Models Acquired with an Unmanned Aircraft System. *Remote Sensing*, 6 (11), 10335–10355. https://doi.org/10.3390/rs61110335

Geodata Extraction Tool (2025). *SLU*. https://dike.slu.se/get/ [2025-01-23]

Goswami, S., Gamon, J., Vargas, S. & Tweedie, C. (2015). Relationships of NDVI, Biomass, and Leaf Area Index (LAI) for six key plant species in Barrow, Alaska. PeerJ PrePrints. https://doi.org/10.7287/peerj.preprints.913v1

Hjälp (2024). http://www.vallprognos.se/hjalp.html [2025-01-25]

Huang, C.Z., Ling, J. & Wang, J. (2019). *Elastic Light Scattering Spectrometry*. De Gruyter. https://doi.org/10.1515/9783110573138

Huang, S., Tang, L., Hupy, J.P., Wang, Y. & Shao, G. (2021). A commentary review on the use of normalized difference vegetation index (NDVI) in the era of popular remote sensing. *Journal of Forestry Research*, 32 (1), 1–6. https://doi.org/10.1007/s11676-020-01155-1

Irisarri, J.G.N., Durante, M., Derner, J.D., Oesterheld, M. & Augustine, D.J. (2022). Remotely Sensed Spatiotemporal Variation in Crude Protein of Shortgrass Steppe Forage. *Remote Sensing*, 14 (4), 854. https://doi.org/10.3390/rs14040854

Jordbruksmarkens användning 2023. Slutlig statistik (2023). [text]. https://jordbruksverket.se/om-jordbruksverket/jordbruksverkets-officiella-statistik/jordbruksverkets-statistikrapporter/statistik/2024-02-07-jordbruksmarkens-anvandning-2023.-slutlig-statistik [2025-01-08]

Krizsan, S.J., Chagas, J.C., Pang, D. & Cabezas-Garcia, E.H. (2021). Sustainability aspects of milk production in Sweden. *Grass and Forage Science*, 76 (2), 205–214. https://doi.org/10.1111/gfs.12539

Li, L., Zhang, Q. & Huang, D. (2014). A Review of Imaging Techniques for Plant Phenotyping. *Sensors*, 14 (11), 20078–20111. https://doi.org/10.3390/s141120078
Lim, J., Kawamura, K., Lee, H.-J., Yoshitoshi, R., Kurokawa, Y., Tsumiyama, Y. & Watanabe, N. (2015). Evaluating a hand-held crop-measuring device for estimating the herbage biomass, leaf area index and crude protein content in an Italian ryegrass field. *Grassland Science*, 61 (2), 101–108. https://doi.org/10.1111/grs.12083
Lugassi, R., Zaady, E., Goldshleger, N., Shoshany, M. & Chudnovsky, A. (2019). Spatial and Temporal Monitoring of Pasture Ecological Quality: Sentinel-2-Based Estimation of Crude Protein and Neutral Detergent Fiber Contents. *Remote Sensing*, 11 (7), 799. https://doi.org/10.3390/rs11070799
Lum, C., Mackenzie, M., Shaw-Feather, C., Luker, E. & Dunbabin, M. (2016). *Multispectral Imaging and Elevation Mapping from an Unmanned Aerial System for Precision Agriculture Applications*.
Multispectral imaging (2024). *GEO Innoter*. https://innoter.com/en/articles/multispectral-imaging/ [2024-12-31]
Ngabire, M., Wang, T., Xue, X., Liao, J., Sahbeni, G., Huang, C., Duan, H. & Song, X. (2022). Soil salinization mapping across different sandy land-cover types in the Shiyang River Basin: A remote sensing and multiple linear regression approach. *Remote Sensing Applications: Society and Environment*, 28, 100847. https://doi.org/10.1016/j.rsase.2022.100847
Oba, M. & Allen, M.S. (1999). Evaluation of the Importance of the Digestibility of Neutral Detergent Fiber from Forage: Effects on Dry Matter Intake and Milk Yield of Dairy Cows. *Journal of Dairy Science*, 82 (3), 589–596. https://doi.org/10.3168/jds.S0022-0302(99)75271-9
Oregon State University (2019). *Factors affecting forage quality*. *Forage Information System*. https://forages.oregonstate.edu/oregon/topics/quality-testing/factors-affecting-forage-quality [2025-01-02]
Pix4D Documentation (2024). https://support.pix4d.com/hc/en-us/articles/202739409 [2024-12-31]
Plant, R.E. (2018). *Spatial Data Analysis in Ecology and Agriculture Using R*. 2. ed. CRC Press. https://doi.org/10.1201/9781351189910
Pontes, L.S., Carrère, P., Andueza, D., Louault, F. & Soussana, J.F. (2007). Seasonal productivity and nutritive value of temperate grasses found in semi-natural pastures in Europe: responses to cutting frequency and N supply. *Grass and Forage Science*, 62 (4), 485–496. https://doi.org/10.1111/j.1365-2494.2007.00604.x
Pranga, J., Borra-Serrano, I., Aper, J., De Swaef, T., Ghesquiere, A., Quataert, P., Roldán-Ruiz, I., Janssens, I.A., Ruysschaert, G. & Lootens, P. (2021). Improving Accuracy of Herbage Yield Predictions in Perennial Ryegrass with UAV-Based Structural and Spectral Data Fusion and Machine Learning. *Remote Sensing*, 13 (17), 3459. https://doi.org/10.3390/rs13173459
Rodriguez, I.R. & Miller, G.L. (2000). Using a Chlorophyll Meter to Determine the Chlorophyll Concentration, Nitrogen Concentration, and Visual Quality of St. Augustinegrass. *HortScience*, 35 (4), 751–754. https://doi.org/10.21273/HORTSCI.35.4.751
Santin-Janin, H., Garel, M., Chapuis, J.-L. & Pontier, D. (2009). Assessing the performance of NDVI as a proxy for plant biomass using non-linear models: a case study on the Kerguelen archipelago. *Polar Biology*, 32 (6), 861–871. https://doi.org/10.1007/s00300-009-0586-5
Shu, M., Wang, Z., Guo, W., Qiao, H., Fu, Y., Guo, Y., Wang, L., Ma, Y. & Gu, X. (2024). Effects of Variety and Growth Stage on UAV Multispectral

Estimation of Plant Nitrogen Content of Winter Wheat. *Agriculture*, 14 (10). https://doi.org/10.3390/agriculture14101775
Tilman, D., Balzer, C., Hill, J. & Befort, B.L. (2011). Global food demand and the sustainable intensification of agriculture. *Proceedings of the National Academy of Sciences*, 108 (50), 20260–20264. https://doi.org/10.1073/pnas.1116437108
Tsouros, D.C., Bibi, S. & Sarigiannidis, P.G. (2019a). A Review on UAV-Based Applications for Precision Agriculture. *Information*, 10 (11), 349. https://doi.org/10.3390/info10110349
Tsouros, D.C., Bibi, S. & Sarigiannidis, P.G. (2019b). A Review on UAV-Based Applications for Precision Agriculture. *Information*, 10 (11), 349. https://doi.org/10.3390/info10110349
Vallprognos.se (2025). http://www.vallprognos.se/ [2025-01-21]
Wachendorf, M., Fricke, T. & Möckel, T. (2018). Remote sensing as a tool to assess botanical composition, structure, quantity and quality of temperate grasslands. *Grass and Forage Science*, 73 (1), 1–14. https://doi.org/10.1111/gfs.12312
Watanabe, N., Sakanoue, S., Lee, H.-J., Lim, J., Yoshitoshi, R. & Kawamura, K. (2014). Use of a hand-held crop growth measuring device to estimate forage crude protein mass of pasture. *Grassland Science*, 60 (4), 214–224. https://doi.org/10.1111/grs.12064
World Administrative Boundaries - Countries and Territories (2025). https://public.opendatasoft.com/explore/dataset/world-administrative-boundaries/export/ [2025-02-10]
Zhu, X. & Liu, D. (2015). Improving forest aboveground biomass estimation using seasonal Landsat NDVI time-series. *ISPRS Journal of Photogrammetry and Remote Sensing*, 102, 222–231. https://doi.org/10.1016/j.isprsjprs.2014.08.014

Popular science summary

For livestock farmers, forage quality is a key factor affecting animal health, productivity, and the nutritional value of meat and milk. However, laboratory analyses required to assess its characteristics can be costly and time-consuming. This study explored the use of drones equipped with advanced sensors to collect data on forage crops, highlighting the potential of remote sensing in precision agriculture. Due to their small size, drones can be effectively used in agriculture to capture high-resolution images. In this study, the drone was equipped with a multispectral camera capable of capturing images at different wavelengths. The data obtained were processed using specialized software, allowing the calculation of statistical parameters to develop predictive models for estimating forage quantity and quality. Four different statistical models were tested, giving variable results: some proved to be more accurate than others in making predictions. A major advantage of this approach is that the models were tested across different locations and time periods, showing adaptability. However, further validation using additional datasets will be needed to confirm their reliability. Despite its potential, integrating drone technology into agriculture comes with challenges. The vast amount of data collected by onboard sensors requires sophisticated software and specialized expertise to interpret effectively. Overcoming these obstacles will be crucial for fully leveraging drone-based remote sensing in precision agriculture. As technology continues to evolve, drones may soon become indispensable tools for farmers, helping them optimize crop management and improve efficiency.

Appendix

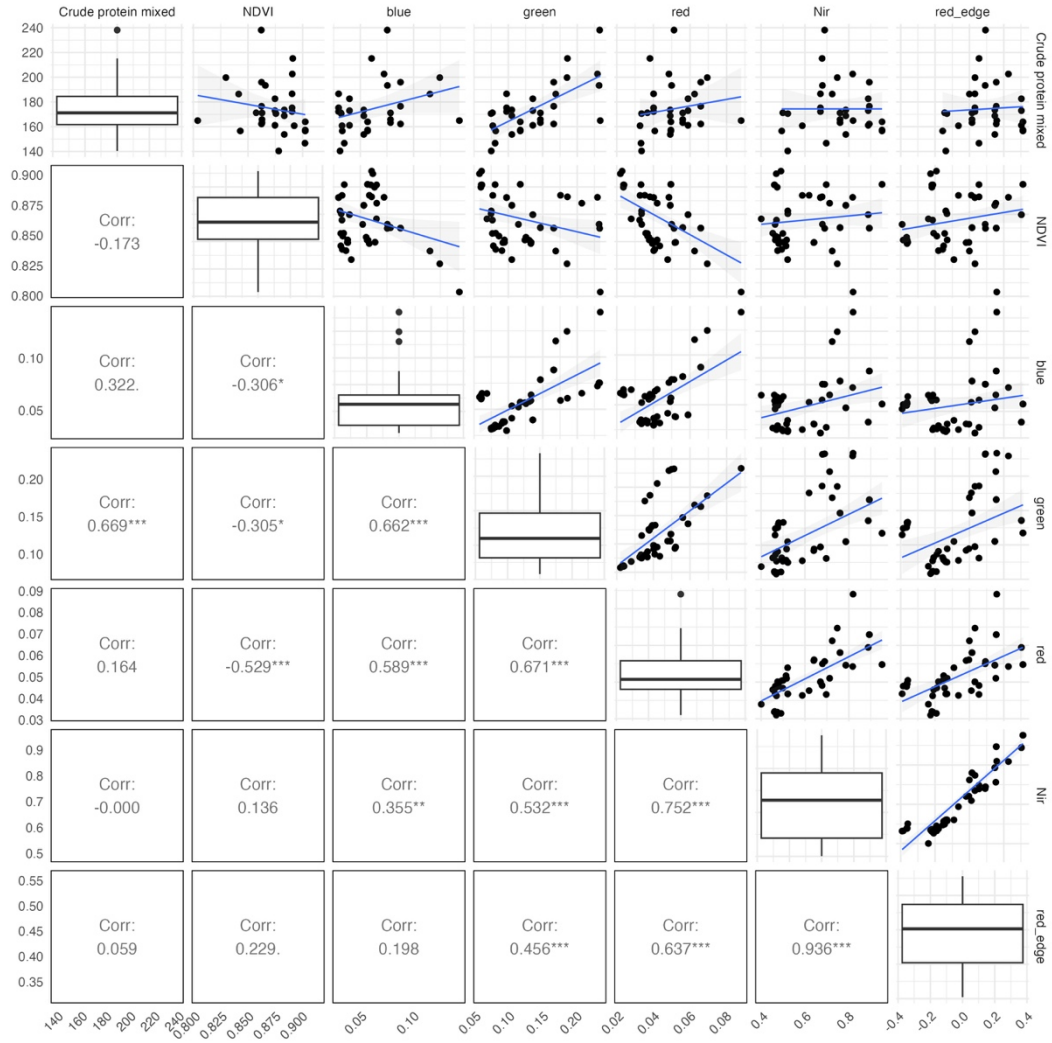

*Figure 15. Correlation matrix and scatterplots depicting the relationships between Crude Protein (CP) and several spectral indices and reflectance values, including NDVI, blue, green, red, NIR (near-infrared), and red-edge bands. * indicates statistically significant differences.*

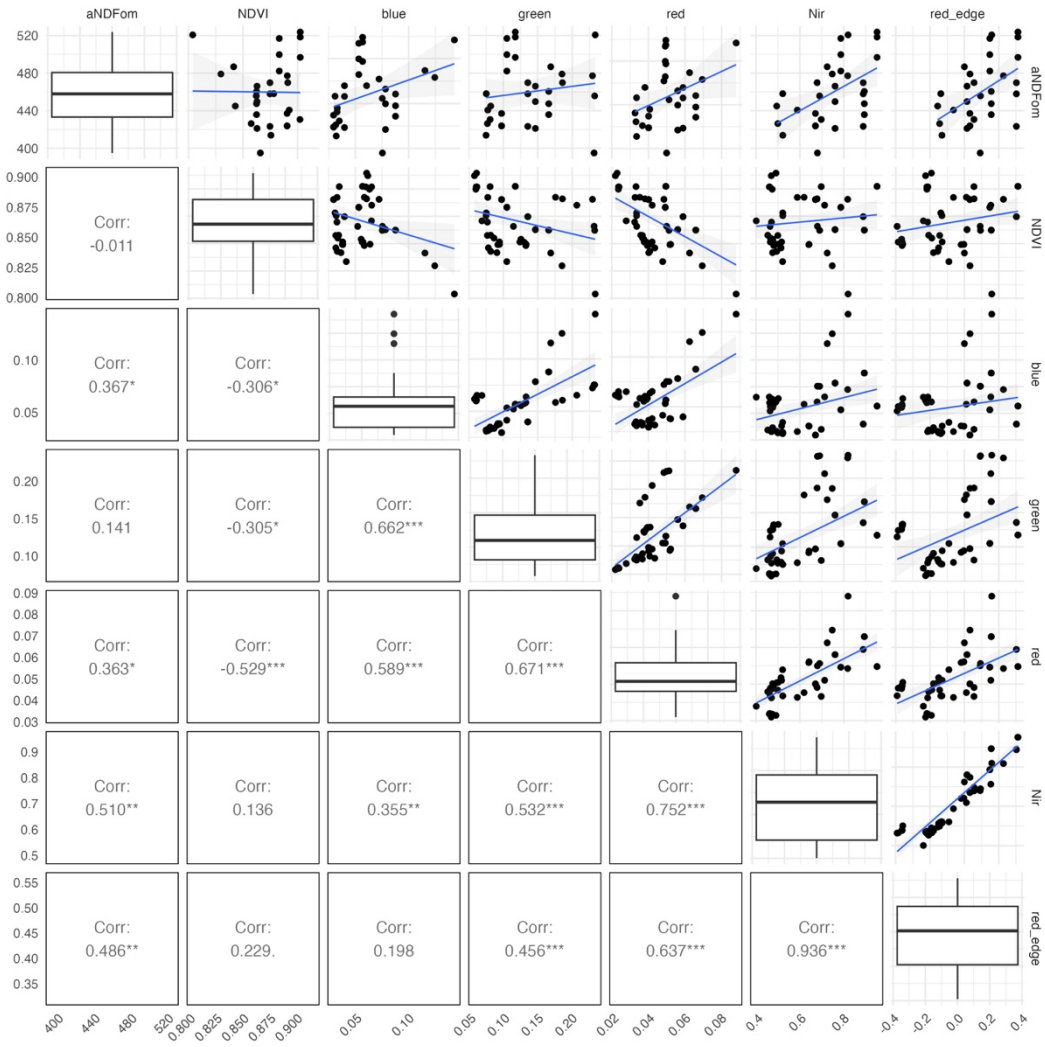

*Figure 16. Correlation matrix and scatterplots depicting the relationships between aNDF and several spectral indices and reflectance values, including NDVI, blue, green, red, NIR (near-infrared), and red-edge bands. * indicates statistically significant differences.*

Publishing and archiving

Approved students' theses at SLU can be published online. As a student you own the copyright to your work and in such cases, you need to approve the publication. In connection with your approval of publication, SLU will process your personal data (name) to make the work searchable on the internet. You can revoke your consent at any time by contacting the library.

Even if you choose not to publish the work or if you revoke your approval, the thesis will be archived digitally according to archive legislation.

You will find links to SLU's publication agreement and SLU's processing of personal data and your rights on this page:

- https://libanswers.slu.se/en/faq/228318

☒ YES, I, Sarah Carola Vecchietti, have read and agree to the agreement for publication and the personal data processing that takes place in connection with this

☐ NO, I, Sarah Carola Vecchietti do not give my/our permission to publish the full text of this work. However, the work will be uploaded for archiving and the metadata and summary will be visible and searchable.