Evaluation of cut alfalfa moisture content and operative conditions by hyperspectral imaging combined with chemometric tools: In-field application

Evaluation of cut alfalfa moisture content and operative conditions by hyperspectral imaging combined with chemometric tools: in-field application

Chiara Cevoli[a], Luca Di Cecilia[b], Luca Ferrari[b], Angelo Fabbri[a], Giovanni Molari[a]

[a] *Department of Agricultural and Food Sciences Alma Mater Studiorum, University of Bologna, Italy.*

[b] *CNH Industrial Italia SpA - Innovation Sensing & Controls Engineer Modena, Italy*

corresponding author: chiara.cevoli3@unibo.it

Abstract

High quality of hay or silage can be obtained by monitoring the moisture level at mowing or harvest time to reduce the losses of yield and nutrients and preserve feeds in the long term. Mechanical conditioning is one of the most common ways to increase the rate of water loss from forage during drying. Near Infrared (NIR) spectroscopy is a good alternative to the classical gravimetric method to timely provide information on forage moisture content. The aim of this work was to evaluate the potential of in-field Vis/NIR hyperspectral imaging combined with chemometric tools to monitor alfalfa parameters after conditioning. Partial Least-Squares Discriminant Analysis (PLSDA) models were developed to discriminate samples according to several operative conditions (level of conditioning, field type, time after conditioning, and time of day) and yielded good discrimination power (mean sensitivity=87%, mean probability=89%). Moisture content was estimated by Partial Least-Squares (PLS) models obtaining determination coefficient (R^2)=0.86 and Root Mean Square Error (RMSE)=2.00% (external validation). The number of spectral bands was reduced by using the Variable Importance in Projection (VIP) method, passing from 272 to 74 bands. Further PLS models were built considering the reduced variable numbers and achieved comparable results to those

obtained with the full spectra, demonstrating that reduction of the number of variables did not affect the goodness of the models. Finally, the best PLS model was applied to each pixel of the hyperspectral images to obtain false colour images. Pixels having similar colours were characterized by comparable moisture content. The models developed may be useful to determine moisture content of alfalfa remotely and in real-time during mowing or harvest procedures.

Keywords: forage, hyperspectral imaging, partial least square regression, moisture, harvest, field.

1. Introduction

Alfalfa is a valuable and intensively produced perennial forage crop and plays an important role in the food supply chain as feedstock for animals. In 2020, alfalfa mixtures were harvested from 6.6 million hectares corresponding to about 48.1 million metric tons of yield (Feng et al., 2021). In general, forage is harvested as hay or silage. In the first case, the product is sun-dried in the field after cutting to reach a moisture level less than 12-15%. For silage, a shorter drying time is needed until a moisture content of about 50% is reached (Feng, Zhang, Ma, Du, & Williams, 2020; Lim, Watanabe, Yoshitoshi, & Kawamura, 2020). A high quality final product can be obtained by monitoring the moisture level, as this can reduce the losses of yield and nutrients and preserve feeds in the long term (Lim et al., 2020).

Together with swath manipulation practices, mechanical conditioning at cutting is the most common procedure to increase the rate of water loss from forage during drying. Mechanical conditioning, such as flail or roll conditioners, increases the drying rate by physical rupture of the plant tissue that allows the slower-drying stem to dry at a rate approaching that of faster-drying leaves (Borreani, Tabacco, & Ciotti, 1999). Especially for alfalfa, roll conditioning greatly reduces the dry time without causing high losses of dry matter. Furthermore, the level of conditioning, in terms of distance between rolls, can affect both dry time and leaf loss in the field.

The standard method to measure forage moisture content is based on the oven gravimetric method (103°C for 24 hours). This procedure is not a real-time technique and is clearly incompatible with field operations (Lim et al., 2020). Alternatively, several methods and instruments have been proposed to rapidly monitor forage moisture content in real time and in a non-destructive manner. In particular, considering water sensitivity in the infrared region, many spectroscopy techniques operating in the short-wave infrared (SWIR, from 1300 to 2500 nm) and near infrared (NIR, from 750 to 1300 nm) ranges have been developed to determine water content in vegetation (Feret et al., 2008). Although the strongest relationship with the plant water content (leaf level) is observed in the SWIR region, wavelengths in the NIR region, characterized by weak absorption of water, show better predictive performance (Serrano, Ustin, Roberts, Gamon, & Pen, 2000; Sims & Gamon, 2003). This is probably due to the water vapor in the atmosphere, between the instrument and the forage, that leads to noise in the spectral response, particularly for water-sensitive wavelengths (Colombo, Busetto, Meroni, Rossini, & Panigada, 2012). Most of these spectroscopic techniques have applied to evaluate the plant drought stress, and only a few applications have investigated moisture content determination in cut forage to make silage or hay. Digman & Shinners, 2008 proposed a real-time method to measure forage (alfalfa and corn silage) moisture content using NIR reflectance spectroscopy. The system was mounted on a forage harvest spout and worked in the spectral region from 950 to 1,530 nm. Partial least square (PLS) models were developed to predict moisture content, obtaining a determination coefficient (R^2) of 0.83 and a root mean square error (RMSE) of 3.7 % (cross-validation) for cut alfalfa. Recently, the precision and accuracy of this system were evaluated by other authors (Long, Ketterings, Russell, Vermeylen, & Degloria, 2016).

Hyperspectral imaging (HIS) combines spectroscopy with digital imaging, allowing spatial and spectral information to be obtained simultaneously (Manley, 2014). Due to the rapid development of computer systems with high data processing capability and miniaturization of HSI systems, the opportunity for real-time analysis directly in field has become more interesting. Hyperspectral images can be acquired remotely by installing the system on airborne or unmanned aerial vehicles (UAV) or

on ground vehicles which produce proximal images with high spatial resolution (Benelli, Cevoli, & Fabbri, 2020). Most applications of HSI to estimate forage yield and quality parameters used UAV systems (Feng et al., 2020, 2021; Geipel, Bakken, Jørgensen, & Korsaeth, 2021). Very few applications have utilized ground-based systems, especially to monitor the composition of cut forage. Suzuki et al., (2008) mounted a hyperspectral camera (360-1010 nm) on the roof of a vehicle to scan the mowed grass immediately after the cutting. Multiple linear regression models based on the mean spectra of each image, were developed to estimate grass digestible nutrient (R^2=0.74) and crude protein (R^2=0.76).

In general, studies applying multivariate approaches to elaborate hyperspectral data in the forage sector are scarce. The conventional approach is based on the use of simple vegetation indices using only a few wavelengths (Geipel et al., 2021). This procedure is certainly suitable for reducing the computation time required to real-time data elaboration, although there is the risk that significant information is left out by using a very small number of wavelengths, (Stein I Øvergaard, Isaksson, & Korsaeth, 2013; Stein Ivar Øvergaard, Isaksson, & Korsaeth, 2010). Capolupo, Kooistra, Berendonk, & Boccia, (2015) compared different statistical approaches to elaborate data from UAV-acquired HSI and concluded that the multivariate method (PLS) was superior to methods based on vegetation indices.

The aim of this study was to evaluate the potential of in-field (ground-based) Vis/NIR HSI combined with chemometric techniques (PLS models) to predict the moisture content of cut alfalfa and to discriminate samples according to several operative parameters (PLS discriminant analysis, PLSDA). The hyperspectral acquisitions were made considering several combinations of conditioning level, cuts, time of day, and time after the conditioning.

2. Materials and methods

Tests were done in two different alfalfa fields located in Ozzano (Italy) during summer 2020, which had an area of about 40 x 80 meters. Three forage cuts (1=June, 2=August, and 3=September) for

each field, corresponding to different growth periods, were carried out. Air temperature and relative humidity were measured before each acquisition session (TESTO 605i, Testo SE & Co. KgaA). Each field was split into two parts: the first was cut and conditioned in the morning, and the second in the afternoon. Two different levels of conditionings (distance between the rollers) (FENDT 3160) were considered: close (C) and wide (W). C and W were the minimum and maximum distances between the rollers allowed by conditioner. A swath was conditioned and immediately after it was scanned by the camera in three different zones (about one meter) along its length. After 15 minutes and 2 hours from the conditioning, the same swath was scanned again in the same spots. The above procedure was repeated for all conditioning levels. Several combinations of operative conditions, including, different fields, time of day, condition level, time after conditioning and cuts, allow to cover a large variability in terms of both spectral information and moisture content, which is advantageous for developing predictive models.

2.1 Hyperspectral acquisitions

A hyperspectral camera (HSC) (Nano-Hyperspec VNIR, Headwall Photonics, Inc., Fitchburg, MA, USA) working in the Vis/NIR range (400–1000 nm) was installed on the front loader of a New Holland 4050 Tractor (Figure 1). The optical axis was perpendicular to the ground (scanned surface) analysed. The camera was directly powered by the tractor battery through a DC to AC inverter, and data were sent to a laptop by a WiFi access point connected to the camera and fixed on the loader. Camera-swath distance was about 80 cm. During acquisition, the tractor moved with a constant speed of 2 km h^{-1} and each scan lasted for about 5 seconds. The exposure time, depending on the light intensity, was achieved through calibration by framing a white high-reflectance panel placed at the same distance as the swath. The calibrated reflectance (Rc) was extracted from images considering the raw (Rr), white (Rw), and dark (Rd) reflectance, respectively:

$$Rc = \frac{Rr - Rd}{Rw - Rd}$$

128 **2.2 Moisture content measurement**

129 Immediately after each image acquisition, about 20 g of alfalfa were harvested (in the same scanned
130 area) and directly weighed in the field by using the scale KB 3600-2N (Kern, Italy). Considering the
131 small sample amount, to obtain samples as homogeneous as possible, only the leaves were collected.
132 Moisture content was determined by drying the samples in an oven at 103°C until constant weight
133 (ASABE standard S358.1-2012 Moisture Measurement-Forage).

134

135 **2.3 Data elaboration**

136 The mean spectrum was calculated for each image for a total of 216 mean spectra. These spectra were
137 used for chemometric elaboration. The first step was spectra pre-treatment by the standard normal
138 variate (SNV), first derivative (D1) and mean centring (MC). The SNV is one of the most common
139 pre-processing methods used to correct spectra for changes in optical path length and light scattering,
140 while the derivatives have the capability to remove both additive and multiplicative effects in the
141 spectra (Rinnan, van den Berg, & Engelsen, 2009).
142 PLS discriminant analysis (PLSDA) was used to discriminate samples according to: I) conditioning
143 level (two groups: C and W); II) field type (two groups: field-1 and field-2); III) time of day (two
144 groups: morning, afternoon); and IV) time after conditioning (three groups: 0, 15, and 120 min).
145 PLSDA models were developed for explorative purpose to evaluate the ability to discriminate the
146 samples as a function to operative conditions only on the base of the spectral signatures, regardless
147 the moisture content. From an operative point of view, the models could be used in field during
148 control and traceability processes to trace it back to operating, and timing conditions concerning the
149 cut/conditioning.
150 Before developing the models, the sample data set (216 mean spectra) was split into calibration/cross
151 validation set (80% of samples) and test set (20% of samples) using the Kennard-Stone method
152 (selects samples that best span the same range as the original data, but with an even distribution of
153 samples across the same range) (Daszykowski, Walczak, & Massart, 2002).

PLSDA assumes that each sample is assigned to a single class using a probabilistic approach based on Bayes rule. To avoid model over-fitting, the optimal number of latent variables (LV) was chosen by detecting the global minimum of RMSE cross validation (RMSECV). Subsequently, a permutation test was used to evaluate the significance of the model. Permutation test provides the probability that the given model is significantly different from the one built under the same conditions but on random data. Repeatedly and randomly reordering of the y-block (shuffling the class assignments on each sample, assigning samples to the "wrong" classes), rebuilding the model with the current modelling settings after each reordering, was considered. In this specific case, the class assignment was permuted 50 times.

PLS models were developed to predict the moisture content of all samples. The experimentally determined moisture contents were used as reference values (Y), while the mean spectra (72) calculated considering the three acquisitions carried out along each swath were used as independent variables (X). This is because the hyperspectral acquisition and the sampling for the moisture determination were not done in the same perfect points. PLS models were validated by the venetian blind cross-validation method (segments: 10). Subsequently, the PLS prediction power was evaluated using an external data set composed of 36 samples acquired during summer 2019 (field 1) and adopting the same instrumental set-up. The results obtained for the external validation set were also used to evaluate the robustness of the model in terms of the degree of model generalization (Putri & Fukusaki, 2015).

The application with fewer wavebands is preferable for a more stable model and easier implementation in a subsequent multispectral imaging system

To reduce the computational complexity, improve the predictive ability of calibration models, and simplify the calibration models, a selection of the original wavelengths was used. Furthermore, a limited number of wavelengths could be directly used for developing a low-cost multispectral equipment compatible with market requirements.

179 The wavelength selection methods used in this study are based on the variable importance in
180 projection (VIP method). This method estimates the importance of each variable in the projection
181 used in a PLS model. In particular, the VIP score calculates the contribution of each variable
182 according to the variance explained by each PLS component. The 'greater than one' is conventionally
183 used as the criterion for variable selection; accordingly, only the variables with VIP scores greater
184 than one are selected.
185 The data elaboration was carried out by using PLS Toolbox and MIA Toolbox for Matlab2018a®
186 (https://www.mathworks.com).
187

188 **3. Results and discussion**

189 Air temperature and relative humidity data acquired before each set of acquisitions are reported in
190 Table 1. The temperature ranged from 25.34° to 34.36°C, while the RH from 25.25% to 65.93%.
191 Monitoring these parameters is important because environmental conditions can affect the drying
192 rate. Moisture content values (means and deviation standards) determined on leaf samples are shown
193 in Table 2 as a function of cut, conditioning level, time after conditioning, and time of day. The
194 highest initial moisture content values (from 75.91% to 76.67%) were achieved in the morning during
195 the second cut (field 2) at the same time as the highest RH (65.93%), while the lower ones (from
196 68.52% to 69.68%) belonged to the samples analysed during the day with lower RH (25.25%).
197 Changes in moisture content from 0 to 120 min calculated for all the combinations of daytime, field,
198 cut, and condition levels are shown in Table 3. Due to more intense breaks of stems and leaves, the
199 highest moisture losses were achieved for conditioning level C. This result is in agreement with those
200 reported by (Borreani et al., 1999), where an alfalfa 36% faster drying rate was proved with roll
201 nearness passing from 40% to 80%. The authors reported a drying rate that was 36% faster for alfalfa
202 passing from 40% to 80% of the roll nearness. After 120 min, the moisture content of the samples C
203 and W decreased on average by 11.4±1.8% and 8.4±2.0%, respectively.
204 The mean spectra pre-treated by SNV are shown in Figure 2a and compared with those reported in

literature: the reflectance signatures are similar to those reported for alfalfa acquired with a hyperspectral camera installed on UAV (Feng et al., 2020), more in generally, to typical reflectance spectra arising from green vegetation (Mishra et al., 2017). Typically, the visible region (400-700 nm) is dominated by absorption of light by photosynthetic pigments, while the NIR (700-1000) region by dry matter and water. Particularly, high reflectance in NIR region is due to the scattering in spongy mesophyll and low reflectance in the blue (450–500 nm) and red regions (600–700 nm), corresponding to chlorophyll absorption (Hansen & Schjoerring, 2003). In Figure 2b the mean spectra pre-treated by SNV+D1 are shown.

PLSDA results, in terms of numbers of latent variables (LV), percentage of samples correctly discriminated (model sensitivity), and probability to correctly discriminate, are reported in Table 4. For all the models, the following pre-treatment sequence was adopted: SNV+D1+MC. Considering both data sets (calibration and external validation), the highest sensitivity was achieved for the PLSDA model, which was able to discriminate between field 1 and field 2, with sensitivity values of 96.0% and 93.7%, respectively. The mean probability of belonging to the correct group (external validation) was higher than 80%. The PLSDA model developed to discriminate the samples in three classes (0, 15, and 120 min after conditioning) showed low sensitivity for class 2 (15 min), especially for the external validation. All the samples that were wrongly discriminated were attributed to group 1 (0 min). The significance of the PLSDA models was evaluated by employing permutation tests. The results in terms of the correlation coefficient between the original and permuted class vectors versus the standardized sums-of-squares (SSQ) are shown in Figures 3 and 4. In general, the cross-validated and self-prediction (calibration) values should be relatively close to each other but should be less than the results for the non-permuted y-block (right side of the plot), independently of their correlation with the real class values. The further away the un-permuted results are from the mean, the more unlikely it is the original model is over-fit. Consequently, the higher the y-value, the more likely it is that the original model is significant (Thomas, 2004). For all the binary (two classes) PLSDA models, the y-values (un-permuted) were higher than 3, suggesting good model significance.

231 Furthermore, p-values lower than 0.05 were obtained for cross-validation by employing a
232 randomization t-test, thus confirming the significance of the original models at a 95% confidence
233 level. For the PLSDA model developed to discriminate samples as a function of time after
234 conditioning (three classes), three permutation tests were employed, one for each class (Fig. 4). The
235 results suggested that the model concerning class 2 (15 min) is not significant. The y-values (cross-
236 validation) of the un-permuted model (right side of the plot) were slightly larger than one, and the p-
237 value obtained by the randomization t-test was 0.688.
238 PLS models were developed to predict the moisture content considering the entire spectral range. To
239 avoid model over-fitting, the optimal number of LV was selected by evaluating the root mean square
240 error in cross validation (RMSECV) as a function of the number of LV. The results in terms of
241 measured *versus* predicted moisture content (%) of the cross validated models are reported in Figures
242 5a and 5b for pre-treatment sequences SNV+MC and SNV+D1+MC, respectively. Furthermore,
243 determination coefficients (R^2) RMSE and BIAS in calibration and cross-validation are shown.
244 Moisture content was predicted with R^2 of 0.839 (RMSECV=2.08%) and 0.826 (RMSECV=2.14%),
245 for SNV+MC and SNV+D1+MC, respectively. The results are comparable with those obtained by
246 Digman & Shinners, (2008) which measured the moisture content of cut alfalfa using NIR reflectance
247 spectroscopy. In particular, the authors elaborated the spectral data by PLSr reported, in cross-
248 validation, R^2=0.83 and RMSE= 3.7 % (indicatively for a mean moisture content of 60%).
249 Considering the good results achieved on calibration and cross-validation, the model prediction power
250 was also evaluated using an external data set (36 samples acquired in 2019). The results reported in
251 Figures 5c and 5d (R^2 higher than 0.86) confirm the model's predictive power. The robustness of the
252 models, in terms of degree of model generalization, was expressed as the ratio between the apparent
253 performance (R_C^2 and RMSEC) and external validation performance ($R_P^2$and RMSEP). When the
254 apparent and external validation predictive powers are similar, it is possible to affirm that the model
255 is robust. The values of RMSEC/RMSEP from 0.75 to 0.79 0.75 and R_P^2/ R_C^2 from 0.94 0.98, prove
256 the model's robustness.

257 VIP scores obtained by the PLS model (Fig. 6) were considered to evaluate the importance of each
258 variable in the projection used in a PLS model and to reduce the spectral range. A variable with a VIP
259 score greater than unity can be considered important in a given model. Consequently, the variable
260 selection method based on the VIP scores (>1) was used to reduce the original data set and
261 consequently to remove unnecessary spectral regions. The spectral region with VIP>1 is placed in
262 the range of the absorption bands of water (760 nm and 960–970 nm). Two other PLS models were
263 built and validated considering the reduction of the data set (74 and 81 points for SNV+MC and
264 SNV+D1+MC, respectively). Measured *versus* predicted moisture content (%) are reported in Figure
265 7 for the calibration (grey symbols) and external validation (black symbols) data sets. Prediction R^2
266 (higher than 0.80) were slightly lower than those obtained considering the entire spectral range, but
267 completely acceptable. The best prediction (R^2=0.818, RMSEP=2.54%) was achieved by the model
268 developed when pre-treating the data by SNV+D1+MC. A decrease of wavelength number by more
269 than 70% was achieved, this could be a first step in developing a low cost-multispectral camera to
270 measure the alfalfa moisture content directly in field.
271 The spectral and spatial information of each pixel in hyperspectral images allow for the evaluation of
272 the reference parameter of each pixel with chemometric models. The pixels having similar spectra
273 showed similar colours in the images, and consequently similar predicted values. Using the PLS
274 model developed considering the reduced spectral range (SNV+D1+MC), false colour images were
275 obtained. The prediction maps of two representative samples (acquired immediately after
276 conditioning and after 120 min) are shown in Figure 8. The colour bar (from blue to yellow) indicates
277 the scale of the reference values (moisture content from 50 to 80%). Moisture spatial distributions are
278 in alignment with the measured values, and the mean predicted moistures were 72.3±14.3% and
279 65.3±18.1% for the sample acquired immediately and after 120 min from conditioning, respectively.
280 The distributions of the pixel frequency as a function of the predicted moisture content are shown as
281 histograms. The noise of the hyperspectral image, influenced by a non-homogeneous alfalfa
282 distribution, and by the presence of holes, flowers, any foreign bodies, can affect the spectrum of each

pixel, which may result in the under/over estimation of the moisture. Probably, several and specific masks to apply during a hypercube pre-processing phase could be developed to remove the pixels related to flowers, foreign bodies, showdown and so no. However, considering the high heterogeneity of the hyperspectral images due to many experimental conditions, the mask creation would make the application of the proposed method rather complicated, without a sure, and prominent improvement in the model's performance.

Conclusions

This research presents a solution for in-field and non-destructive evaluation of mowed and conditioned alfalfa quality parameters by Vis/NIR HSI combined with chemometric techniques. PLSDA models allowed discrimination of samples according to several operative conditions (level of conditioning, field type, time after conditioning, and time of day) with good discrimination power in external validation (mean sensitivity=87%, mean probability=89%). These models may be useful to remotely determine operative conditions during mowing or harvest procedures in real time. PLS models were developed to estimate the moisture content considering different spectral pre-treatments and obtained R^2=0.86 and RMSE=2.00%, in external validation. To reduce the computational time required for data elaboration, the spectral range (numbers of x-variables) was reduced by using the VIP method, passing from 272 to 74 bands. Further PLS models were built by considering the reduced variable numbers and achieved results that were comparable to those reported with the full spectra, demonstrating that a reduction of the number of variables did not affect the goodness of the models. The best PLS model was applied to each pixel of the hyperspectral images to obtain false colour images. Pixels having similar colours are characterized by comparable moisture content. The results confirm the potentiality of in-field, non-destructive, and real-time determination of cut alfalfa moisture content using HSI combined with linear chemometric techniques. The limits concerning the acquisition time that influence the tractor speed of the tractor must be considered. Moreover, since the chemometric models were developed considering alfalfa samples (hyperspectral images and

309 moisture content) of the swath surface, the relationship between surface and internal moisture of the
310 swaths should be evaluated.

Funding: This work was supported by the CNH Industrial Italia SpA.

References

Benelli, A., Cevoli, C., & Fabbri, A. (2020). In-field hyperspectral imaging: An overview on the ground-based applications in agriculture. *Journal of Agricultural Engineering*, *LI*(1030), 129–139. https://doi.org/10.4081/jae.2020.1030

Borreani, G., Tabacco, E., & Ciotti, A. (1999). Effects of Mechanical Conditioning on Wilting of Alfalfa and Italian Ryegrass for Ensiling. *Agronomy Journal*, *91*, 457–463. https://doi.org/10.2134/agronj1999.00021962009100030016x

Capolupo, A., Kooistra, L., Berendonk, C., & Boccia, L. (2015). Geo-Information Estimating Plant Traits of Grasslands from UAV-Acquired. *ISPRS International Journal of Geo-Information*, *4*, 2792–2820. https://doi.org/10.3390/ijgi4042792

Colombo, R., Busetto, L., Meroni, M., Rossini, M., & Panigada, C. (2012). Optical Remote Sensing of Vegetation Water Content. In P. S. Thenkabail, J. G. Lyon, & A. Huete (Eds.), *Hyperspectral Remote Sensing of Vegetation* (pp. 227–239). CRC Press.

Daszykowski, M., Walczak, B., & Massart, D. L. (2002). Representative subset selection. *Analytica Chimica Acta*, *468*(July), 91–103.

Digman, M. F., & Shinners, K. J. (2008). Real-time moisture measurement on forage harvest using near-infrared reflactance spectroscopy. *Transactions of the ASABE*, *51*(5), 1801–1810.

Feng, L., Zhang, Z., Ma, Y., Du, Q., & Williams, P. (2020). Alfalfa Yield Prediction Using UAV-Based Hyperspectral Imagery and Ensemble Learning. *Remote Sensing*, *12*, 2–24.

Feng, L., Zhang, Z., Ma, Y., Sun, Y., Du, Q., Williams, P., … Luck, B. (2021). Multitask Learning of Alfalfa Nutritive Value From UAV-Based Hyperspectral Images. *IEEE GEOSCIENCE AND REMOTE SENSING LETTERS*, *19*, 1–5.

Feret, J., François, C., Asner, G. P., Gitelson, A. A., Martin, R. E., Bidel, L. P. R., … Jacquemoud,

337 S. (2008). Remote Sensing of Environment PROSPECT-4 and 5 : Advances in the leaf optical
338 properties model separating photosynthetic pigments. *Remote Sensing of Environment*, *112*,
339 3030–3043. https://doi.org/10.1016/j.rse.2008.02.012
340 Geipel, J., Bakken, A., Jørgensen, M., & Korsaeth, A. (2021). Forage yield and quality estimation by
341 means of UAV and hyperspectral imaging. *Precision Agriculture*, (1).
342 https://doi.org/10.1007/s11119-021-09790-2
343 Hansen, P. M., & Schjoerring, J. K. (2003). Reflectance measurement of canopy biomass and nitrogen
344 status in wheat crops using normalized difference vegetation indices and partial least squares
345 regression. *Remote Sensing of Environment*, *86*, 542–553. https://doi.org/10.1016/S0034-
346 4257(03)00131-7
347 Lim, J., Watanabe, N., Yoshitoshi, R., & Kawamura, K. (2020). Simple in-field evaluation of
348 moisture content in curing forage using normalized differece vegetation index (NDVI).
349 *Grassland Science*, *66*, 238–248. https://doi.org/10.1111/grs.12275
350 Long, E. A., Ketterings, Q. M., Russell, D., Vermeylen, F., & Degloria, S. (2016). Assessment of
351 yield monitoring equipment for dry matter and yield of corn silage and alfalfa/grass. *Precision*
352 *Agriculture*, *17*(5), 546–563. https://doi.org/10.1007/s11119-016-9436-y
353 Manley, M. (2014). Near-infrared spectroscopy and hyperspectral imaging: non-destructive analysis
354 ofbiological materials. *Chem.Soc.Rev*, *43*, 8200–8214. https://doi.org/10.1039/c4cs00062e
355 Mishra, P., Shahrimie, M., Asaari, M., Herrero-langreo, A., Lohumi, S., & Scheunders, P. (2017).
356 Close range hyperspectral imaging of plants : A review. *Biosystems Engineering*, *4*, 49–67.
357 https://doi.org/10.1016/j.biosystemseng.2017.09.009
358 Øvergaard, Stein I, Isaksson, T., & Korsaeth, A. (2013). Prediction of wheat yield and protein using
359 remote sensors on plots — Part I : Assessing near infrared model robustness for year and site
360 variations. *Journal of Near Infrared Spectroscopy*, *131*(April), 117–131.
361 https://doi.org/10.1255/jnirs.1042
362 Øvergaard, Stein Ivar, Isaksson, T., & Korsaeth, A. (2010). Comparisons of two hand-held ,

multispectral field radiometers and a hyperspectral airborne imager in terms of predicting spring wheat grain yield and quality by means of powered partial least squares regression. *Journal of Near Infrared Spectroscopy*, *261*(October), 247–261. https://doi.org/10.1255/jnirs.892

Putri, S. P., & Fukusaki, E. (2015). *Mass spectrometry-based metabolomics: A practical guide.* CRC Press.

Rinnan, Å., van den Berg, F., & Engelsen, S. B. (2009). Review of the most common pre-processing techniques for near-infrared spectra. *TrAC Trends in Analytical Chemistry*, *28*(10), 1201–1222. https://doi.org/10.1016/J.TRAC.2009.07.007

Serrano, L., Ustin, S. L., Roberts, D. A., Gamon, J. A., & Pen, J. (2000). Deriving Water Content of Chaparral Vegetation from AVIRIS Data. *Remote Sensing of Environment*, *581*(May), 570–581.

Sims, D. A., & Gamon, J. A. (2003). Estimation of vegetation water content and photosynthetic tissue area from spectral reflectance : a comparison of indices based on liquid water and chlorophyll absorption features. *Remote Sensing of Environment*, *84*, 526–537.

Suzuki, Y., Tanaka, K., Kato, W., Okamoto, H., Kataoka, T., Shimada, H., … Shima, E. (2008). Field mapping of chemical composition of forage using hyperspectral imaging in a grass meadow. *Japanese Society of Grassland Science*, *54*, 179–188. https://doi.org/10.1111/j.1744-697X.2008.00122.x

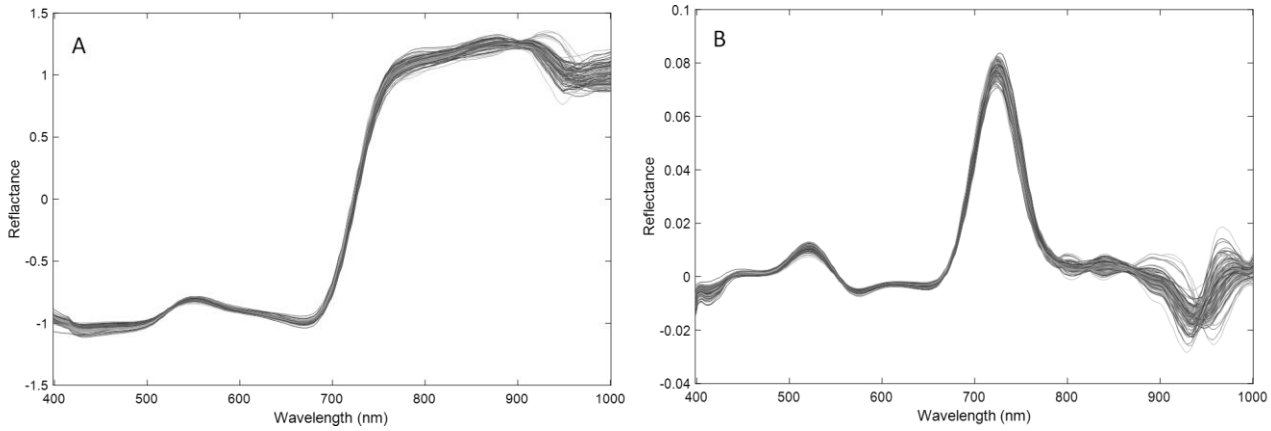
Thomas, V. E. (2004). Non-parametric statistical methods for multivariate calibration model selection and comparison y. *Journal of Chemometrics*, *17*, 653–659. https://doi.org/10.1002/cem.833

383

384 **Figure 1.** Instrumental set-up installed on the front linkage of the tractor. A: hyperspectral camera.

385 B.white reference material, C: Wi-Fi access point.

386



387

388 **Figure 2.** Reflectance mean spectra of all acquired samples. A: standard normal variate (SNV) pre-

389 treatment; B: first derivative (D1) pre-treatment.

390

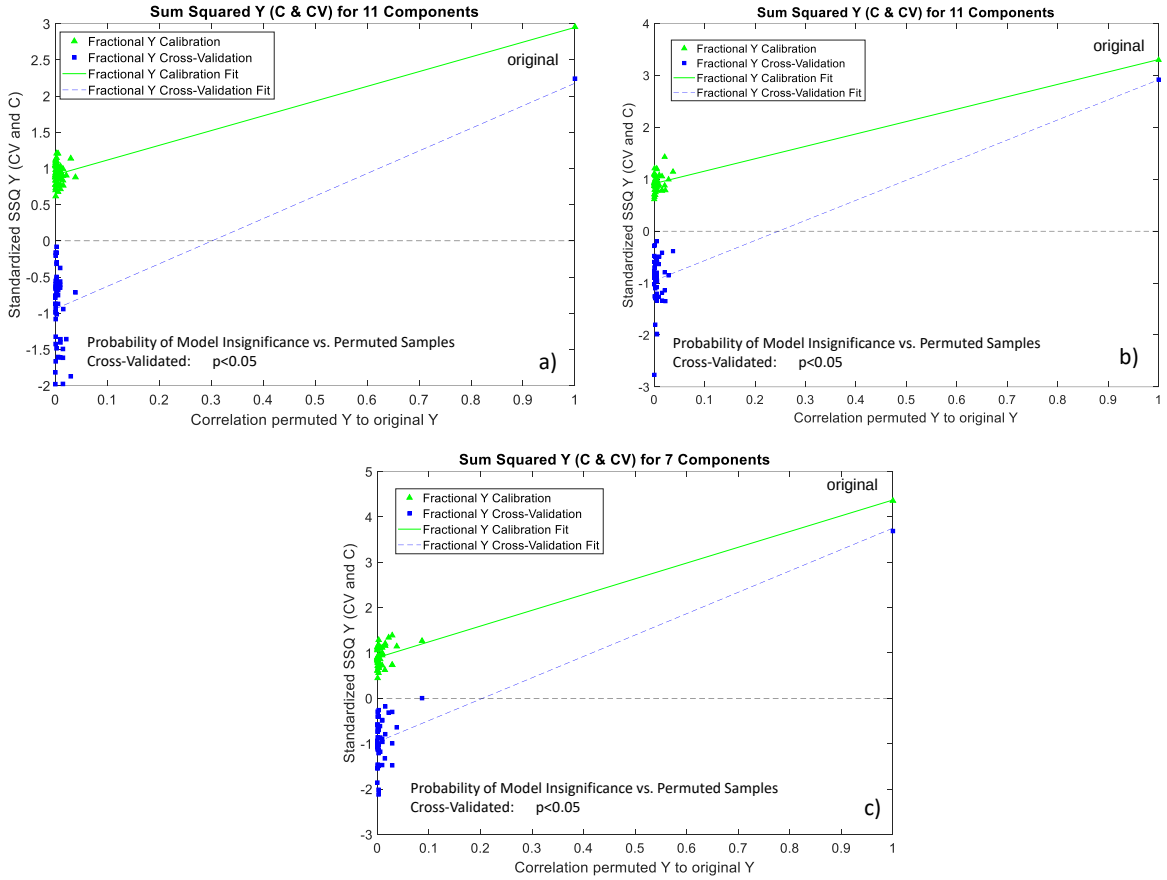


Figure 3. Permutation test results of PLSDA models (a: conditioning level; b: field type; c: time of day).

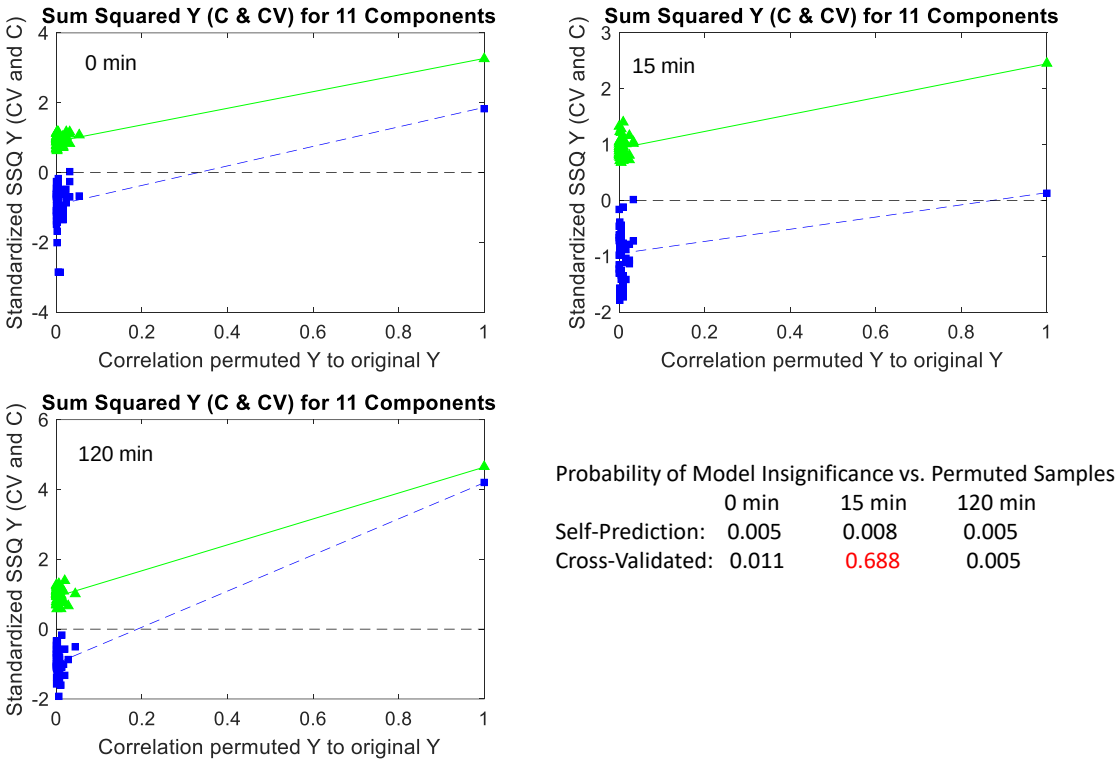
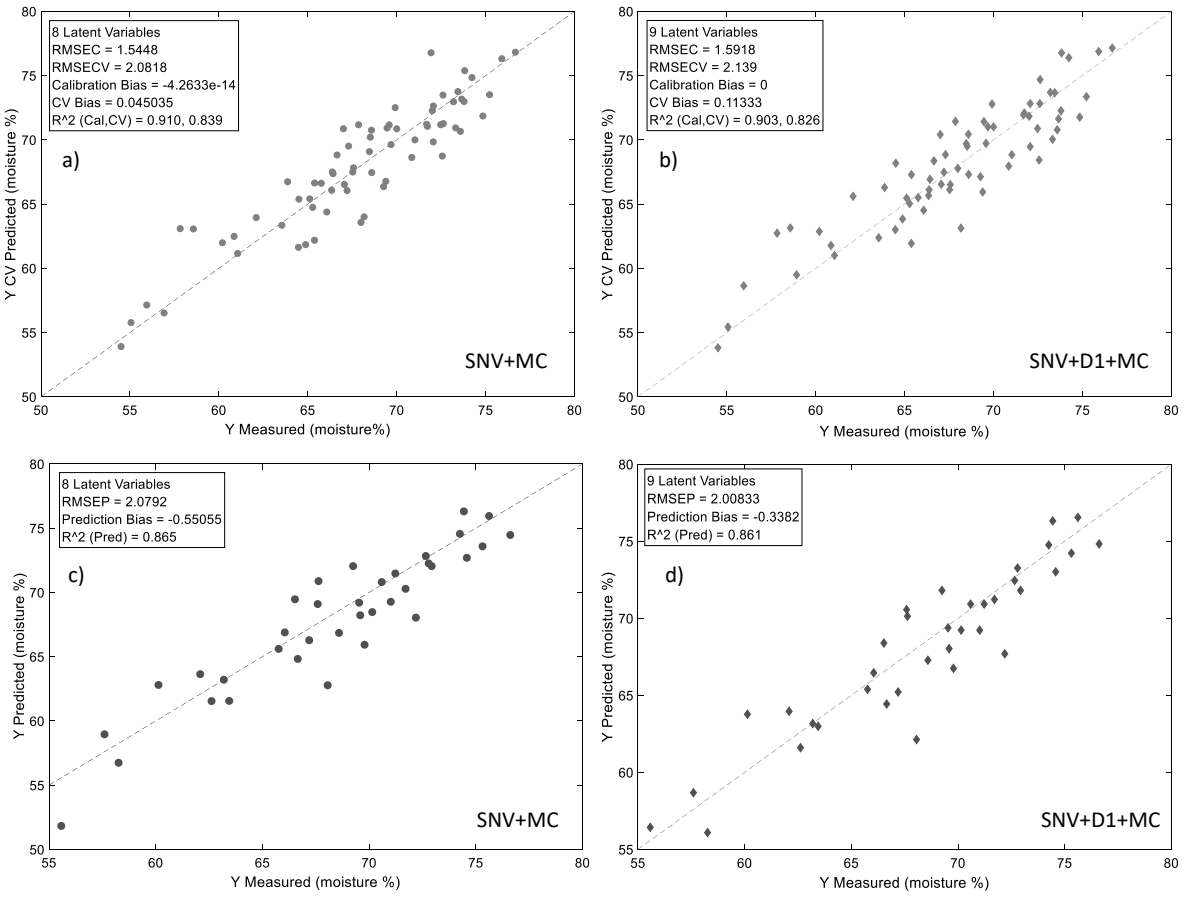


Figure 4. Permutation test results of the PLSDA model developed to discriminate as a function of time after conditioning (0, 15, and 120 min)

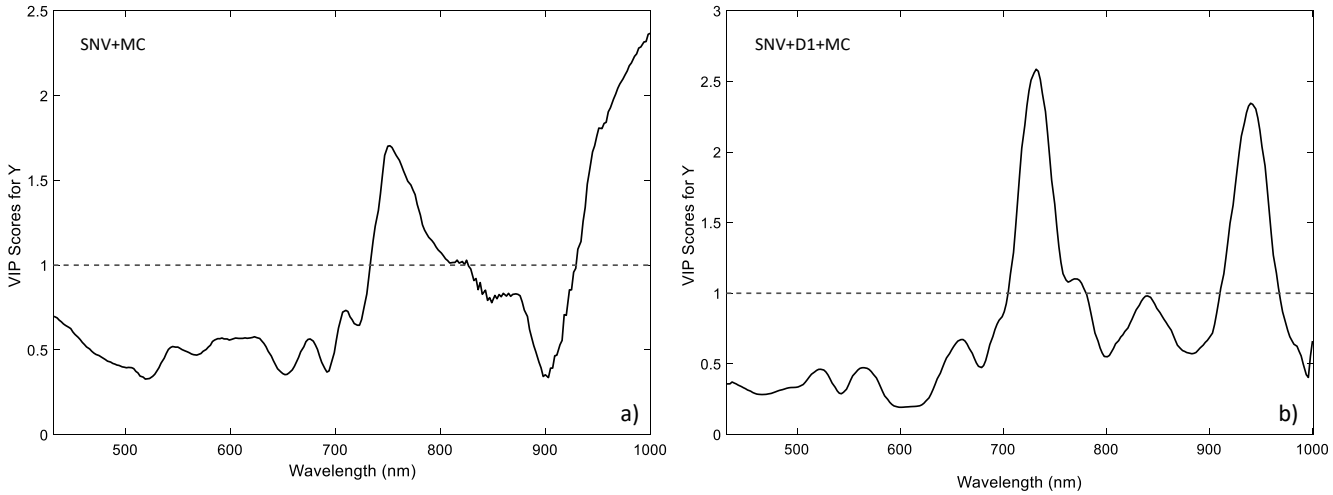


401

402 **Figure 5.** PLS results in terms of measured *versus* predicted moisture content (%) values as a function
403 of spectral pre-treatment (standard normal variate: SNV; first derivative: D1; mean cantering: MC)
404 and data set (a and b: cross-validation data: c and d: external validation data).

405

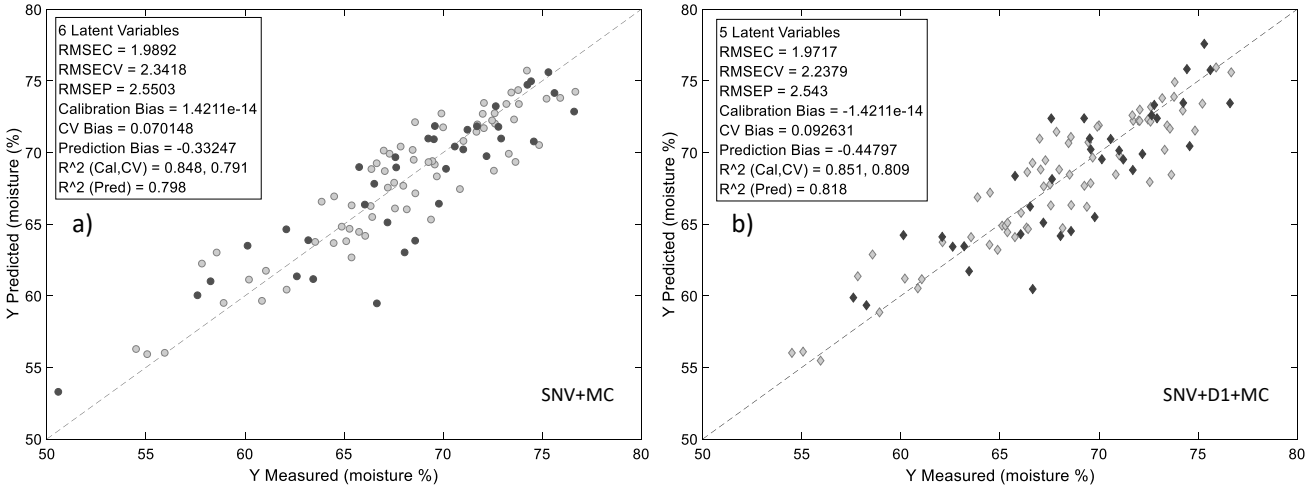
406



407

408 **Figure 6.** VIP scores obtained by PLS regression models developed considering the full spectra, as a
409 function of spectral pre-treatment (standard normal variate: SNV; first derivative: D1; mean
410 cantering: MC).

411



412

413 **Figure 7.** PLS results in terms of measured *versus* predicted moisture content (%) values as a function

414 of spectral pre-treatment (standard normal variate: SNV; first derivative: D1; mean cantering: MC)

415 (grey symbols: cross validation data, black symbols: external validation data).

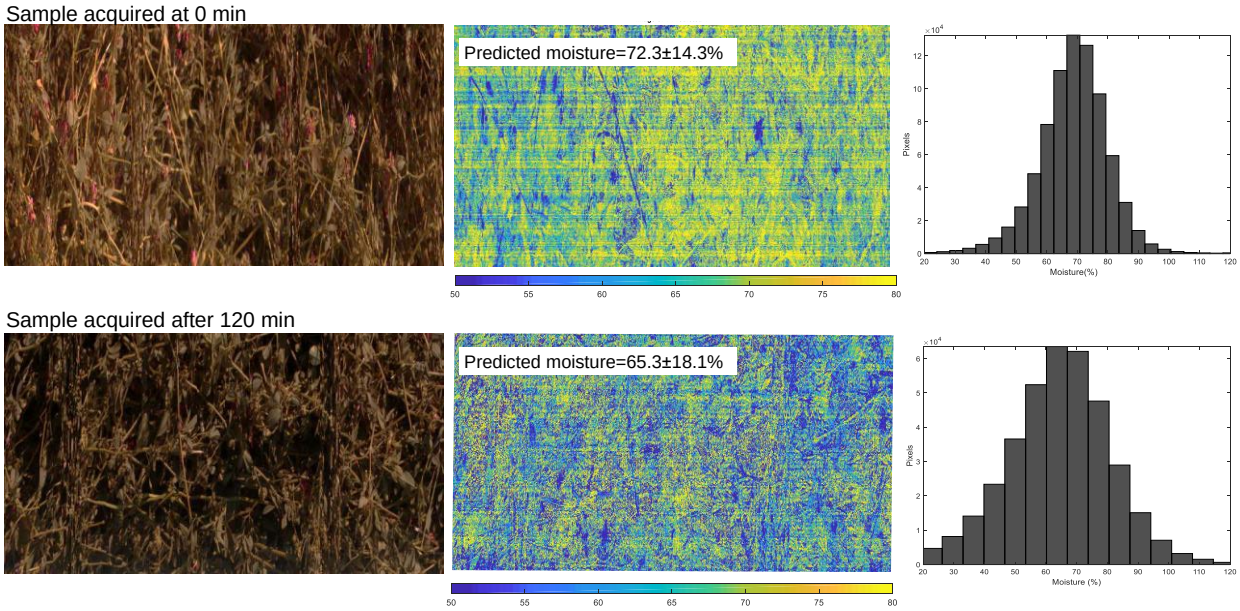


Figure 8. Prediction maps (centre) of two representative samples (left) acquired immediately after conditioning and 120 min later.

422 **Table 1.** Environmental temperature and relative humidity (RU) acquired in field.

cut	field	Morning		Afternoon	
		temperature (°C)	*RH (%)*	*temperature (°C)*	*RH (%)*
1	1	29.71	36.21	31.93	28.26
	2	32.83	34.82	34.36	25.25
2	1	26.26	55.32	32.43	42.16
	2	25.34	65.93	29.21	48.38
3	1	26.26	56.24	30.92	40.87
	2	26.98	51.45	30.38	40.46

423

Table 2. Moisture content (means and standard deviations) of alfalfa samples depending on forage cut (1, 2, and 3), conditioning level (C: close and W: wide), time after conditioning (0, 15, and 120 min), field type (1 and 2), and time of day (morning and afternoon).

Cut	Conditioning level	Time after conditioning (min)	Moisture content (%)							
			field 1				field 2			
			morning		*afternoon*		*morning*		*afternoon*	
1	C	0	73.31	0.93	75.22	0.25	69.53	0.21	69.58	1.58
		15	70.85	0.05	72.00	0.62	66.87	0.25	66.34	0.28
		120	61.06	1.45	63.54	0.74	54.50	1.50	55.06	1.06
	W	0	73.19	0.18	73.82	0.31	69.68	1.19	68.52	0.21
		15	71.73	0.86	72.62	1.70	64.50	0.50	67.21	0.07
		120	65.17	0.52	65.60	0.18	58.57	0.57	57.82	1.25
2	C	0	73.79	0.75	73.65	0.94	76.67	1.00	72.06	0.73
		15	68.58	0.42	70.00	0.94	74.85	1.42	70.00	0.94
		120	62.92	0.78	65.38	1.06	65.76	0.22	62.53	1.06
	W	0	72.47	1.05	72.06	0.93	75.91	0.22	72.60	1.30
		15	66.65	0.50	67.30	2.98	71.68	1.66	71.03	0.81
		120	64.47	0.57	66.38	1.20	66.42	1.84	61.11	1.60
3	C	0	67.86	1.20	69.46	0.86	73.44	1.20	72.56	0.73
		15	67.06	0.96	67.99	0.69	73.58	1.42	69.20	0.94
		120	55.95	0.32	58.11	0.98	63.28	0.22	62.10	1.06
	W	0	69.39	0.32	67.00	1.36	74.23	0.22	69.91	1.30
		15	66.07	1.65	67.57	1.46	71.93	1.66	69.26	0.81
		120	61.46	1.89	60.86	0.56	65.38	1.84	64.88	1.60

Table 3. Moisture content decrease (%) from 0 to 120 min of alfalfa samples depending on forage cut (1, 2 and 3), conditioning level (C: close and W: wide), field type (1 and 2), and time of day (morning and afternoon).

Cut	Conditioning level	Moisture content decrease (%)			
		field 1		field 2	
		morning	*afternoon*	*morning*	*afternoon*
1	C	12.3	11.7	15.0	14.5
	W	8.0	8.2	11.1	10.7
2	C	10.9	8.3	10.9	9.5
	W	8.0	5.7	9.5	11.5
3	C	11.9	11.3	10.2	10.5
	W	7.9	6.1	8.8	5.0

433 **Table 4.** Results of the PLSDA developed to discriminate the samples according to conditioning level
434 (C and W), field type (field-1 and field-2), time of day (morning and afternoon), and time after
435 conditioning (0, 15, and 120 min).

	Spectral pre-reatment	LV	Samples correctly grouped (%)			Probability (external validation)
			calibration n=173	cross-validation n=173, 10 segments	external validation n=43	
Conditioning level						
C	SNV+D1+MC	11	90.1	86.4	86.9	93.3(14.4)
W			92.8	86.7	89.0	92.1(11.6)
Field						
1	SNV+D1+MC	11	96.2	95.1	96.1	95.5(6.8)
2			97.6	95.2	93.7	94.9(7.5)
Time of day						
morning	SNV+D1+MC	7	92.4	88.6	88.2	80.0(19.9)
afternoon			89.4	87.1	81.2	85.5(15.5)
Time after conditioning (min)						
0	SNV+D1+MC	11	81.2	82.1	82.4	81.5(15.6)
15			80.7	71.2	65.3	82.4(16.7)
120			94.5	89.3	98.2	97.8(5.2)

436 Note: SNV: standard normal variate; D1: first derivative; MC: mean centering; C: close; W: wide; LV: latent variables;
437 standard deviations in parentheses.
438

439