Computers and Electronics in Agriculture 216 (2024) 108463

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

Onfield estimation of quality parameters in alfalfa through hyperspectral spectrometer data

Angie L. Gámez [a], Thomas Vatter [b,c], Luis G. Santesteban [d], Jose Luis Araus [b,c], Iker Aranjuelo [e,*]

[a] NAFOSA company, Avenida Leizaur 79, 31350, Peralta, Navarre, Spain
[b] Integrative Crop Ecophysiology Group, Plant Physiology Section, Faculty of Biology, University of Barcelona, 08028, Barcelona, Spain
[c] AGROTECNIO (Center of Research in Agrotechnology), Av. Rovira Roure 191, 25198, Lleida, Spain
[d] Department of Agronomy, Biotechnology and Food Science, Public University of Navarre (UPNA), 31006, Pamplona, Spain
[e] Agrobiotecnology Institute (IdAB), CSIC, Government of Navarre, 31192, Mutilva Baja, Spain

ARTICLE INFO

Keywords:
Hyperspectral technique
Quality parameters
Alfalfa
Canopy
Trait prediction

ABSTRACT

Alfalfa is a forage of vast importance around the world. In the past, near-infrared spectroscopy (NIRS) technique have been explored in the lab to determine quality traits such as fibre content in dried and ground material. During the last decade, portable hyperspectral devices have emerged as a tools for in-field prediction, of not only crop yield but also a large range of quality and physiological markers. The objective of this study was to estimate quality parameters in an alfalfa crop using hyperspectral data acquired from a full-range (350–2500 nm) spectrometer under field conditions. Reflected spectra were measured in single leaves as well as at the canopy level, then reflectance was related to target parameters such as biomass, leaf pigments, sugars, protein, and mineral contents. Due to their large effect on crop quality parameters, meteorological conditions and phenological stages were included as predictors in the models. We found that meteorological and phenological variables improved the accuracies and percentage of variance explained (R^2) for most of the parameters evaluated. Based on R^2 values, the best prediction models were obtained for biomass (0.71), sucrose (0.65), flavonoids (Flav) (0.56) and nitrogen (0.70) with normalized root mean squared errors of 0.196, 0.32, 0.087 and 0.08, respectively. These parameters were associated mainly with visible (VIS) (approx. 350–700 nm) and near infrared (NIR) (700–1250 nm) regions of the spectrum. Regarding mineral composition, the best prediction models were developed for P (0.51), B (0.50) and Zn (0.44), associated with the short-wave infra-red (SWIR) region (1250–2500 nm). The results of this study demonstrated the potential of hyperspectral techniques to be used as a base for performing initial evaluations in the field of quality traits in alfalfa crops.

1. Introduction

Alfalfa (*Medicago sativa* L.) is the most important forage around the world, due to its strong and broad adaptability to different environmental conditions (Feng et al., 2020). Because of this, alfalfa is called the "queen of forage", considering its crucial role in animal feeding, which then provides dairy and meat products that are important constituents for the human diet (Fan et al., 2018). The forage provides several compounds such as fibres, soluble carbohydrates, proteins, minerals, etc., that directly affect animal performance and the profitability of livestock (Feng et al., 2022). This perennial legume has the potential to produce high yields and withstands multiple harvests in a single year without loss in quality (Shi et al., 2017). Furthermore, planting of alfalfa contributes to soil protection and quality, improving the soil's structure due to its capacity to develop deep root systems (Hrbáčková et al., 2020; Radović et al., 2009). In addition, like other legumes, alfalfa plants have the capacity to fix atmospheric molecular nitrogen, therefore contributing to the natural N fertilization of soils, minimizing the use of synthetic fertilizers. For this reason, forages are becoming an appropriate crop for inclusion in agricultural systems with low inputs that could

Abbreviations: ADF, acid detergent fibre; Anth, anthocyanins; Chl, chlorophyll; CP, crude protein; E.net, elastic net regression model; Flav, flavonoids; GDD, growing degree days; LASSO, least absolute shrinkage and selection operator; NDF, neutral detergent fibre; NIR, near infrared; NIRS, NIR spectroscopy; nRMSE, normalized root-mean square error; PCA, principal component analysis; PLS, partial least square regression; RMSE, root-mean square regression; SWIR, short wave infrared; VIS, visible.

* Corresponding author.
E-mail address: iker.aranjuelo@csic.es (I. Aranjuelo).

https://doi.org/10.1016/j.compag.2023.108463
Received 6 June 2023; Received in revised form 18 November 2023; Accepted 20 November 2023
Available online 12 December 2023

mitigate the emission of greenhouses gases (Stagnari et al., 2017).

For forages, the correct determination of the best time to harvest is a crucial technical management factor that has the greatest effect on production quantity, quality, persistence, and profitability of the crop, and as a consequence has an impact on its digestibility (McDonald et al., 2021). For instance, an equilibrium between structural and non-structural carbohydrates is required to provide sufficient energy while avoiding the development of slowly digested fibre (Chamberlain et al., 2016; Fulgueira et al., 2007). The forage reaches its optimal quality period prior to the appearance of the flower buds, while maximum production is reached during full bloom. Therefore, the choice of the harvest time is based on a complex balance between quality and production (Chamberlain et al., 2016).

The most important quality parameters considered by the alfalfa market are: (i) crude protein (CP) involving all the N content in both protein and non-protein forms; (ii) starch and soluble carbohydrates (glucose, fructose, sucrose) that can be digested quickly and used as energy by rumen microbes (Chamberlain et al., 2016); (iii) mineral nutrients required for animal reproduction, health, regulation of metabolism and immune functions, especially the macronutrients potassium (K), phosphorus (P), calcium (Ca), magnesium (Mg), sulfur (S) and the micronutrients zinc (Zn) and iron (Fe), which are essentials precursors for enzymatic activities to enhance digestibility (Chand et al., 2022) and (iv) fibres (cellulose, hemicellulose, and lignin) that form the structure of the cell wall (Fulgueira et al., 2007). Previous authors have indicated that ruminants need to consume dietary fibre to assist the function of the rumen, although high fibre consumption decreases digestibility of the forage. Indeed, the increase in the fibre content in advance phenological stages of alfalfa plants has been implicated in reduced digestibility in cattle (Chamberlain et al., 2016). Therefore, the characterization of these compounds is a very relevant point because their values can be altered by factors such as phenology (Fan et al., 2018) and environmental conditions (Li et al., 2022).

Within the framework of precision agriculture, producers require cheaper, faster, and more precise methods that inform about the state of the plants and the plant food quality at all times. The most standardized analyses (HPLCs, gas chromatographs, etc.) are lab-based, and the requirement to purchase expensive equipment and, consumables, etc. significantly increases the cost of analysis. Furthermore, it has to be considered that these analyses are very laborious and require a high investment in time and consumables (Feng et al., 2020), which decreases throughput. In the case of alfalfa, some of the key parameters related to fibre composition, such as neutral detergent fibre (NDF) and acid detergent fibre (ADF) analyses, need to be determined using time-consuming methods and hazardous chemicals (Fulgueira et al., 2007). Currently, traditional techniques to analyse such traits have been replaced and/or complemented by lab spectroscopy techniques (Bec et al., 2020). Among these, near infrared spectroscopy (NIRS) equipment would stand out as being the most commonly used. NIRS devices use radiation in the 750 – 2500 nm wavelength region of the spectrum and allow rapid quantification of different traits in several crops (Batten, 1998; Vasseur et al., 2022). This technique was first used in pastures and forage, including alfalfa crops, mainly for quality factors determinations such as protein, fibre, lignin and *in vitro* dry matter (Batten, 1998; Brogna et al., 2009). However, despite the rapid and simultaneous determination possibilities of NIRS devices, those analyses require a prior preparation and standardization in the collection and manipulation of samples, i.e., particle size has a significant influence on the spectral responses, thus all samples should be dry and finely ground (Batten, 1998). By contrast, portable spectroradiometers enable *in situ,* non-destructive determinations of the crop, which means no processing of samples and enables repeated measures on the same plant. Also, the wide range and high resolution of some devices (e.g. 350 nm – 2500 nm) allow the capture of information in the entire range of the spectrum, including the visible region of the spectrum related to marked absorption peak of plant pigments that is absent in NIRS devices (Araus et al., 2021). These advantages have together increased the interest in spectral measurements for further exploitation in agricultural research, and the development of portable tools for *in situ* (i.e. field) assessment of small areas, or even individual plants or canopies (Araus et al., 2021; Grzybowski et al., 2021).

Several studies have demonstrated the efficient use of information generated from full-range or hyperspectral portable spectroradiometers to create prediction models of target plant quality and yield markers (Bruning et al., 2019; Buchaillot et al., 2022; Miguel Garriga et al., 2021; Jackman et al., 2021; Vergara-Diaz et al., 2020; Wiegmann et al., 2019). In alfalfa crops in particular, studies that use hyperspectral techniques have been focused on the development of models to predict yield (Feng et al., 2020; Garriga et al., 2020; Marshall & Thenkabail, 2015; Noland et al., 2018), moisture content (Cevoli et al., 2021), leaf area index, gas exchange, water potential, nitrogen content and isotopic composition (Garriga et al., 2020), and CP and fibre (Feng et al., 2020; R. L. Noland et al., 2018). Other parameters in alfalfa like soluble sugars or minerals elements have not been explored with hyperspectral spectroradiometers under field conditions.

Nowadays, agriculture demands cheaper, faster, non-destructive, and easy-to-use methods that facilitate the most appropriate decisions from the agronomic, economic, and environmental point of view. Therefore, effective technical solutions to determine plant quality traits in-field are highly desirable. In this sense, the use of full-range portable spectroradiometers could allow the characterization of relevant alfalfa parameters in the field. The main objective of this study was to build up prediction models for onsite assessment of quality parameters in commercial an alfalfa crop located in Navarre (Spain), based on the reflectance obtained with a portable full range (350 nm – 2500 nm) spectroradiometer. Spectral measurements were performed in individual leaves and at canopy levels during the entire 2021 campaign.

2. Materials and methods

2.1. Plant material and experimental site

The study was carried out in an alfalfa (*Medicago sativa* L.) cv. Aragon field trial located in Olite (municipality of Navarre region in northern Spain; Fig. 1). The trial was performed during the 2021 growing season. Annual accumulated precipitation was 437.8 mm, and the mean temperature was 13.5 °C (Fig. 2). Patterns of temperature and precipitation were obtained from the meteorological station nearest to the field evaluated (42° 25′N, 1° 39′ W), belonging to the openly accessible station network of the regional government of Navarre (www.meteo.navarra.es/estaciones).

The commercial field selected was sown in September of 2019 and covered a total area of 26 ha, with the samples taken from twelve different sub-plots chosen randomly inside the area (detail information of sampling procedure is described below). The trial was conducted under organic agriculture specifications, and water was supplied according to the plants needs via a pivot irrigation system, with a total of 7.453 m^3 ha^{-1} of water applied during the campaign. The soil pH was 8.21, with optimal levels of N-NO_3 (21 mg kg^{-1}) and medium levels of P (16 mg kg^{-1}) and K (141 mg kg^{-1}) (Appendix Table S1).

2.2. Sampling procedure

In this study we analysed four harvests performed in the field on April 16th, June 7th, July 7th and August 11th and the harvests were designated H1, H2, H3 and H4, respectively (Fig. 2). An additional harvest was performed by field managers on September 29th, but it was not included in this study. The plant characteristics determined at the four harvest dates are a good representation of each time point in growth cycle of alfalfa: H1 being representative of winter and early spring, H2 representing spring, and H3 and H4 covering summer conditions. In addition, four different growth stages were sampled prior of each

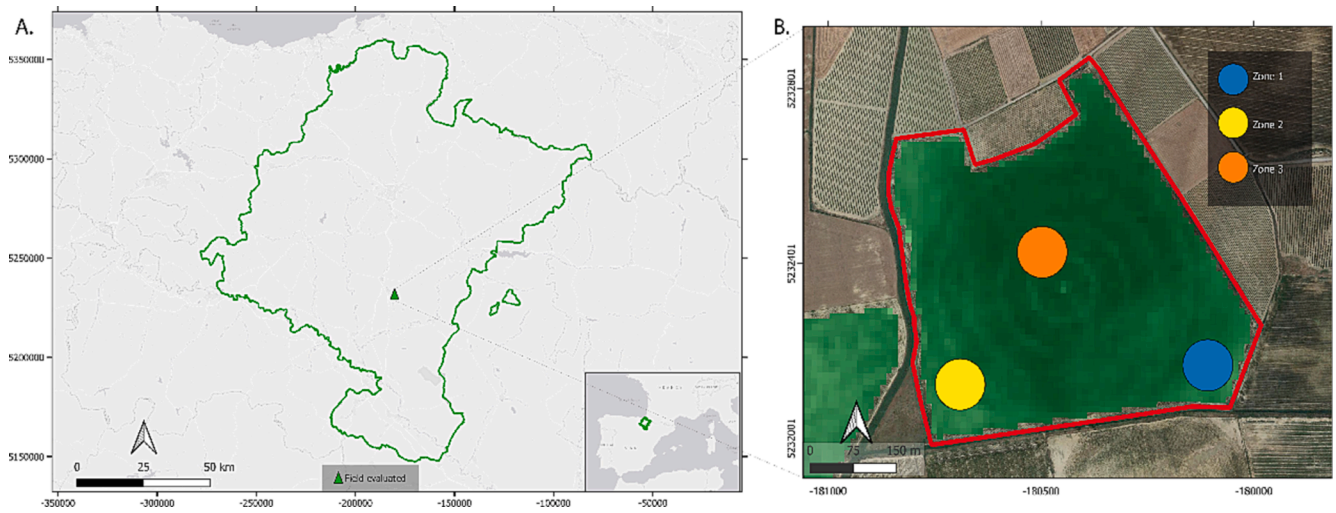

Fig. 1. Geographical location of the field studied in Navarre, Spain. **A.** The geographical location of the field site is marked with a green triangle inside the Navarre region. The lower right inset map shows Navarre's location in Europe. **B.** The right panel shows the field studied and the three zones selected are marked in circles. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

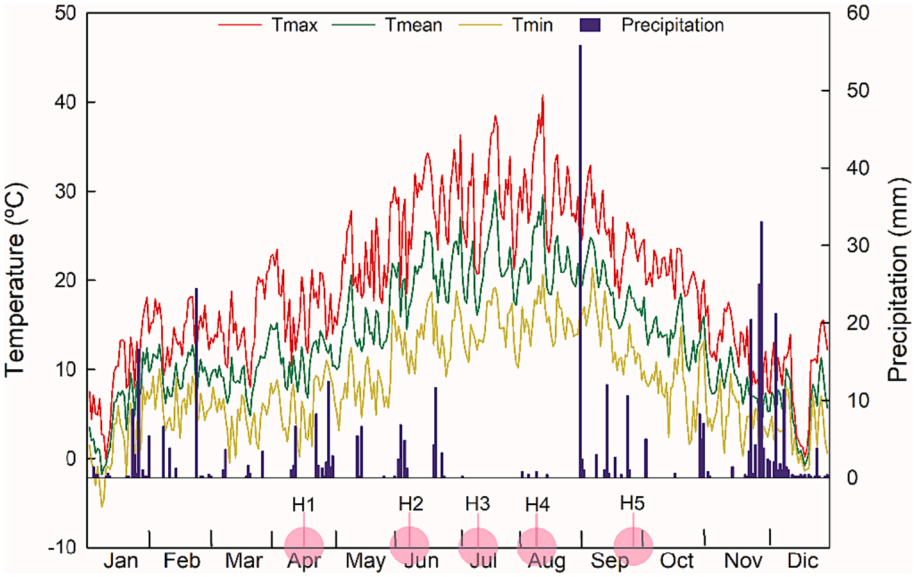

Fig. 2. Monthly temperature (average, minimum and maximum) and precipitation during the 2021 growing season. Harvest dates (H1 to H5) are indicated in the horizontal axis as purple circles. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

harvest, considering the accumulated growing degree days (GDD) calculated according to Charles (1959) (Eq. 1). The dates of sampling were performed with consideration of these GDD and their relationship to phenology, with the dates being indicated in Appendix Table S1.

$$\sum_{Lastharvest}^{Actualharvest} \frac{Tmax + Tmin}{2} - Tbase$$

Where Σ indicates the sum between harvests, depending of the harvest analysed, of daily maximum and minimum temperatures divided by two, minus the base temperature, which in this case was considered 5 °C according to Noland and Wells (2018).

To include the total variability of the field at all sampling dates, three zones were considered within the field (Fig. 1B), and four replicates were taken from each, resulting in sampling of a total of 12 sub-plots. Within each sub-plot, a PVC square of 1 m^2 area was randomly thrown into the field and the biomass inside the square was cut to determine the produced fresh biomass (Appendix Fig. S1). A subsample of the biomass was transferred under refrigeration to the lab, and subsequently processed and stored for later laboratory analyses.

2.3. Spectral field measurements

Prior to harvesting, the total biomass of the random sub-plots, leaf and canopy spectra were measured with a FieldSpec4 full-range portable spectroradiometer (350 – 2500 nm) (ASD Inc. PANanalytical Company, Boulder, USA) (Appendix Fig. A1). In each case, the determinations were conducted, in sunny days, from 11:00 h to 13:00 h. The reflectances of five leaves were recorded for each 1 m^2 sub-plot with a leaf clip accessory coupled to the contact probe of the ASD device, provided with halogen light connected through an optical fibre to the spectroradiometer. Likewise, ten canopy spectra were captured at 45° with respect to the ground and 80 cm above the canopy, with a pistol grip coupled to an optical fibre. The reflectance was calibrated in each sub-plot with a white reference. The averages by sub-plot of the reflectance data were used for analyses.

2.4. Plant biomass production

The alfalfa aboveground biomass of each 1 m^2 subplot was harvested during each harvest date. A subsample of this fresh biomass was kept refrigerated in cold and stored at −80 °C for later laboratory analyses of protein content. The remaining biomass samples collected were placed in an oven at 60 °C for 48 h and the dry biomass was determined. After weighing, the samples were powdered and stored for the later carbohydrate and mineral composition analyses.

2.5. Pigment content

Parallel to spectral determination, chlorophyll *a* + b (Chl), anthocyanins (Anth) and flavonoid (Flav) relative contents were estimated *in vivo*, with a DUALEX ScientificTM a clip sensor (Force-A, Orsay, France), which operates with a red reference at 650 nm and UV light at 375 nm (Cerovic et al., 2012). The data for each sub-plot corresponds to the average of five measurements on different leaves. The measurements were taken on healthy upper leaves of the plants.

2.6. Quality parameter determination in the alfalfa crop

2.6.1. Soluble sugars (glucose, fructose, sucrose), starch, total soluble protein, and N content

The extraction of soluble sugars in the aerial part of alfalfa was carried out in powdered dried material. 25 mg of powder was suspended in 1 ml of ethanol (80 %) in an Eppendorf tube, the samples were incubated and shaken using a thermomixer (90 min, 70 °C, 1100 rpm) and centrifuged at 20800*g* for 10 min at 22 °C. The supernatant was used for soluble sugar quantification using an ionic chromatographer (ICS-3000, Thermo Scientific^{TM}, USA). The pellet obtained was used for starch quantification. Starch was solubilized by adding KOH (0.2 N) to the pellet, and pH was adjusted to 4.8 – 5 with acetic acid (0.1 N). The extraction was carried out with an amyloglucosidase test kit (R-Biopharm AG, Darmstadt, Germany). Finally, the absorbance was measured with a spectroradiometer at 340 nm.

The total soluble protein concentration was determined using the Bradford method (Bradford, 1976) in the samples that were dried, ground and frozen previously. In addition, a total N content determination (organic and inorganic forms) was carried out at the ionomic service of the Centre for Edaphology and Applied Biology (Murcia, Spain) based on Dumas's combustion method using an elemental analyser (TrusSPec CN628, LECO, Michigan, USA) equipped with an autosampler.

2.6.2. Mineral composition

The concentration of a range of minerals (B, Ca, Cu, Fe, K, Mn, Na, P, S and Zn) was analysed on previously dried and ground material. Samples were weighed (≈ 150 mg) in Eppendorf tubes (1.5 ml) and the analyses were carried out at the ionomic service of the Centre for Edaphology and Applied Biology (Murcia, Spain) using ICP/OES (inductively coupled plasma/optical emission spectrometry) on an iCAP 6500 Duo spectrometer (Thermo Fisher Scientific, Waltham, USA). Due to the detection limits and their low concentrations (<0.01 mg kg^{-1}) the elements As, Be, Bi, Cd, Co, La, Pb, Sb, Se, Tl and Ti, were not considered in this study.

2.7. Data analysis

2.7.1. Processing of spectral data

Before using the spectral data for modelling, data processing was performed with SK-UTALCA software (Lobos & Poblete-Echeverría, 2017). Exploratory analysis, spectral noise deletion and spectra outlier exclusion were conducted. Additionally, spectral regions with high levels of noise were removed, especially at the canopy level in wavelengths between 1350 nm and 1465 nm and 1800 nm – 2040 nm, due to their high level of water vapour absorption (Vergara-Diaz et al., 2020). After that, samples with missing values were excluded and then data were normalized using RStudio v4.2.0 (R Core Team, 2020) (R Foundation for Statistical Computing, Vienna, Austria). Finally, principal component analysis (PCA) was applied to reduce the large number of highly correlated wavelengths.

2.7.2. Model building

Biomass and quality parameter prediction models were generated with data from canopy and leaf spectra. To fit the models, single spectra (canopy or leaf) were used and the four phenological stages were combined into one dataset with the objective of increasing the variability of data provided to the model. Considering the distribution of the data in the case of quality parameters, log transformation was applied to some of these traits. Three regression methods were applied for modelling: least absolute shrinkage and selection operator (LASSO), Elastic net (E.net) and Partial least square regression (PLS). PLS is well recognized as an effective method for quantitative correlation of spectral data with reference values (Bec et al., 2020). The ability to reduce the highly correlated predictors to a small number of latent variables allows maximization of the correlations and increases the robustness of the model (Burnett et al., 2021). In addition, LASSO and E.net models apply regularization techniques, which penalizes any coefficients that tend to zero, to avoid overfitting, and effectively identify the subset of significant predictors in high-dimensional data (Pannu & Billor, 2017; Su & Wang, 2021). The data was analysed in Rstudio, principally with the *caret, glmnet* and *pls* packages normally used to run and compare a variety of machine learning models (R Core Team, 2020). As the first step, to increase the robustness of the results, ten-fold cross-validation was conducted, with a total of 100 cross-validation runs being performed. The full dataset was divided randomly into training and validation sets in 70 % and 30 % proportions, respectively. The 70 % data set was used for model training involving parameter optimization using ten-fold cross validation. The 30 % data set, representing the independent test set, was used for assessing the prediction capability of the trained model. The accuracy evaluation of the model was achieved with root-mean-square error (RMSE) and normalized RMSE (nRMSE = RMSE / mean) calculations, and the coefficient of determination (*adj-R2*) was used to estimate the proportion of variance explained by the model. The model showing the greatest robustness and accuracy was chosen for each parameter analysed. In order to increase the robustness of models, the meteorological data (temperature, precipitation and relative humidity) and GDD in each sampling were included into models as predictors.

In addition, selection of the wavelengths in each parameter was directly extracted from the model chosen as a regression coefficient of the respective waveband. For the PLS models, the variables important for projection scores were used as the statistic to select the most important variables. In LASSO and E.net models the variable selection was incorporated into the model-building procedure, with the LASSO method reducing some of the regression coefficients to zero, and non-zero values were selected for use as predictors in the model. E.net is an improvement on the LASSO method using two types of shrinkages of the coefficients. These selected predictors were associated with each parameter evaluated. The bands were aggregated into 50 nm, with the aim of avoiding the highly correlated wavebands and reducing the noise, as described by Hennessy et al. (2020).

2.7.3. Statistical analysis

Statistical analyses were conducted using R software. For all parameters, the normal distribution and homoscedasticity were checked with Shapiro-Wilk and Levene tests, respectively. To study the harvests across the campaign and phenological effects on the biomass yield and quality parameters studied, one -and two- factors analyses of variance (ANOVA) were performed. Differences among factors were assessed using Tukey's HSD test. Significance was accepted at $p < 0.05$ and all figures were created with Sigma-Plot 11.0 software (Systat Software

Inc., California, USA). Furthermore, principal component analysis (PCA) was conducted using R with the *stats* package and the *prcomp* function (R Core Team, 2020).

3. Results

3.1. Yield and quality traits data

The different harvests performed throughout the seasons examined and the phenological stages determined according the GDD reported in Table 1, had a significant effect on the traits analysed (Appendix Table A2 and A3). In this sense, Chl content increased in the last harvest analysed but Flav and Anth decreased significantly. Likewise, glucose, fructose and sucrose showed a decline in H4 compared to H1. The rest of the parameters showed significant changes across the different harvests, although these did not follow a clear trend (Appendix Fig. A1). Regarding the phenological effects, the parameters of biomass, Chl, Flav, glucose, fructose, sucrose, and protein content in leaves increased as a function of advanced phenology. Contrastingly, N content progressively decreased in the last phenological stages. However, starch content reached its highest values in the late vegetative stage but then declined in subsequent stages. In particular, the Anth content decreased progressively up to bud stage and finally increased during flower development (Fig. 3).

In relation to mineral content, significant differences among harvests were observed, but there was no trend among the results (Appendix Fig. A3). Regarding the effect of phenology (Fig. 4), there was a significant decline during the advanced phenological stage in Cu, K, Mn, P, S and Zn, but an increase in Na. The other minerals analysed, such as B, Ca and Fe did not show a clear trend.

Overall, considering the coefficient of variation (CV) a wide variation was observed in traits evaluated in this study, showing a CV higher than 0.20 in most of them, except for leaf pigments (Appendix Table A2), which is important for generating robust models. The traits with highest CV corresponded to the sugars: glucose (CV = 0.66), fructose (CV = 0.53) and starch (CV = 0.43), followed by biomass yield (CV = 0.26), sucrose (CV = 0.25) and protein (CV = 0.25). Finally, the lowest CV was observed for the leaf pigments Chl (CV = 0.08), Flav (CV = 0.12) and Anth (CV = 0.15). In the case of minerals, Fe, B and Na showed the highest CVs with 0.57, 0.49 and 0.30, respectively. The CV for the remaining minerals was in a range between 0.12 and 0.22 (Appendix Table A3).

3.2. Biomass yield and quality trait predictions

The accuracy and R^2 of the prediction models for estimating quality parameters in aboveground material of alfalfa were highly influenced by the GDD (associated with phenology) and the meteorological parameters of temperature, relative humidity, and precipitation during the plants' growth. The results showed that the inclusion of these parameters as predictors in the models enhanced the predictive ability of the traits. In addition, we have developed the models with a dataset that include the four phenological stages for increasing the variability within the parameters. Further each phenological stage was not considered separately because not enough data was available for developing robust models.

Table 1
Growing degree days (GDD) for the four analysed phenological stages in each harvest.

GDD				
Harvest analysed	**Mid vegetative**	**Late vegetative**	**Early bud**	**Early flowering**
1	577	616	665	720
2	371	440	490	518
3	396	424	463	523
4	415	453	505	574

3.2.1. Estimation of biomass yield

The reflectance spectra obtained from canopy and leaves were clearly differentiated by principal component analysis (PCA) (Fig. 5A), although these differences could not be so clearly separated when plotted as individual reflectance values (Fig. 5B).

Canopy and leaf reflectance spectra were used to predict biomass yield (n = 192), and the best prediction was obtained using canopy spectra (Table 2). Canopy spectra were able to explain 71 % of variability in test sets with high accuracy (nRMSE = 0.196), using the PLS regression model. Although, the E.net regression models provided the best prediction accuracy, it was very close to the PLS prediction accuracy. In addition, the R^2 and accuracy was increased up to 8.2 % when GDD and meteorological conditions were included as predictors into the PLS model.

3.2.2. Estimation of quality traits

The quality parameters measured in this study in alfalfa plants were associated with spectra in both canopy and leaf reflectance (Table 3). In the leaf pigments estimation (Chl, Flav, Anth) (n = 191) using a handheld leaf pigment meter (i.e., DUALEX), the best predictions were obtained for Flav and Anth, where the LASSO regression model was able to explain 56 % and 45 % of variability in validation sets, respectively; with high accuracy ($nRMSE_{Flav}$ = 0.087, $nRMSE_{Anth}$ = 0.01). For the estimation of Chl, E.net regression explained 37 % of variability in validation sets with high accuracy (nRMSE = 0.089). Similar to biomass, the GDD and meteorological conditions improved the accuracy of the models by 4 % − 10 %. For Anth in particular, the raw reflectance from the canopy, without the inclusion of meteorological parameters as predictors, were sufficient to generate the best model with low RMSE, while the external parameters did not increase model accuracy.

For most of the sugars, the best models were generated with the inclusion of GDD and meteorological conditions as predictors, besides the spectral reflectance variables. According to the R^2 in the validation sets, the sugars with the best predictions were sucrose (n = *192,* R^2 = 0.65, nRMSE = 0.32) using canopy spectra in the PLS regression model, followed by fructose (n = *187,* R^2 = 0.50, nRMSE = 0.46) using canopy spectra in E.net. The models for the remaining sugars including starch (n = *192*) and glucose (n = 186) were able to explain 23 % and 41 %, although the accuracies of the predictions obtained for these were low ($nRMSE_{starch}$ = 0.56, $nRMSE_{Glucose}$ = 0.53). Regarding soluble protein content (n = *192*), the LASSO regression model was able to explain 54 % of variability in the validation sets with high accuracy (nRMSE = 0.22). In addition, the LASSO regression model explained 70 % of the variability in total N content with very high accuracy (n = *192,* nRMSE = 0.08).

Concerning the estimation of mineral content (n = *192*) in the field, the models were able to explain more than 30 % of the variability in validation sets for seven of the ten minerals assessed (Table 4). The best predictions were obtained for P (R^2 = 0.51, nRMSE = 0.17), B (R^2 = 0.50, nRMSE = 0.38) and Zn (R^2 = 0.44, nRMSE = 0.05). For these three minerals, no strong differences were observed in the statistical parameters for training and test sets (Table 4). The percentage of variance explained by the models for the remaining minerals was lower than 40 %. For mineral content, the GDD and climatic parameters also improved the accuracy of the model between 14 % and 70 %.

3.3. Wavelength selection for the best prediction model for quality parameters

Fig. 6 illustrates the wavebands selected for each quality parameter. We analysed the full spectrum across VIS-NIR-SWIR for all the parameters evaluated. Despite the different methods used for selecting the predictors for each parameter (either canopy or leaf-based

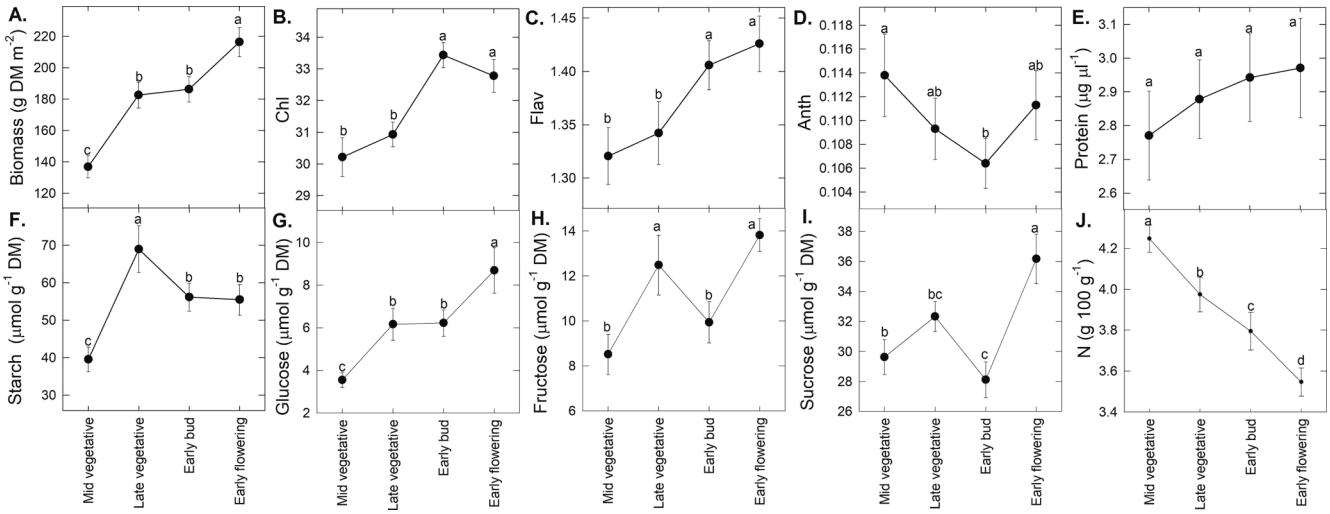

Fig. 3. Biomass, leaf pigments, sugars and N content evaluated throughout different phenological stages of alfalfa. Values represent means ± SE. Different letters indicate significant differences among phenological stages at *p-value* < 0.05.

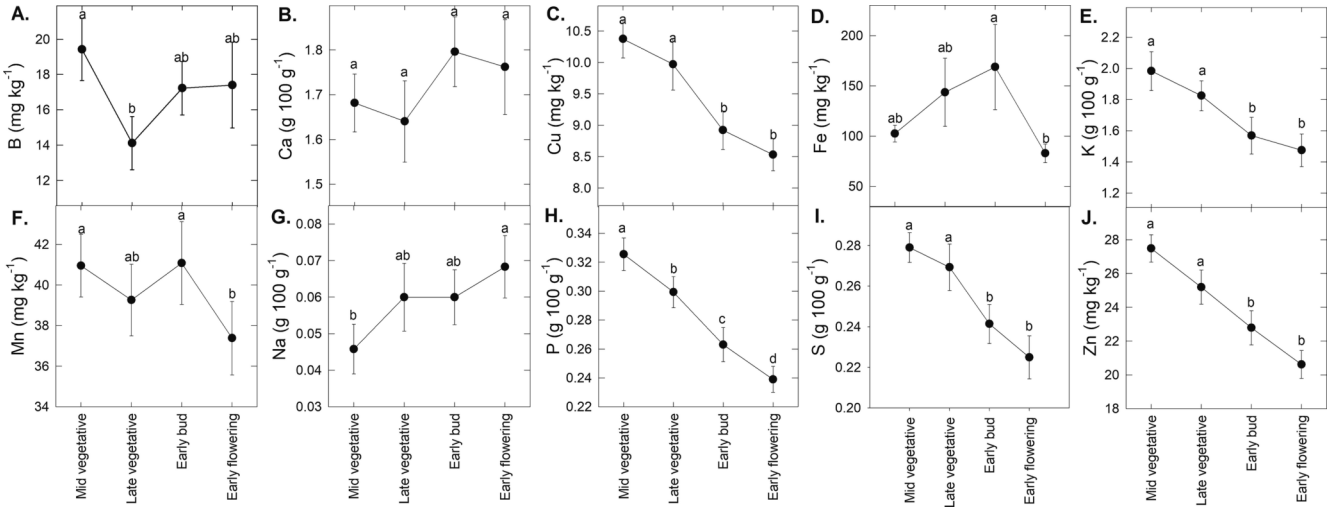

Fig. 4. Mineral contents evaluated throughout different phenological stages of alfalfa. Values represent means ± SE. Different letters indicate significant differences among phenological stages at *p-value* < 0.05.

measurement) common bands were selected in some cases. The best waveband selections for biomass yield prediction occurred in the VIS (350 – 700 nm) and an extensive portion of the NIR (700 nm – 1100 nm) regions of the spectrum. Moreover, when leaf pigments were analysed, the principal bands selected were those that belonged to the VIS region between 350 nm and 700 nm, irrespective of the recording method (canopy or leaf). For the prediction of starch, soluble sugars, protein and nitrogen, the wavebands with the highest impact on prediction were found in the VIS and NIR regions.

A wide range of wavebands was selected in the basis of coefficients obtained for the estimation of mineral composition. Three major regions were identified according to the accumulation of bands: the first was the VIS region, the second in the near SWIR (1250 nm – 1800 nm) and the last corresponded to the far SWIR (2300 nm – 2500 nm). These results show a great variation in the selected bands as a function of the parameters evaluated. In summary, the VIS and NIR regions were mainly selected for biomass, leaf pigments, sugars, soluble protein, and N content, although for minerals the VIS-NIR and a wide portion of the SWIR region of the spectrum were selected.

4. Discussion

The main objective of this research was to estimate yield and quality parameters through generation of statistical models using canopy and leaf reflectance collected with a portable and non-destructive spectroradiometer in an alfalfa crop under field conditions. The evaluation of different phenological stages from harvests undertaken throughout the year, and consequently under different meteorological conditions due to the season, have allowed a dataset to be obtained with a high range of variation and thus better parameter estimation (Table S2, S3). This is the ideal basis for evaluating the potential of spectral data from field conditions to generate quality trait prediction (Vatter et al., 2022) and such an approach benefits the creation of accurate models (Burnett et al., 2021; Hastie et al., 2008). In this context, it is relevant to note that plants are capable of adjusting metabolism to the surrounding environmental conditions and phenological stages to ensure survival. In particular, in alfalfa, the physiological and metabolic modifications involve cold acclimation in autumn, dormancy in winter, and de-acclimation in spring, with evident changes in photoperiod and ambient temperature (Li et al., 2022). Likewise, several authors (Fan et al., 2018; Lamb et al., 2003; McDonald et al., 2021; Sheaffer et al., 2000) have reported the

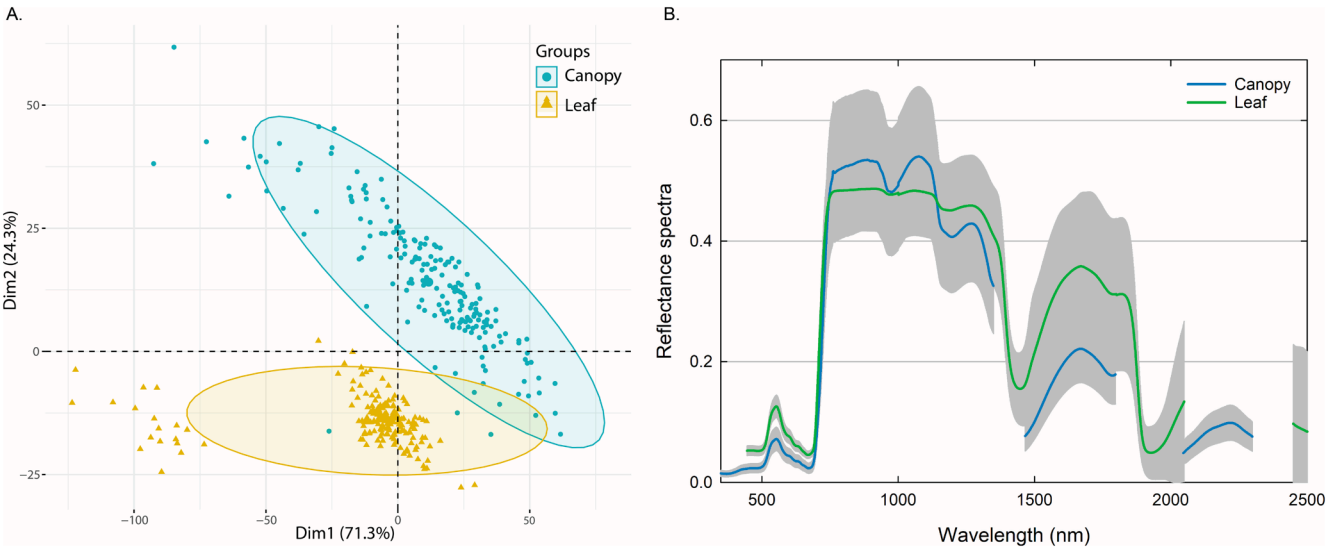

Fig. 5. A. Principal component analysis (PCA) of the reflectance spectra and **B.** Means of canopy and leaf VIS-NIR-SWIR reflectance spectra.

Table 2
Prediction statistics for biomass yield (g DM m^{-2}) of alfalfa for training and test sets evaluated by three regression models (LASSO, E.net and PLS) at the canopy and leaf level.

	Raw spectra as predictors					Raw spectra, GDD and climatic parameters as predictors				
	Training set		Test set			Training set		Test set		
Model	*adj-R^2*	RMSE	*adj-R^2*	RMSE	nRMSE	*adj-R^2*	RMSE	*adj-R^2*	RMSE	nRMSE
	Canopy									
LASSO	0.634	40.05	0.621	42.08	0.233	0.703	34.15	0.675	38.17	0.213
E.net	0.642	40.17	0.615	38.54	0.213	0.712	34.99	0.704	35.03	0.194
PLS	0.656	40.53	0.655	38.79	0.218	0.708	34.37	0.709	35.22	0.196
	Leaf									
LASSO	0.382	49.75	0.389	52.53	0.294	0.536	44.05	0.464	46.25	0.249
E.net	0.382	50.74	0.366	50.88	0.281	0.538	43.81	0.446	47.04	0.254
PLS	0.368	54.94	0.304	52.04	0.281	0.505	46.03	0.456	47.56	0.254

Table 3
Prediction statistics for quality parameters in training and test sets of alfalfa from canopy or leaf spectra. The predictors of phenology and meteorological variables added as predictors next to the reflectance data in the model with higher influence are marked for each parameter.

	Training set		*Test set*			Predictors with higher influence in the model*						
Parameter	*adj-R^2*	RMSE	*adj-R^2*	RMSE	Spectra selected**	GDD	Tmax	Tmin	Tmean	RH	Prec	Model selected***
Biomass	0.708	34.37	0.709	35.22	Canopy	X	X	X	X		X	PLS
Chl	0.390	2.894	0.374	2.854	Canopy	X	X	X		X	X	E.net
Flav	0.559	0.126	0.557	0.121	Leaf	X	X					LASSO
Anth	0.500	0.017	0.449	0.014	Canopy	–	–	–	–	–	–	LASSO
Starch	0.276	28.21	0.227	28.371	Leaf	X	X			X		E.net
Glucose	0.427	4.535	0.410	3.562	Leaf	X	X			X	X	E.net
Fructose	0.487	5.873	0.496	5.390	Canopy		X				X	E.net
Sucrose	0.709	11.558	0.649	12.391	Canopy		X	X			X	PLS
Protein	0.550	0.624	0.541	0.621	Leaf	X	X	X		X		LASSO
Nitrogen	0.691	0.330	0.697	0.309	Leaf				X			LASSO

* GDD: Growing degree days; Tmax: Maximum temperature, Tmin: Minimum temperature; Tmean: Average temperature; RH: Relative humidity; Prec: Precipitation
** The spectra with the best prediction and highest accuracy were selected.
*** The model with the best prediction and highest accuracy was select.

adjustment of plant metabolism as a function of the phenological stage. In this sense, our results have highlighted that phenology significantly altered the parameters analysed. As an example, degradation of starch during advanced growth stages contributed to accumulation of glucose (Fig. 3), as reported by Fan et al. (2018). In relation to mineral content, advanced stages had a negative effect on some of the minerals analysed (Table 4). Similar results were observed by Kume et al. (2001), who reported that P, K, and Mg content decreased in bloom stage of alfalfa and these authors concluded that the most important factor influencing the mineral content in alfalfa is the growth stage (Kume et al., 2001). Given the important role of climatic conditions and phenological factors, as also reported by Noland et al. (2018), these parameters were included as predictors in the current study, resulting in a considerably increased prediction ability in the most of parameters analysed.

Parameter predictions during alfalfa growth period *in vivo* using hyperspectral spectroradiometry has been mainly applied to yield,

Table 4
Prediction statistics for minerals in training and test sets of alfalfa from canopy or leaf spectra. The spectra and models with the best prediction are indicated.

	Training set		Test set				
Mineral	***adj-R²***	**RMSE**	***adj-R²***	**RMSE**	**nRMSE Test**	**Spectra selected**	**Model selected**
B	0.51	7.36	0.50	6.91	0.38	Canopy	Lasso
Ca	0.41	0.34	0.33	0.34	0.20	Leaf	Lasso
Cu	0.36	1.30	0.30	1.33	0.14	Leaf	Lasso
Fe	0.39	0.17	0.33	0.17	0.09	Canopy	Lasso
K	0.29	0.42	0.28	0.41	0.24	Leaf	E.net
Mn	0.34	0.08	0.33	0.08	0.05	Leaf	Lasso
Na	0.21	0.03	0.21	0.03	0.49	Canopy	E.net
P	0.51	0.05	0.51	0.05	0.17	Canopy	Lasso
S	0.28	0.04	0.28	0.04	0.15	Leaf	PLS
Zn	0.54	0.06	0.44	0.07	0.05	Leaf	E.net

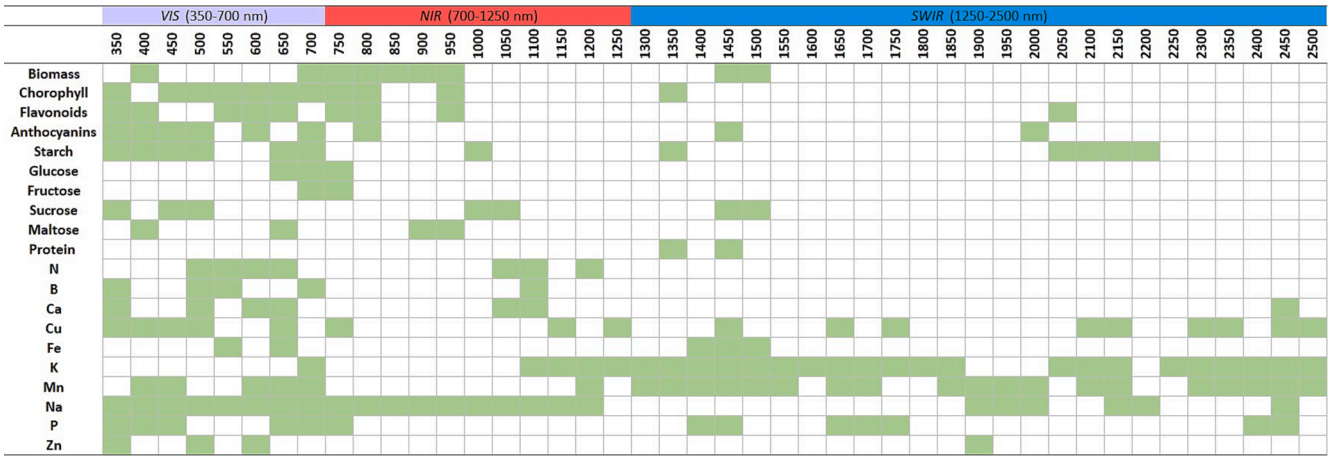

Fig. 6. Wavelength selection at 50 nm intervals for the VIS-NIR-SWIR regions (350 nm – 2500 nm). Green filled cells represent at least one of the wavelengths selected for each parameter evaluated. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

physiological traits (Garriga et al., 2020; Marshall & Thenkabail, 2015), NDF and crude protein content (R. L. Noland et al., 2018). Despite studies of *in vivo* parameters prediction considered in this study, such as leaf pigments, starch, soluble sugars, and mineral content have not been explored for alfalfa crops.

4.1. Biomass yield prediction

The main focus of our study was the prediction of traits based on the entire spectrum from the VIS-NIR-SWIR regions. Yield prediction studies undertaken in alfalfa have commonly focused on the application of vegetation indices calculated with the combination of single bands of spectra (Cevoli et al., 2021; Chandel et al., 2021; Feng et al., 2020). Nevertheless, these method have been limited by using only a few wavebands, and therefore the information available for introduction into machine learning models has been restricted (Garriga et al., 2020; Vatter et al., 2022). Within this context, it should be also considered that indices calculated with NIR bands tend to be saturated at advanced phenological stages as result of a high leaf area index, which could result in poor correlations between traits and indices (Kayad et al., 2016). Thus, we used single wavelengths as inputs into the models evaluated.

Biomass yield prediction obtained in this study was informative, being the percentage of variance explained intermediate between the one obtained by Noland et al. (2018) (who reported $R^2 = 0.79$), and the value achieved in Garriga et al. (2020) ($R^2 = 0.65$). However, previous authors have restricted these estimations to harvest time, while our study performed the estimation at different phenological stages.

4.2. Prediction of leaf pigment contents

The prediction of leaf chlorophyll content through reflectance parameters is well established (Tayade et al., 2022). Different authors have reported considerable prediction for chlorophyll measured by the handheld SPAD and MC-100 devices in wheat and maize crops (Cotrozzi et al., 2020; Ge et al., 2019; Silva-Perez et al., 2018). However, our results in alfalfa showed lower success for relative chlorophyll estimation when the DUALEX device was used ($R^2 = 0.37$). Regarding prediction of other leaf pigments, Flav and Anth have been less studied, and to the best of our knowledge, no previous investigations in alfalfa crops exist. We achieved relatively high accuracies in the models in this study, although only intermediate R^2 values were generated ($R^2 = 0.58$ and 0.45 for Flav and Anth, respectively). Different authors have suggested that the presence of chlorophyll could mask the concentration of other pigments (Blackburn, 2007; Gitelson et al., 2001), which could explain the moderate prediction in Flav and Anth content.

4.3. Predictions of quality trait

On the other hand, we obtained an intermediate accuracy for sucrose ($R^2 = 0.65$) and fructose ($R^2 = 0.50$) predictions. This result could be explained by the rapid metabolization of stored sugars during the day to prevent down regulation of photosynthesis by avoiding the accumulation of solutes in the organs; as a consequence, excess sugars are limited under optimal conditions and their detection *in vivo* become more difficult (Ely et al., 2019; Zeeman et al., 2010). For this reason, Ely et al. (2019) indicated that for many metabolites, developing robust models is a challenge due to regulation by external and internal factors. Nevertheless, similar results to ours were found in maize plants by Yendrek

et al. (2017) for sucrose prediction ($R^2 = 0.62$) and in rice by Das et al. (2018) for total sugars ($R^2 = 0.72$). In addition, Ely et al. (2019) evaluated eight different crops and found a low R^2 for fructose ($R^2 = 0.29$) in comparison to our results, but had a strong prediction for starch ($R^2 = 0.75$).

The prediction of soluble protein content was analysed in this study due to its importance as a principal N source. Although studies that estimate protein *in vivo* are limited, we obtained a prediction of soluble protein with high accuracy, but the R^2 (0.54) was moderate in comparison to estimations in alfalfa of crude protein reported by Noland et al. (2018) ($R^2 = 0.66 – 0.87$) and Feng et al. (2022) ($R^2 = 0.84$). In relation to N, we found similar R^2 (0.70) to Ramoelo et al. (2011) who obtained R^2 values of 0.60, when original reflectance was used in savanna grass *in vivo*.

Moreover, in the current study, we were able to estimate N, Zn, P, B, Ca, Fe and Mn with relatively high accuracies (Table 4). Because some inorganic components do not have specific absorbance, mineral predictions based on spectroscopy techniques are more difficult than organic molecules. For this reason, minerals should be estimated by their association with other compounds that produce a spectral signal with specific vibrational bonds (Horta et al., 2015). Likewise, water's absorption in different regions of the spectrum can mask minerals, especially in fresh material (Ramoelo et al., 2011). However, previous authors obtained low R^2 for P (0.18) in comparison to our results ($R^2 = 0.52$), while Ge et al. (2019) reported a similar R^2 value to ours (0.48) for P in maize. The results of the current study demonstrate the potential of using hyperspectral techniques for an initial screening for prediction of minerals in the field.

4.4. Selection of wavelength for quality parameters

As mentioned before, we used the entire spectrum to develop the models. In this sense, we extracted the most relevant wavebands for each trait (Fig. 6). In this study, selection for biomass yield corresponded mainly to the VIS region and NIR regions, which have previously been associated with photosynthetic processes and water absorption, respectively (Hernandez et al., 2015; Vergara-Diaz et al., 2020). Moreover, as we observed, the reflection in the VIS region was the most associated with leaf pigments (Blackburn, 2007). Furthermore, the NIR region, which includes the red edge slope (680 – 780 nm), is one of the most frequently selected regions in other studies, as detailed in Hennessy et al. (2020). This region has been associated with structural features as well as the nitrogen and sugar contents of leaves (Araus et al., 2021). This is in agreement with our results, where the NIR region was linked to starch and soluble sugars.

In addition, the SWIR region can be related to absorptions by proteins, lignin, cellulose and other biochemical components (Araus et al., 2021; Hennessy et al., 2020; Vergara-Diaz et al., 2020). In our study, mineral predictions were associated with these regions. Concerning mineral composition, most of the bands selected are found in the SWIR region. Our results are similar to those reported by Ramoelo et al. (2011), who analysed N and P in savanna grass and found that many of the important bands were located between 1000 and 2500 nm. In the same sense, Pimstein et al., (2011) found high correlation of the SWIR region with the N and K content in wheat plants.

5. Conclusions

The current manuscript showed that the full range of reflectance spectra recorded *in vivo* from the canopy and leaves can be used as an initial screening to characterize the yield and quality parameters of alfalfa before harvest and may serve as a rapid and non-destructive method to implement management practices in real-time during the crop campaign. Additionally, this study established that the inclusion of environmental and phenological data as predictors can improve such predictions and the accuracy of the models. Nevertheless, we suggest the constant addition of new data from future harvests in the models, with the aim of enhancing the accuracy of parameter predictions. Incorporation of these remote sensing techniques opens the possibilities of monitoring crops at different phenological stages with the advantage of fast evaluation of quality under field conditions and inform efficient decisions in relation to time of harvest.

CRediT authorship contribution statement

Angie L. Gamez: Conceptualization, Methodology, Data curation, Writing – original draft. **Thomas Vatter:** Data curation, Writing – review & editing. **Luis G. Santesteban:** Writing – review & editing. **Jose Luis Araus:** Conceptualization, Writing – review & editing. **Iker Aranjuelo:** Conceptualization, Methodology, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: [Angie Lorena Gamez Guzman reports financial support was provided by Navarra Government].

Data availability

Data will be made available on request.

Acknowledgments

Angie L. Gámez is the recipient of a PhD grant (reference 0011-1408-2020-000005) funded by the Government of Navarre and Nafosa S.L. This manuscript has been conducted within the context of the CropEqualT-CEC project funded by the European Union's Horizon 2020, Belgium Marie Curie Rise research and innovation programme. Garazi Ezpeleta for technical support provided during the sample analyses in the lab and José María Oses and Bernardo Monreal for the being in charge of crop management.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compag.2023.108463.

References

Araus, J.L., Kefauver, S.C., Díaz, O.V., Gracia-Romero, A., Rezzouk, F.Z., Segarra, J., Buchaillot, M.L., Chang-Espino, M., Vatter, T., Sanchez-Bragado, R., Gallego, J.A.F., Serret, M.D., Bort, J., 2021. Crop phenotyping in a context of global change: What to measure and how to do it. J. Integr. Plant Biol. 64 (2), 592–618. https://doi.org/10.1111/jipb.13191.

Batten, G.D., 1998. Plant analysis using near infrared reflectance spectroscopy: the potential and the limitations. Australian Journal of Exp. Agric. 38, 697–706.

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 (August), 1–27. https://doi.org/10.3389/fpls.2020.01226.

Blackburn, G.A., 2007. Hyperspectral remote sensing of plant pigments. J. Exp. Bot. 58 (4), 855–867. https://doi.org/10.1093/jxb/erl123.

Bradford, M.M., 1976. A rapid and sensitive method for the quantitation microgram quantities of protein utilizing the principle of Protein-Dye binding. Anal. Biochem. 254, 248–254.

Brogna, N., Pacchioli, M.T., Immovilli, A., Ruozzi, F., Ward, R., Formigoni, A., 2009. The use of near-infrared reflectance spectroscopy (NIRS) in the prediction of chemical composition and in vitro neutral detergent fiber (NDF) digestibility of Italian alfalfa hay. Ital. J. Anim. Sci. 8 (SUPPL. 2), 271–273. https://doi.org/10.4081/ijas.2009.s2.271.

Bruning, B., Liu, H., Brien, C., Berger, B., Lewis, M., Garnett, T., 2019. The development of hyperspectral distribution maps to predict the content and distribution of Nitrogen and water in wheat (Triticum aestivum). Front. Plant Sci. 10, 1–16. https://doi.org/10.3389/fpls.2019.01380.

Buchaillot, M.L., Soba, D., Shu, T., Liu, J., Aranjuelo, I., Araus, J.L., Runion, G.B., Prior, S.A., Kefauver, S.C., Sanz-Saez, A., 2022. Estimating peanut and soybean

photosynthetic traits using leaf spectral reflectance and advance regression models. Planta 255 (4), 1–19. https://doi.org/10.1007/s00425-022-03867-6.
Burnett, A.C., Anderson, J., Davidson, K.J., Ely, K.S., Lamour, J., Li, Q., Morrison, B.D., Yang, D., Rogers, A., Serbin, S.P., 2021. A best-practice guide to predicting plant traits from leaf-level hyperspectral data using partial least squares regression. J. Exp. Bot. 72 (18), 6175–6189. https://doi.org/10.1093/jxb/erab295.
Cerovic, Z.G., Masdoumier, G., Ghozlen, N., Latouche, G., 2012. A new optical leaf-clip meter for simultaneous non-destructive assessment of leaf chlorophyll and epidermal flavonoids. Physiol. Plantarum 146, 251–260. https://doi.org/10.1111/j.1399-3054.2012.01639.x.
Cevoli, C., Di Cecilia, L., Ferrari, L., Fabbri, A., & Molari, G. (2021). Potential of in-field Vis/NIR hyperspectral imaging to monitor quality parameters of alfalfa. *2021 IEEE International Workshop on Metrology for Agriculture and Forestry, MetroAgriFor 2021 - Proceedings*, 341–345. 10.1109/MetroAgriFor52389.2021.9628816.
Chamberlain, Lindsay, Ketterings, Q., Lyons, S., Cerosaletti, P., Czymmek, K., Cherney, D., & Kilcer, T. (2016). *Forage Quality Parameters Explained Agronomy Fact Sheet 94*. http://ccedelaware.org/wp-.
Chamberlain, L., Ketterings, Q., Lyons, S., Cerosaletti, P., Czymmek, K., Cherney, D., Kilcer, T., 2016. Forage quality parameters explained. Agronomy Fact Sheet Series. http://ccedelaware.org/wp-.
Chandel, A.K., Khot, L.R., Yu, L.X., 2021. Alfalfa (Medicago sativa L.) crop vigor and yield characterization using high-resolution aerial multispectral and thermal infrared imaging technique. Comput. Electron Agric. 182 (June 2020), 105999 https://doi.org/10.1016/j.compag.2021.105999.
Chand, S., Indu, Singhal, R.K., Govindasamy, P., 2022. Agronomical and breeding approaches to improve the nutritional status of forage crops for better livestock productivity. Grass Forage Sci. 77 (1), 11–32. https://doi.org/10.1111/gfs.12557.
Charles, A., 1959. The determination and significance of the base temperature in a linear heat unit system. Proc. Am. Soc. Hortic. Sci. 74 (3).
Cotrozzi, L., Peron, R., Tuinstra, M.R., Mickelbart, M.V., Couture, J.J., 2020. Spectral phenotyping of physiological and anatomical leaf traits related with maize water status. Plant Physiol. 184 (3), 1363–1377. https://doi.org/10.1104/pp.20.00577.
Das, B., Sahoo, R.N., Pargal, S., Krishna, G., Verma, R., Chinnusamy, V., Sehgal, V.K., Gupta, V.K., Dash, S.K., Swain, P., 2018. Quantitative monitoring of sucrose, reducing sugar and total sugar dynamics for phenotyping of water-deficit stress tolerance in rice through spectroscopy and chemometrics. Spectrochimica Acta - Part a: Molecular and Biomolecular Spectroscopy 192, 41–51. https://doi.org/10.1016/j.saa.2017.10.076.
Ely, K.S., Burnett, A.C., Lieberman-Cribbin, W., Serbin, S.P., Rogers, A., 2019. Spectroscopy can predict key leaf traits associated with source-sink balance and carbon-nitrogen status. J. Exp. Bot. 70 (6), 1789–1799. https://doi.org/10.1093/jxb/erz061.
Fan, W., Ge, G., Liu, Y., Wang, W., Liu, L., Jia, Y., 2018. Proteomics integrated with metabolomics: Analysis of the internal causes of nutrient changes in alfalfa at different growth stages. BMC Plant Biol. 18 (1), 1–15. https://doi.org/10.1186/s12870-018-1291-8.
Feng, L., Zhang, Z., Ma, Y., Du, Q., Williams, P., Drewry, J., Luck, B., 2020. Alfalfa yield prediction using UAV-based hyperspectral imagery and ensemble learning. Remote Sens. 12, 1–24. https://doi.org/10.3390/rs12122028.
Feng, L., Zhang, Z., Ma, Y., Sun, Y., Du, Q., Williams, P., Drewry, J., Luck, B., 2022. Multitask learning of alfalfa nutritive value from UAV-Based hyperspectral images. IEEE Geosci. Remote Sens. Lett. 19 https://doi.org/10.1109/LGRS.2021.3079317.
Fulgueira, C.L., Amigot, S.L., Gaggiotti, M., Romero, L.A., Basílico, J.C., 2007. Forage Quality : Techniques for testing. Fresh Produce 1 (2), 121–131.
Garriga, M., Ovalle, C., Espinoza, S., Lobos, G.A., Del Pozo, A., 2020. Use of Vis-NIR reflectance data and regression models to estimate physiological and productivity traits in lucerne (Medicago sativa). Crop Pasture Sci. 71 (1), 90–100. https://doi.org/10.1071/CP19182.
Garriga, M., Romero-Bravo, S., Estrada, F., Méndez-Espinoza, A.M., González-Martínez, L., Matus, I.A., Castillo, D., Lobos, G.A., Del Pozo, A., 2021. Estimating carbon isotope discrimination and grain yield of bread wheat grown under water-limited and full irrigation conditions by hyperspectral canopy reflectance and multilinear regression analysis. International J. Remote Sens. 42 (8), 2848–2871. https://doi.org/10.1080/01431161.2020.1854888.
Ge, Y., Atefi, A., Zhang, H., Miao, C., Ramamurthy, R.K., Sigmon, B., Yang, J., Schnable, J.C., 2019. High-throughput analysis of leaf physiological and chemical traits with VIS-NIR-SWIR spectroscopy: A case study with a maize diversity panel. Plant Methods 15 (1), 1–12. https://doi.org/10.1186/s13007-019-0450-8.
Gitelson, A.A., Merzlyak, M.N., Chivkunova, O.B., 2001. Optical properties and nondestructive estimation of anthocyanin content in plant leaves. Photochem. Photobiol. 74 (1), 38. https://doi.org/10.1562/0031-8655(2001)074<0038:opaneo>2.0.co;2.
Grzybowski, M., Wijewardane, N.K., Atefi, A., Ge, Y., Schnable, J.C., 2021. Hyperspectral reflectance-based phenotyping for quantitative genetics in crops: Progress and challenges. Plant Commun. 2 (4), 1–11. https://doi.org/10.1016/j.xplc.2021.100209.
Hastie, T., Tibshirani, R., Friedman, J., 2008. The elements of statistical learning: Data mining, inference and prediction (Second edi). Springer.
Hennessy, A., Clarke, K., Lewis, M., 2020. Hyperspectral classification of plants: A review of waveband selection generalisability. Remote Sens. 12 (1) https://doi.org/10.3390/RS12010113.
Hernandez, J., Lobos, G.A., Matus, I., del Pozo, A., Silva, P., Galleguillos, M., 2015. Using ridge regression models to estimate grain yield from field spectral data in bread wheat (Triticum Aestivum L.) grown under three water regimes. Remote Sens. 7 (2), 2109–2126. https://doi.org/10.3390/rs70202109.
Horta, A., Malone, B., Stockmann, U., Minasny, B., Bishop, T.F.A., McBratney, A.B., Pallasser, R., Pozza, L., 2015. Potential of integrated field spectroscopy and spatial analysis for enhanced assessment of soil contamination: A prospective review. Geoderma 241–242, 180–209. https://doi.org/10.1016/j.geoderma.2014.11.024.
Hrbáčková, M., Dvořák, P., Takáč, T., Tichá, M., Luptovčiak, I., Šamajová, O., Ovečka, M., Šamaj, J., 2020. Biotechnological perspectives of omics and genetic engineering methods in Alfalfa. Front. Plant Sci. 11 (May) https://doi.org/10.3389/fpls.2020.00592.
Jackman, P., Lee, T., French, M., Sasikumar, J., O'Byrne, P., Berry, D., Lacey, A., Ross, R., 2021. Predicting key grassland characteristics from hyperspectral data. AgriEngineering 3 (2), 313–322. https://doi.org/10.3390/agriengineering3020021.
Kayad, A.G., Al-Gaadi, K.A., Tola, E., Madugundu, R., Zeyada, A.M., Kalaitzidis, C., 2016. Assessing the spatial variability of alfalfa yield using satellite imagery and ground-based data. PLoS One 11 (6). https://doi.org/10.1371/journal.pone.0157166.
Kume, S., Toharmat, T., Nonaka, K., Oshita, T., 2001. Relationships between crude protein and mineral concentrations in alfalfa and value of alfalfa silage as a mineral source for periparturient cows. Animal Feed Sci. Tech. 93, 157–168.
Lamb, J.F.S., Sheaffer, C.C., Samac, D., 2003. Population density and harvest maturity effects on leaf and stem yield in alfalfa. Agron. J. 95, 635–641.
Li, Z., He, F., Tong, Z., Li, X., Yang, Q., 2022. Metabolomic changes in crown of alfalfa (Medicago sativa L.) during de - acclimation. Sci Rep 12 (14977), 1–13. https://doi.org/10.1038/s41598-022-19388-x.
Lobos, G.A., Poblete-Echeverría, C., 2017. Spectral knowledge (SK-UTALCA): Software for exploratory analysis of high-resolution spectral reflectance data on plant breeding. Front. Plant Sci. 7, 1–15. https://doi.org/10.3389/fpls.2016.01996.
Marshall, M., Thenkabail, P., 2015. Developing in situ non-destructive estimates of crop biomass to address issues of scale in remote sensing. Remote Sens. 7 (1), 808–835. https://doi.org/10.3390/rs70100808.
Mcdonald, I., Min, D., Baral, R., 2021. Effect of a fall cut on dry matter yield, nutritive value, and stand persistence of alfalfa. J Anim Sci Technol 63 (4), 799–814.
Noland, R., & Wells, S. (2018). *Using growing degrees days to plan early-season alfalfa harvests*. https://extension.umn.edu/forage-harvest-and-storage/using-growing-degree-days-plan-early-season-alfalfa-harvests#sources-1049160.
Noland, R.L., Wells, M.S., Coulter, J.A., Tiede, T., Baker, J.M., Martinson, K.L., Sheaffer, C.C., 2018. Estimating alfalfa yield and nutritive value using remote sensing and air temperature. Field Crop Res 222 (August 2017), 189–196. https://doi.org/10.1016/j.fcr.2018.01.017.
Pannu, J., Billor, N., 2017. Robust group-Lasso for functional regression model. Communications in Statistics—Simulation and Computation 46 (5), 3356–3374. https://doi.org/10.1080/03610918.2015.1096375.
Pimstein, A., Karnieli, A., Bansal, S.K., Bonfil, D.J., 2011. Exploring remotely sensed technologies for monitoring wheat potassium and phosphorus using field spectroscopy. Field Crop Res 121 (1), 125–135. https://doi.org/10.1016/j.fcr.2010.12.001.
Radović, J., Sokolović, D., Marković, J., 2009. Alfalfa-most important perennial forage legume in animal husbandry. Biotechnol. Animal Husbandry 25 (5–6), 465–475. https://doi.org/10.2298/bah0906465r.
Ramoelo, A., Skidmore, A.K., Schlerf, M., Mathieu, R., Heitkönig, I.M.A., 2011. Water-removed spectra increase the retrieval accuracy when estimating savanna grass nitrogen and phosphorus concentrations. ISPRS J. Photogramm. Remote Sens. 66 (4), 408–417. https://doi.org/10.1016/j.isprsjprs.2011.01.008.
Sheaffer, C.C., Martin, N.P., Lamb, J.F.S., Cuomo, G.R., Jewett, J.G., Quering, S.R., 2000. Leaf and stem properties of alfalfa entries. Agron. J. 92, 733–739.
Shi, S., Nan, L., Smith, K.F., 2017. The current status, problems, and prospects of alfalfa (Medicago Sativa L.) breeding in China. Agronomy 7 (1), 1–11. https://doi.org/10.3390/agronomy7010001.
Silva-Perez, V., Molero, G., Serbin, S.P., Condon, A.G., Reynolds, M.P., Furbank, R.T., Evans, J.R., 2018. Hyperspectral reflectance as a tool to measure biochemical and physiological traits in wheat. J. Exp. Bot. 69 (3), 483–496. https://doi.org/10.1093/jxb/erx421.
Stagnari, F., Maggio, A., Galieni, A., Pisante, M., 2017. Multiple benefits of legumes for agriculture sustainability: An overview. Chem. Bio. Techn. Agri. 4 (1), 1–13. https://doi.org/10.1186/s40538-016-0085-1.
Su, M., Wang, W., 2021. Elastic net penalized quantile regression model. J. Computational and Applied Mathematics 392. https://doi.org/10.1016/j.cam.2021.113462.
Tayade, R., Yoon, J., Lay, L., Khan, A.L., Yoon, Y., Kim, Y., 2022. Utilization of spectral indices for high-throughput phenotyping. Plants 11 (13). https://doi.org/10.3390/plants11131712.
Team, R. C. (2020). *R: a language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.r-project.org/.
Vasseur, F., Cornet, D., Beurier, G., Messier, J., Rouan, L., Bresson, J., Ecarnot, M., Stahl, M., Heumos, S., Gérard, 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. Frontiers Plant Science 13 (May). https://doi.org/10.3389/fpls.2022.836488.
Vatter, T., Gracia-Romero, A., Kefauver, S.C., Nieto-Taladriz, M.T., Aparicio, N., Araus, J. L., 2022. Preharvest phenotypic prediction of grain quality and yield of durum wheat using multispectral imaging. Plant J. 109 (6), 1507–1518. https://doi.org/10.1111/tpj.15648.
Vergara-Diaz, O., Vatter, T., Kefauver, S.C., Obata, T., Fernie, A.R., Araus, J.L., 2020. Assessing durum wheat ear and leaf metabolomes in the field through hyperspectral data. Plant J. 102 (3), 615–630. https://doi.org/10.1111/tpj.14636.

Wiegmann, M., Backhaus, A., Seiffert, U., Thomas, W.T.B., Flavell, A.J., Pillen, K., Maurer, A., 2019. Optimizing the procedure of grain nutrient predictions in barley via hyperspectral imaging. PLoS One 14 (11), 1–22. https://doi.org/10.1371/journal.pone.0224491.

Yendrek, C.R., Tomaz, T., Montes, C.M., Cao, Y., Morse, A.M., Brown, P.J., McIntyre, L. M., Leakey, A.D.B., Ainsworth, E.A., 2017. High-throughput phenotyping of maize leaf physiological and biochemical traits using hyperspectral reflectance. Plant Physiol. 173 (1), 614–626. https://doi.org/10.1104/pp.16.01447.

Zeeman, S.C., Kossmann, J., Smith, A.M., 2010. Starch: Its metabolism, evolution, and biotechnological modification in plants. Annu. Rev. Plant Biol. 61, 209–234. https://doi.org/10.1146/annurev-arplant-042809-112301.