This is the accepted version of the article:

Rubio-Delgado, J., Pérez, C.J. & Vega-Rodríguez, M.A. Predicting leaf nitrogen content in olive trees using hyperspectral data for precision agriculture. Precision Agric 22, 1–21 (2021). https://doi.org/10.1007/s11119-020-09727-1

Predicting leaf nitrogen content in olive trees using hyperspectral data for precision agriculture

Judit Rubio-Delgado[1], Carlos J. Pérez[1], Miguel A. Vega-Rodríguez[2]

Corresponding author: juditrd@unex.es

1. Department of Mathematics, University of Extremadura, 10003, Cáceres, Spain.

2. Department of Computer and Communications Technologies, University of Extremadura, 10003, Cáceres, Spain.

Abstract

Olive orchard is one of the main crops in the Mediterranean basin and, particularly, in Spain, with 56% of European production. In semi-arid regions, nitrogen (N) is the main limiting factor of olive trees after water and its quantification is essential to carry out accurate fertilization planning. In the present study, N status of an olive orchard located in Carmonita (southwest Spain) was analysed using hyperspectral data. Reflectance data were recorded with a high precision spectro-radiometer through the full spectrum (350–2500 nm). Different vegetation indices (VI), combining two or three wavelengths, and partial least squares regression (PLSR) models were developed, and the prediction capabilities were compared. Different pre-processing (smoothing, SM; standard normal variate, SNV; first and second derivative) were applied to analyse the influence of the noise generated by the spectro-radiometer measurements when computing the determination coefficient between leaf N content (LNC) and spectra data. Results showed that second derivative combined with SNV pre-processing produced the best determination coefficients. The wavelengths most sensitive to N variation used to perform VI were selected from the visible and the short-wave infrared spectrum regions, which relate to chlorophyll a+b and N absorption features. DCNI and TCARI showed the best fittings for the LNC prediction (R^2=0.72, R^2_{cv}=0.71; and R^2=0.64, R^2_{cv}=0.63, respectively). PLSR models yielded higher accuracy than the models based on VI (R^2=0.98, R^2_{cv}=0.56), although the large difference between calibration and cross-validation showed more uncertainty in the PLSR models.

Keywords: Leaf nutritional status, Linear regression, Nitrogen indices, Olive orchards, Partial Least Squares Regression, SWIR spectral region.

Introduction

Olive (Olea europaea L.) is one of the main crops in the Mediterranean basin. In 2016, the area under olive trees accounted for about 5 Mha in the European Union, with a production of 11.8 Mt of olives (61% of the worldwide yield; FAOSTAT 2017). 51% of the total European olive orchard surface is concentrated in Spain, 23% in Italy, 17% in Greece, 7% in Portugal and the 2% remaining in other countries (EUROSTAT 2017). In Spain, in recent years, the oleic sector contributed between 8% and 13% of annual agricultural production (FAOSTAT 2017).

Traditionally, the yield of most Spanish olive orchards was limited by water supply, and soil management was mainly based on plough, disk and harrow tillage operations (López-Granados et al. 2004). Currently, precision agriculture techniques are being introduced in the sector and are targeted to reduce economic costs and adverse environmental effects of agricultural inputs (Akdemir et al. 2018).

Nitrogen (N) is an essential element for plant growth and is the mineral nutrient most commonly applied in agriculture and in olive orchards fertilization. Between 2002 and

2014, more than 1,200 Mt of N were consumed in agriculture worldwide, compared to the 500 and 400 Mt of potassium (K) and phosphorus (P), respectively. During this period, N consumption increased by 32% (FAOSTAT 2017). Over-fertilization has negative economic and environmental effects. N loss by leaching and ground water can result in ground and aquifer contamination (Puckett et al. 2011). This affects carbon storage (Schulze et al. 1989) and produces a N deficiency that results in a reduction in yields and, therefore, economic losses to the farmers (Haboudane et al. 2002). This situation motivates the need to carry out efficient and sustainable management programmes of agricultural fertilization.

Traditional techniques to estimate crop nutrient status are based on leaf sampling and foliar analysis. However, these methods are destructive, time-consuming and expensive. Moreover, traditional N estimates provide limited information, as sampling is based on only a limited number of sites in a given field (Camino et al. 2018). Therefore, it is necessary to consider efficient alternatives.

In recent years, remote sensing for characterizing biophysical parameters of vegetation has had important potential implications for predicting chlorophyll (Chl) content as a proxy of the plant N status (Clevers et al. 2017). Several studies have demonstrated strong correlations between N content and Chl in crops, because of the large amount of protein that complexes the photosynthetic pigment (Cilia et al. 2014; Clevers et al. 2017; Singh et al. 2017; Miphokasap and Wannasiri 2018).

Spectral sensing is a spectroscopic method commonly used for assessing N content (Alchanatis et al. 2009). Hyperspectral data contain large amounts of redundant information due to the relatively few parameters that really and effectively control the spectral signatures of vegetation (Im and Jensen 2008). This relatively low number of variables contrasts with the often more than 100 wavelengths available through imaging spectroscopy and from commercially available spectro-radiometers (Atzberger et al. 2010). Therefore, in order to develop suitable indices, it is necessary to determine those wavelengths that presented a better correlation with the measured biophysical parameters.

Reflectance spectra in the short-wave infrared (SWIR) spectrum region (1100–2500 nm) have been found to be highly correlated with N content. Reflectance in this region is directly related to N-hydrogen (H) stretch, first overtone and absorption features of protein (Curran 1989). Several authors have shown improvements in N content estimations using SWIR spectral region to determine vegetation indices (VI), sometimes in combination with the visible and near-infrared (VNIR) region (350–1100 nm) (Cohen and Alchanatis 2018). Serrano et al. (2002) showed that $NDNI_{1510\_1680}$ (Normalized Difference Nitrogen Index), using log_{10} transformed reflectance, was sensitive to N concentration in chaparral vegetation. Ferwerda (2005) recommended NRI (Normalized Ratio Index) using the combination of reflectance at 1770 nm and 693 nm for the best correlation with N content of different species: olive, willow, mopane, grass and shrubs. Herrmann et al. (2010) obtained the best estimated N content in potatoes plants using the NRI_{1510} and $MCARI_{1510}$ indices, which combined information from the 1510 nm and 660 nm wavelengths. Camino et al. (2018) reported that the use of the SWIR spectral range to determine VI provided better quantification of N concentration in wheat plants than when only the VNIR region was used. They found the highest agreement with N concentration using $MCARI_{1510}$ and $NDI_{850,1510}$ (vegetation index estimated in a similar way to NDVI) indices. Table 1 summarizes the VIs used by other authors to determine the N status of several crops using wavelengths within the range 400-1800 nm.

Table 1. Vegetation indices (VI) used by other authors to determine the N status of crops, using wavelengths within the range 400–1800 nm (SR, Simple Ratio; NDI, Normalized Difference Index; MCARI, Modified Chlorophyll Absorption in Reflectance Index/ Optimized Soil Adjusted Vegetation Index; TCARI, Transformed Chlorophyll Absorption Ratio Index/ Optimized Soil Adjusted Vegetation Index; DCNI, Double-Peak Canopy Nitrogen Index).

VI	Equation	Reference
$SR_{550,670}$	R_{550}/R_{670}	Gómez-Casero et al. (2007)
$SR_{780,550}$	R_{780}/R_{550}	Gómez-Casero et al. (2007)
$SR_{780,670}$	R_{780}/R_{670}	Gómez-Casero et al. (2007)
$NDI_{780,670}$	$(R_{780} - R_{670})/(R_{780} + R_{670})$	Gómez-Casero et al. (2007)
DCNI	$[(R_{720} – R_{700})/(R_{700} – R_{670})]/(R_{720} – R_{670} + 0.03)$	Chen et al. (2010)
$MCARI_{1510}$	$[(R_{700} – R_{1510}) - 0.2(R_{700} – R_{550})](R_{700}/R_{1510})$	Herrmann et al. (2010)
$TCARI_{1510}$	$3[(R_{700} – R_{1510}) - 0.2(R_{700} – R_{550})(R_{700}/R_{1510})]$	Herrmann et al. (2010)
$NDI_{1645,1715}$	$(R_{1645} – R_{1715})/(R_{1645} + R_{1715})$	Pimstein et al. (2011)
$NDI_{870,1450}$	$(R_{870} – R_{1450})/(R_{870} + R_{1450})$	Pimstein et al. (2011)
$NDI_{850,1450}$	$(R_{850} – R_{1450})/(R_{850} + R_{1450})$	Camino et al. (2018)

The use of spectroscopy technology applied to olive yield has been mainly focused on the determination and identification of adulterants in olive oils (Muik et al. 2004). Few studies have been found about N content estimation and other nutrient deficiencies in olive orchards. Zarco-Tejada et al. (2004) used spectral data within the range 500–800 nm to study the effects of scene components (soil, shadow and crown reflectance) on the estimation of canopy Chl content in olive orchards, using a digital airborne imaging spectrometer and reflective optics system imaging spectrometer images at 1 m spatial resolution. Ferwerda and Skidmore (2007) estimated the chemical composition (N, F, K, Ca, Na and Mg) of three tree species (willow, mopane and olive) and one shrub species (heather) using hyperspectral data (300–2500 nm) and different spectral pre-processing (reflectance spectral, derivative spectral data and band depth). Gómez-

Casero et al. (2007) used the VNIR spectral region (in this case, 400–900 nm) to determine the hyperspectral reflectance curves of olive trees with different N or K leaf content and to select the best estimating VI for N and K content. Rotbart et al. (2013) analyzed the feasibility of determining N concentration in olive leaves using spectra reflectance in the VNIR-SWIR (in this case, 450–1700 nm) range under laboratory conditions and evaluated different types of spectrometers, different levels of sample preparation and different types of mathematical pre-processing to generate spectral models.

In conclusion, despite the wide literature on N estimation by spectral measurements, very few studies are related to the approach of this work. Ferwerda and Skidmore (2007) selected the bands with optimal fits to estimate leaf N content through a stepwise regression routine, without taking into account the formulation of any VI, and requiring four predictors. With respect to the study by Rotbart et al. (2013), its objective was not to determine suitable bands to estimate leaf N content (LNC). They have showed the power of using the whole spectra to estimate LNC in olive trees and five pre-processing methods to develop PLRS models. Additionally, they have done that under laboratory and not under field conditions.

Taking into account the importance of olive crop expansion, the yield in the Mediterranean region and the few available studies about estimating N content in olive trees using hyperspectral data, the present paper contributes to assess the usefulness of olive leaf spectral features for the objective of estimating the nutritional status of this crop. In this sense, the main aim of this study was to optimize the estimation of leaf N content in olive trees. To do this the following sub-objectives have been addressed:

i) Analysing the potentiality of the entire spectral region (350–2500 nm) dataset to estimate LNC in olive trees under field conditions.

ii) Testing eight different pre-processing methods (smoothing, standard normal variate, first and second derivative, and different combinations among them) to reduce the noise of the reflectance curves.
iii) Developing, implementing and applying a wavelength combination-based method (using three predictors) to consider band combinations that produce the optimal fit to the LNC.
iv) Comparing the predictive power of partial least squares regression (PLSR) and VI estimations.

Material and methods

Study area and field data collection

The fieldwork was conducted during two consecutive days in July 2018 in an organic olive orchard located in Carmonita (Badajoz) (39°09'04" N, 6°19'17" W), in southern Spain.

The climate of the study area is semi-arid Mediterranean, with an average annual temperature of 16 ºC and mean annual rainfall amounts to 550 mm. The summers are hot (24.5 ºC) and dry (12.3 mm rainfall). Winters are warm (8.5 ºC) and humid (79.1 mm rainfall) (Ninyerola et al. 2005). The study site has an area of 1.22 ha and a soil depth about 0.60 m, with an average elevation of 400 m.a.s.l. (meters above sea level), a medium slope gradient of 2.3º, and a predominantly south aspect (Figure 1). The olive tree density is 150 trees/ha. The olive orchard is organic, and no fertilization treatments were applied.

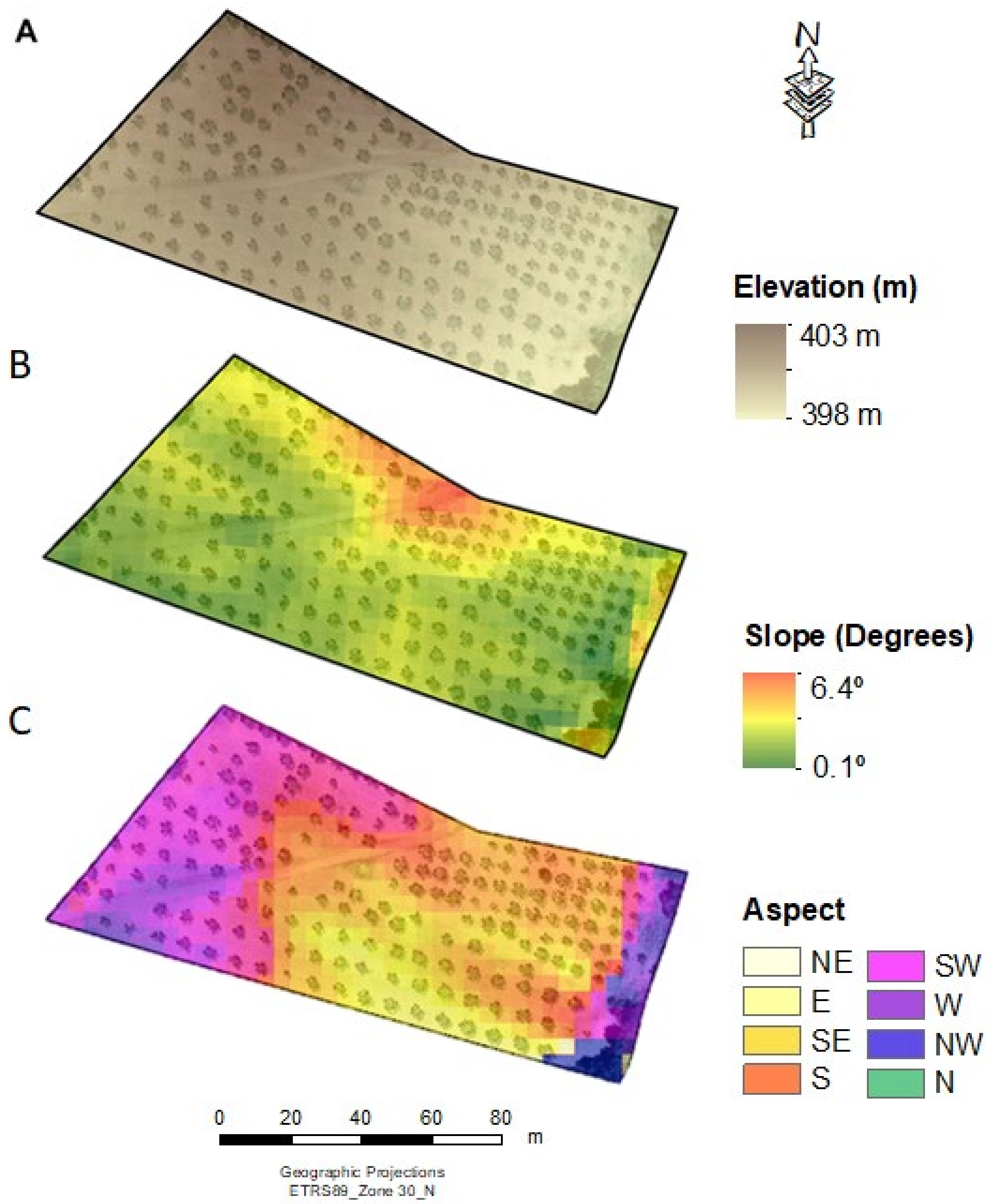

Fig. 1 Topographic digital models (pixel size of 25 m^2) of the study area: A) elevation, B) slope and C) aspect, generated using the cartography available at http://centrodedescargas.cnig.es (*Centro Nacional de Información Geográfica, Spain*).

To determine the real nutritional status of olive trees, a leaf analysis was carried out, collecting samples from 42 olive trees. To ensure that samples represent properly the study area, each sample was composed of 100 g of healthy leaves that were collected from the middle portion of current-season shoots, about 1.5 m above the soil surface, at the four cardinal points for every olive tree. Leaves from each sample were placed in a pile on the soil, with direct incidence of sunlight (Ferwerda and Skidmore 2007) and reflectance spectra were recorded using an Analytical Spectral Device (ASD) FieldSpec® 4 spectro-radiometer (Malvern, United Kingdom). This instrument records reflectance in the spectral range between 350 and 1000 nm, with a sampling interval of 1.4 nm, and between 1000 and 2500 nm, with a sampling interval of 1.1 nm. Spectral data were interpolated to a spectral band width of 1 nm using the ASD software. In order to calculate absolute reflectance, a reference spectrum was measured from a Spectralon reference target between readings, at 15-minutes time intervals. For each sample, 20 reflectance spectra were recorded and averaged, obtaining a single spectral curve per olive tree. The spectral measurements were collected ±2 h around solar noon, under clear-sky conditions and in nadir orientation. Finally, leaves were placed in paper bags and were taken to the laboratory following the protocol established by the Agrifood Laboratory of the Junta de Extremadura (Cáceres, Spain).

The critical and sufficiency threshold of LNC of each olive tree were estimated at 1.4% and 1.5%, respectively (Barranco et al. 1997; Fernández-Escobar et al. 1999).

Spectral data pre-processing

Spectral datasets were considered to predict N status of olive orchards and to identify wavelengths directly related to LNC in olive trees. Spectra are often disturbed by different interferences in the signal acquisition process due to the structure and physical properties of the samples. Frequently, derivative transformations of the reflectance data provide the best explanation of the variation, removing part of the noise (Hruschka 1987). Therefore, to minimize the undesired influence of physical attributes of samples on measured spectra, mathematical pre-processing has become an important issue in NIR spectral modelling (Bi et al. 2016).

In the present study, raw data were pre-processed using the following empirical methods and mathematical operators: Savitzky-Golay Smoothing, which applies a convolution method (SM; Savitzky and Golay 1964), Standard Normal Variate (SNV; Barnes et al. 1989,1993), first derivative of the reflectance (D_1R), and second derivative of the reflectance (D_2R) (Moros et al. 2010). In addition, the different methods and operators were combined, obtaining a total of 8 spectral datasets: SM, SNV, D_1R, D_1R+SNV, D_1R+SNV+SM, D_2R, D_2R+SNV, D_2R+SNV+SM. Data processing was carried out using Unscrambler indices, different numbers of wavelengths (with a limit of three predictors) were systematically combined, following the formulations presented in Table 2, and the different pre-processing methods were tested.

Table 2. Formulations used to calculate systematic combinations of wavelengths, with a limit of three bands.

VI	Equation
SR	W_1/W_2
NDI	$(W_1 - W_2)/(W_1 + W_2)$
DCNI	$[(W_1 - W_2)/(W_2 - W_3)]/(W_1 - W_3 + 0.03)$
MCARI	$[(W_1 - W_2) - 0.2(W_1 - W_3)](W_1/W_2)$
TCARI	$3[(W_1 - W_2) - 0.2(W_1 - W_3)(W_1/W_2)]$

All individual wavelengths were considered first, i.e. a total of 1,721 per each pre-processed dataset. Then, a total of 5,920,240 combinations using two wavelengths were computed for each pre-processed dataset. Finally, the same was performed with 3 wavelengths and a total of 15,265,338,840 combinations were evaluated per each pre-processed dataset. Therefore, an exhaustive computation of all possible wavelength combinations for the considered indices and pre-processed datasets has been carried out. Third, wavelengths associated with known absorption features by the N and providing the largest determination coefficients (R^2) when correlating with LNC were selected to optimize the VI and considered for deeper analysis.

The simple linear regression model (between predicted and measured LNC) and the paired t-test were used. Results were considered statistically significant when p-values were lower than 0.05. Software R (Version 3.5.1) was used for calculations (R Development Core Team, 2008).

Partial Least Squares Regression (PLSR)

PLS is a bilinear calibration method using data compression by reducing the large number of measured collinear spectral variables to non-correlated principal components (PCs), which represent the relevant structural information that can be used to predict the dependent variable. In this way, PLSR is a method often used for the retrieval of vegetation biophysical parameters using spectral data because it is an efficient method when predictors present multi-collinearity and when the number of wavelengths is larger than the number of observations (Wold et al. 2001).

Spectral prediction models were constructed based on PLSR using Unscrambler validation data for the considered model. This process is repeated 5 times, with each one of the 5 subsets used exactly once as the validation data. Finally, the 5 measures are averaged to produce an overall measure.

Results

Nutritional status of sampled and spectral data

The foliar analyses (*n=42*) indicated that LNC varied between 1.01 and 1.74%, with an average concentration of 1.3±0.2%. Figure 2 represents the LNC in each sampled olive tree and the levels of nutritional status. Most of the sampled olive trees (88.1%) presented LNC below the sufficiency threshold of 1.5%, and only 11.9% of the olive trees presented an appropriate LNC, within the range 1.5 – 2.0%. In addition, from the olive trees with LNC below sufficiency threshold, 92% presented values below the critical level of 1.4%. The spatial distribution of the sampled olive trees and nutritional levels are presented in Figure 3.

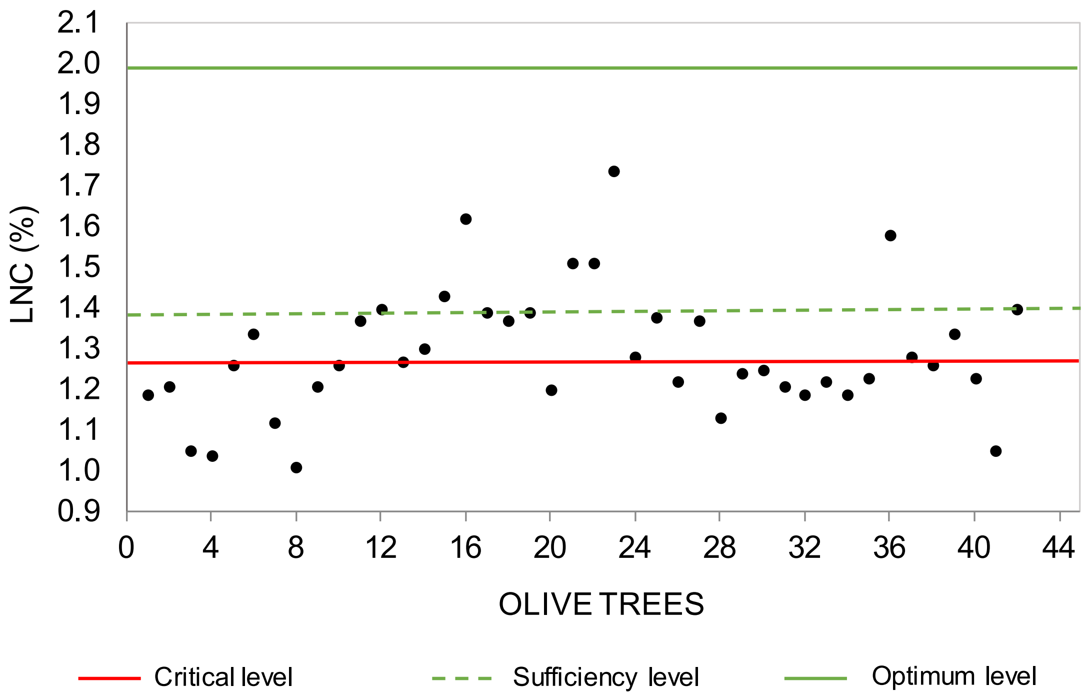

Fig. 2 Leaf N content (LNC %) of sampled olive trees (*n=42*) and nutritional levels.

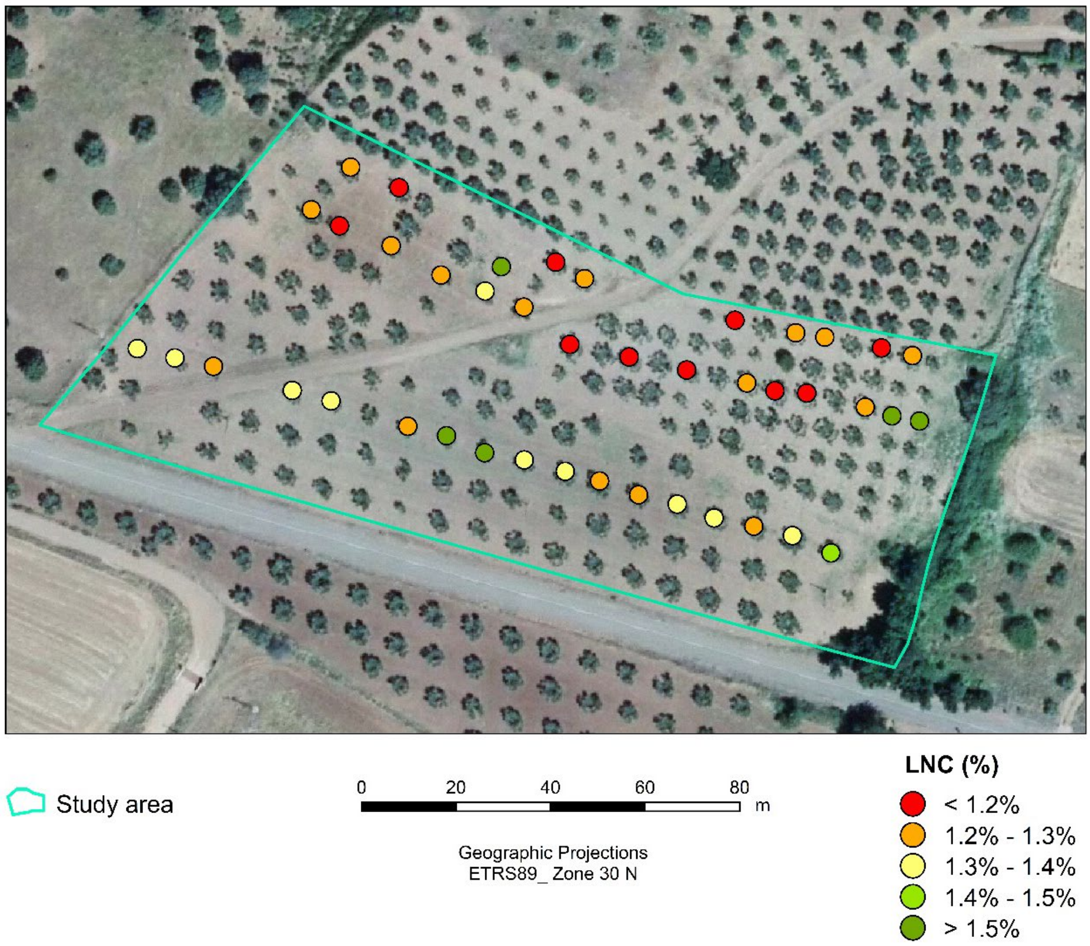

Fig. 3 Spatial distribution of the leaf N content (LNC %) of sampled olive trees ($n=42$).

Reflectance curves of the sampled olive trees leaves are plotted in Figure 4. Overall shapes were similar throughout the wavelengths measured, although the magnitude and amplitude varied, specially in the NIR plateau (750–1100 nm). Reflectance curves showed different reflectance peaks and absorption pits. In the VIS region, a reflectance peak was centered at 554 nm (green region) and two absorption pits were centered at 390 nm and 680 nm (blue and red region, respectively). The NIR plateau presented a higher reflectance than the VIS region, corresponding with a typical spectral signature of green leaves (Liang 2005). In the SWIR region, three absorption pits, centered at 1200 nm, 1450 nm and 1720 nm, and three reflectance peaks, centered at 1280 nm, 1650 nm and 2200 nm were identified.

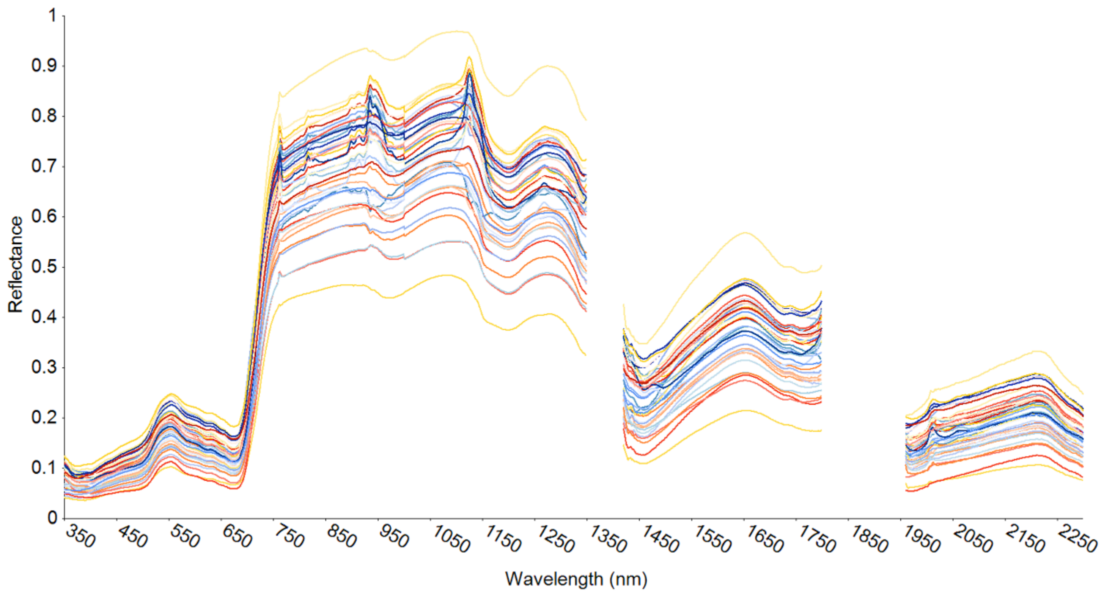

Fig. 4 Reflectance curves of olive leaves within the ranges 350–1350 nm, 1421–1800 nm and 1961–2300 nm. Each colour line represents an olive tree reflectance curve.

Nitrogen indices

Predicted LNC using VI formulated by other authors (see Table 1) presented very low determination coefficients (R^2<0.21) with measured values. No index presented statistically significant correlations with LNC (that is, all of them obtained p-values>0.05), except for the indices that used reflectance at 1450 nm: $NDI_{870,1450}$ (p-value<0.05, R^2=0.21) and $NDI_{850,1450}$ (p-value<0.05, R^2=0.12), matching one of the absorption pits in the SWIR domain (see Figure 4). In spite of the significance, the determination coefficients were very low, and LNC variation was poorly explained by $NDI_{870,1450}$ and $NDI_{850,1450}$.

Results obtained from the systematic combinations of wavelengths, using the VIs in Table 2, indicated an increase of the determination coefficients as the number of combined wavelengths increased. For example, with raw data, R^2=0.10 for one wavelength, R^2=0.36 for two-wavelength VIs and R^2=0.55 for three-wavelength VIs (Table 3). In the same way, the application of pre-processing resulted in an improvement of the determination coefficients. Specifically, from R^2=0.10 to R^2=0.46 for one wavelength, from R^2=0.36 to R^2=0.56 for two-wavelength VIs and from R^2=0.55 to R^2=0.72 for three-wavelength VIs (Table 3). Additionally, considering D_1R+SNV or D_2R+SNV, determination coefficients improved in almost all cases.

Table 3. Maximum determination coefficients (R^2) for leaf N content and individual wavelengths (1W), VIs based on combinations of two wavelengths (2W, i.e. SR and NDI) and on combinations of three wavelengths (3W, i.e. DCNI, MCARI and TCARI) by considering several pre-processed datasets.

Pre-processing	1W	2W	3W
Raw data	0.10	0.36	0.55
SM	0.27	0.54	0.68

SNV	0.30	0.37	0.59
D_1R	0.27	0.54	0.67
D_1R+SNV	0.40	0.54	0.66
D_1R+SNV+SM	0.40	0.54	0.69
D_2R	0.31	0.56	0.69
D_2R+SNV	0.46	0.56	0.72
D_2R+SNV+SM	0.39	0.44	0.69

Table 4 summarizes the results obtained from the wavelength combination process to determine the most suitable index to estimate the LNC in olive trees, considering biophysical properties of the N reflectance curve. As can be observed, the best results were obtained by using the second derivative as pre-processing and, mainly, when the second derivative was combined with SNV.

Table 4. Maximum determination coefficients (R^2) between leaf N content and vegetation indices (VI). *W* considers only a wavelength and no index is defined. *Pp* indicates the pre-processing applied to spectra data, where *A* is D_1R, *B* is D_2R, and *C* is D_2R+SNV. *Selected W* are the wavelength combinations that presented maximum R^2 to each VI, considering biophysical properties of the nitrogen reflectance curve. *Index* is the formula of the specific VI in the present paper. The letters of the *Selected W* columns indicate the region of the spectrum where the wavelengths are located: *b* = blue, *r* = red and *sw* = SWIR region. * indicates repeated wavelengths in different indices.

VI	R^2	R^2_{cv}	Pp	Selected W (nm)	Index
W	0.46	-	*C*	646^{r}	R_{646}
MCARI	0.53	0.50	*A*	1659^{sw}, 1749^{sw}, 1128^{sw}	$((R_{1659}-R_{1749})-0.2(R_{1659}-R_{1128}))(R_{1659}/R_{1749})$
SR	0.56	0.53	*B*	1615^{sw}, 648^{r*}	R_{1615}/R_{648}
NDI	0.56	0.55	*C*	397^{b}, 648^{r*}	$(R_{397}-R_{648})/(R_{397}+R_{648})$

TCARI	0.64	0.63	*C*	1685^{sw}, 412^{b}, 2209^{sw}	$3((R_{1685}-R_{412})-0.2(R_{1685}-R_{2209})(R_{1685}/R_{412}))$
DCNI	0.72	0.71	*C*	395^{b}, 652^{r}, 1275^{sw}	$((R_{395}-R_{652})/(R_{652}-R_{1275}))/(R_{395}-R_{1275}+0.03)$

In Figure 5, scatterplots for VI values and measured LNC with maximum determination coefficients were plotted together with the regression line. DCNI yielded the best goodness-of-fit with LNC (R^2=0.72; R^2_{cv}=0.71), combining the following wavelengths: 395 nm, matching with the absorption pit in the blue region; 652 nm, corresponding to the absorption pit in the red region; and 1275 nm, matching with one of the reflectance peaks in the SWIR region. The second best result was obtained with TCARI. It predicted LNC with R^2=0.64 (R^2_{cv}=0.63), combining the wavelengths: 412 nm, the absorption pit in the blue region; with reflectance peaks at 1685 nm and 2209 nm in the SWIR region.

Determination coefficients obtained using MCARI, SR, and NDI indices were lower than the ones obtained with TCARI and DCNI. SR predicted LNC with R^2=0.56 (R^2_{cv}=0.53) and the selected wavelengths were: 648 nm, the absorption pit in the red region; and 1615 nm, a wavelength close to the reflectance peak at 1650 nm in the SWIR domain. NDI estimated LNC with a similar determination coefficient (R^2=0.56) but with a slightly greater R^2_{cv}=0.55. The combined wavelengths were: 397 nm, the absorption pit in the blue region; and 648 nm, the absorption pit in the red domain. MCARI predicted LNC with R^2=0.53 (R^2_{cv}=0.50), combining three bands in the SWIR region: an absorbance band at 1749 nm with two reflectance bands at 1128 nm and 1659 nm.

The lowest coefficient was obtained from the relationship between LNC and the 646 nm wavelength, close to the absorption pit in the red region, with an R^2=0.46. This case considers only a wavelength and no index is defined.

Only the 648 nm (red region) wavelength correlated with LNC in two indices. The 395 nm, 397 nm and 412 nm wavelengths selected in DCNI, NDI and TCARI indices, respectively, are located around of the chlorophyll$_a$ absorption pit at 395 nm (Curran et al. 2001). The wavelengths 646 nm (W) and 652 nm (DCNI) are very close to 648 nm chlorophyll$_b$ absorbance pit. Wavelengths from the SWIR region were selected to calculate the SR, TCARI, MCARI and DCNI indices. No index combining wavelengths from the NIR region presented a high determination coefficient. These results point to combinations using blue, red and SWIR wavelengths as the most suitable to estimate LNC in the studied olive trees using hyperspectral data.

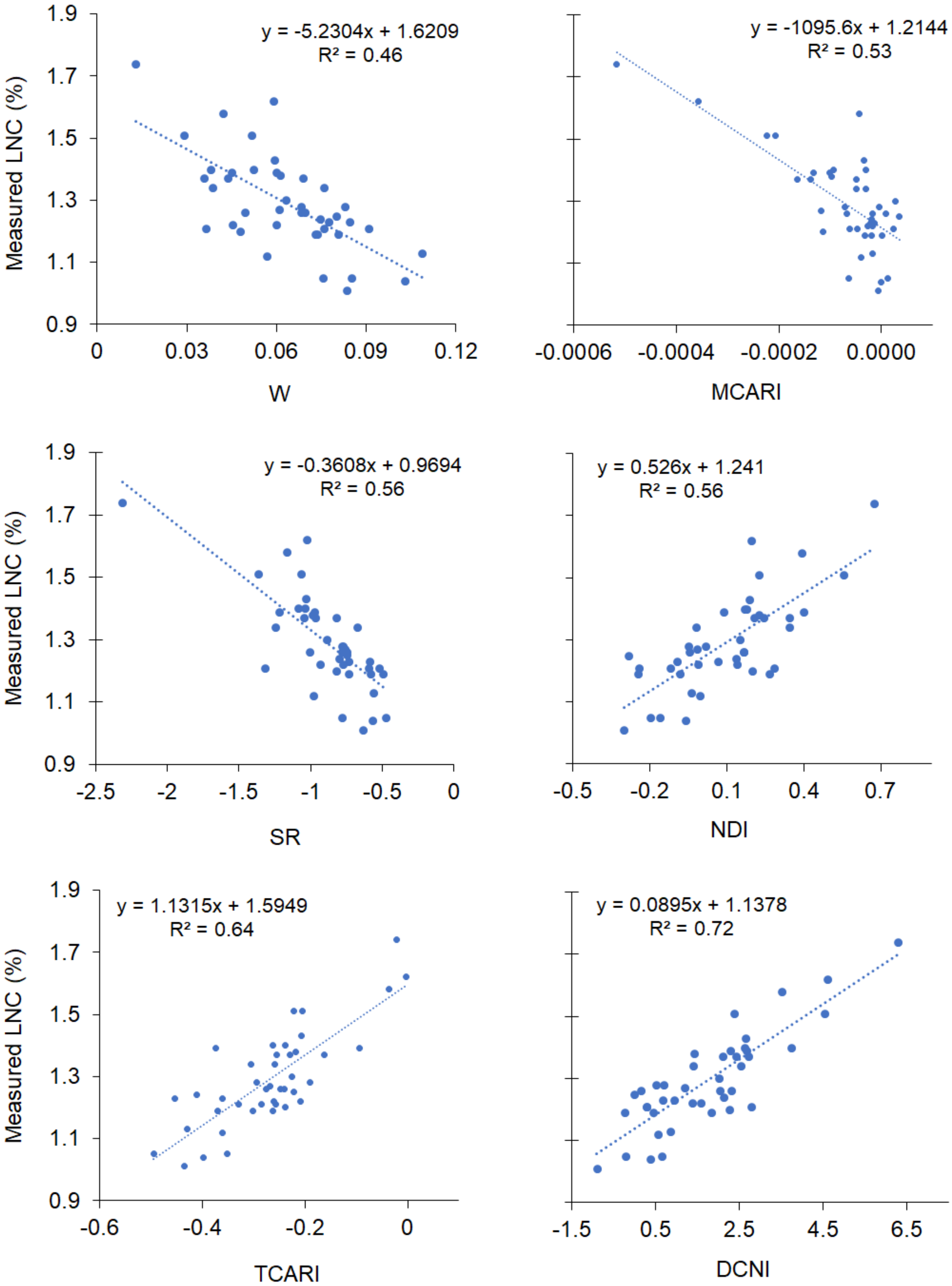

Fig. 5 Scatterplots and regression line for the selected VI and measured leaf N content (%).

Estimated N values obtained by using calibration and 5-fold cross-validation methods were considered to analyze the closeness of the estimations provided with those methods. The linear relationships between N predictions obtained with both methods are analyzed for all cases in Table 5, providing similar R^2 and R^2_{cv} in all these cases. The determination coefficients (R^2_{c-cv}) were close to 1.0 and the differences between LNC estimated with calibration and 5-fold cross-validation linear regressions were not statistically significant (p-values>0.05). RMSE and $RMSE_{cv}$ were similar for each VI and the lowest value was produced by DCNI (Table 5).

Table 5. Comparison between calibration (R^2) and validation results (R^2_{cv}). R^2_{c-cv} is the determination coefficient between predicted LNC with calibration and predicted LNC with 5-fold cross-validation; *RMSE* is the root mean square error of calibration; $RMSE_{cv}$ is the root mean square error of validation.

VI	R^2	R^2_{cv}	R^2_{c-cv}	RMSE	$RMSE_{cv}$	T-test (p-value)
MCARI	0.53	0.50	0.994	0.10	0.10	0.73
SR	0.56	0.53	0.995	0.10	0.10	0.81
NDI	0.56	0.55	0.997	0.10	0.10	0.74
TCARI	0.64	0.63	0.997	0.09	0.09	0.39
DCNI	0.72	0.71	0.997	0.08	0.08	0.44

Multi-dimensional approach based on PLSR

PLSR searches the sensitive information from spectra data and then uses the 5-fold cross-validation procedure to calculate the calibration PLSR model. Different PLSR models were developed using raw and pre-processed data (Table 6).

Table 6. PLSR results for each pre-processed dataset. CV and CV_{cv} are the variation coefficients of calibration and cross-validation models, respectively.

Model/Pre-processing	R^2	R^2_{cv}	LV	RMSE	$RMSE_{cv}$	CV	CV_{cv}	T-test (p-value)
Raw data	0.35	0.23	3	0.12	0.14	0.07	0.07	0.35
SM	0.35	0.23	3	0.12	0.14	0.07	0.07	0.35
SNV	0.38	0.23	3	0.12	0.14	0.07	0.08	0.49
D_1R	0.79	0.33	6	0.07	0.13	0.11	0.10	0.55
D_1R+SNV	0.83	0.52	5	0.06	0.11	0.11	0.10	0.72
D_1R+SNV+SM	0.82	0.52	5	0.07	0.11	0.11	0.11	0.41
D_2R	0.98	0.56	7	0.02	0.10	0.12	0.09	0.60
D_2R+SNV	0.95	0.58	5	0.03	0.10	0.12	0.09	0.87
D_2R+SNV+SM	0.93	0.57	5	0.04	0.10	0.12	0.09	0.99

Analysing the results presented in Table 6, determination coefficients ranged between 0.35 and 0.98 for the calibration models and between 0.23 and 0.58 for the cross-validation models. D_2R model presented the largest R^2=0.98, followed by D_2R+SNV model (R^2=0.95) and by D_2R+SNV+SM model (R^2=0.93) (see Figure 6). The calibration and validation RMSE were relatively small in those cases (≤0.04 N% in the calibration and 0.10 N% in the validation) and the differences between both were lower than 20% (0.20), indicating that the number of LV selected for each pre-processing was suitable, without overfitting the models (Shao et al. 2007). The difference between R^2 and R^2_{cv} was large: 0.42, 0.37 and 0.36, respectively. However, the differences between average LNC estimated with calibration and average LNC estimated with 5-fold cross-validation obtained with PLSR were not statistically significant (p-values>0.05) and the differences between the variation coefficients (CV) were ≤0.03 in all cases. These results indicated that, despite the difference between the determination coefficients of the calibrated and validated models, the LNC estimates were similar.

Predicted LNC using the calibrated and validated PLSR models are plotted in Figure 6 for each dataset, raw and pre-processed data. It can be observed that second derivative pre-processing presented better fitting between measured and predicted LNC than the other pre-processing.

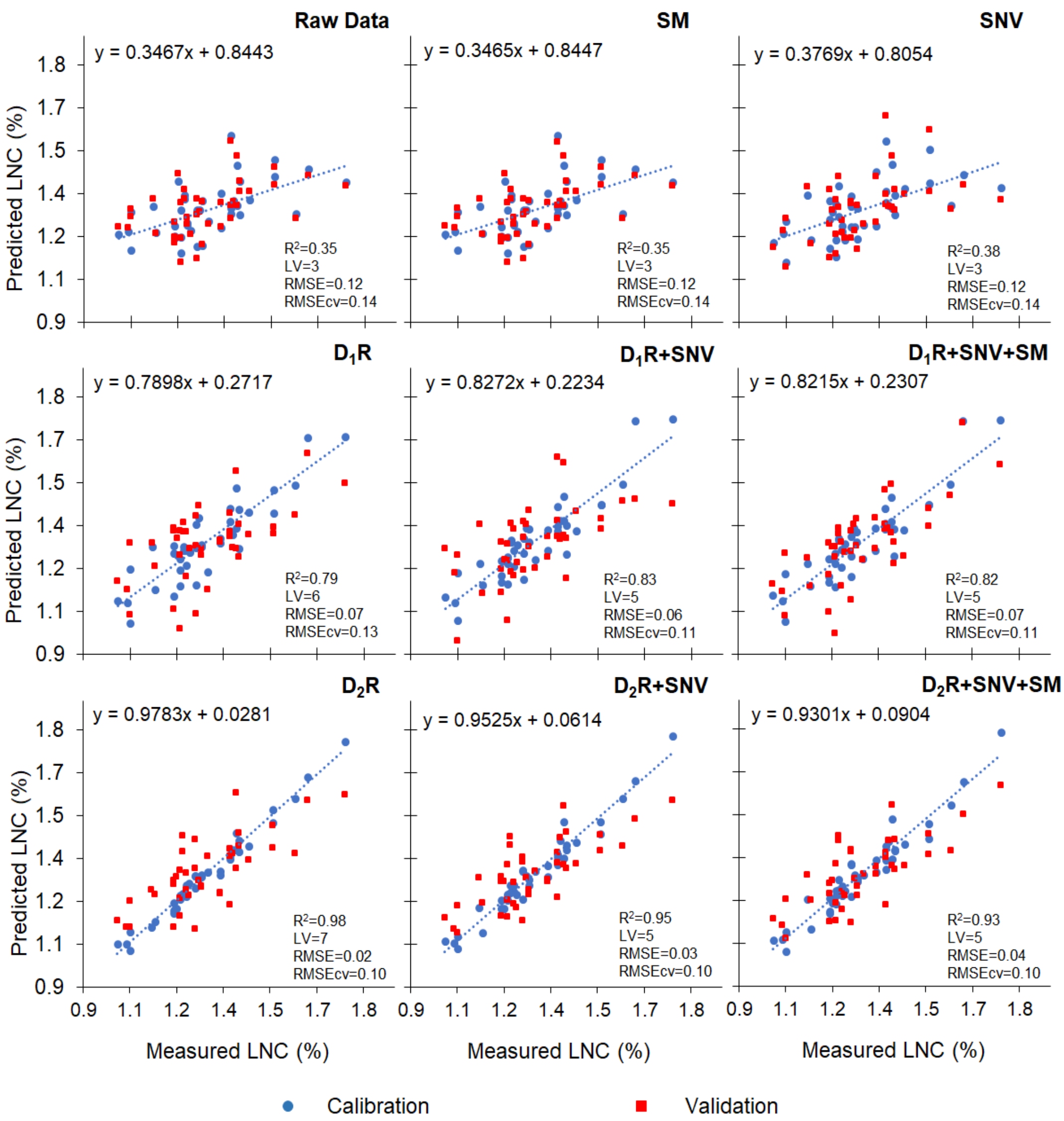

Fig. 6 Measured vs. predicted leaf N content (*LNC %*) using PLSR and different pre-processing methods.

Discussion

VI and PLSR models were used to relate LNC and reflectance data over the full spectrum (350–2500 nm) using a high precision spectro-radiometer. It has been shown that these methods produced high coefficients of determination, particularly when the reflectance spectra were expressed as the second derivative and combined with SNV or SM pre-processing. The wavelengths most sensitive to N variation used to calculate VI were selected from the VIS and SWIR spectral regions, which relate to Chl and N features. PLSR models yielded a higher accuracy (R^2) than VI, although the uncertainties associated with the noise of the hyperspectral data were higher.

Ferwerda and Skidmore (2007) and Rotbart et al. (2013) focused on a similar topic. The novelty of this paper with respect to these studies is the method used to determine the band combinations more suitable to estimate LNC in olive trees. Ferwerda and Skidmore (2007) selected the bands with optimal fits to the N content through a stepwise regression routine, without taking into account the formulation of any VI, and requiring four predictors. Besides, they only analyzed two pre-processing methods (Savitzky–Golay smoothing and continuum-removed spectra). In the present paper, the wavelengths of all the spectral regions are combined, taking into account their biophysical basis and the formulation of the VIs more suitable to estimate LNC according to other authors: SR, NDI, DCNI, MCARI and TCARI (see Tables 1 and 2). The number of predictors is set to three, simplifying the model respect to the one developed by Ferwerda and Skidmore (2007), that used four, at the same time that better results are obtained in this work. Moreover, this paper analyzes eight pre-processing methods, instead of only two. Furthermore, in this paper, the results are compared with those obtained by other authors and those obtained with the PLSR method.

With respect to the study by Rotbart et al. (2013), its objective was not to determine suitable bands to estimate LNC. Besides, Rotbart et al. (2013) only used five pre-processing methods to develop the PLSR models, instead of the eight analyzed in this paper. Even more, Rotbart et al. (2013) used a lower spectral range (450 nm–1000 nm and 1100–1700 nm) due to the use of three different spectrometers (USB-2000, LIGA and Luminar-5030), whereas in this paper all the spectral range (350 nm–2500 nm) has been used. Finally, Rotbart et al. (2013) determined N concentration in olive leaves by using spectral reflectance of the VIS-NIR range under laboratory conditions and not in the field such as is the case of this paper.

Nutritional status of olive trees and spectral data

N status of sampled olive trees was deficient considering a critical and sufficiency threshold of LNC of 1.4% and 1.5%, respectively (Barranco et al. 1997; Fernández-Escobar et al. 1999). Fernández-Escobar et al. (2009) studied the long-term effect of N fertilization on olive trees. The results indicated that yield and growth were maintained in trees non-fertilized and with LNC within the range 1.4% and 2.0% during several consecutive years. In this sense, Fernández-Escobar et al. (2009) recommended that the leaf critical threshold of LNC in olive trees must be revised and is likely to lie between 1.2% and 1.3%. Regarding results presented in the current study, if 1.3 % or 1.2% are considered as critical limits of LNC, then, 59.5% or 21.4%, respectively, of the sampled trees had a deficient N status, instead of the 81.0% obtained with a threshold of 1.4%. All these reports suggest that fertilization management depends on the LNC critical limit established and, therefore, spatial and temporal analyses are necessary to carry out appropriate fertilization planning to improve olive orchards production. According to López-Granados et al. (2004), Gómez-Casero et al. (2007) and Fernández-Escobar et al.

(2009), N fertilization should be applied only to olive trees with an N value that does not exceed the considered threshold, because they reported that, in the case of olive orchards, annual maintenance application of N is not necessary to improve yield and tree growth.

Reflectance curves of the leaves of the sampled olive trees presented a similar shape throughout the spectral range. The reflectance curves obtained within the range 400 nm–900 nm were similar to that obtained by Zarco-Tejada et al. (2004) and Gómez-Casero et al. (2007) for olive orchards without fertilization treatments, presenting absorption pits and reflectance peaks at the same wavelength ranges. Absorption pits were identified at 500±10 nm and 680±10 nm, whereas reflectance peaks were identified at 550±10 nm and 760±20 nm. Wavelength absorption at 680 nm could be related to the absorption feature of Chl_a at 660 nm (Curran 1989). Chl is the main pigment responsible for the properties of the reflectance and transmittance of radiation in the VIS region and has a close correlation with the N content in the leaves of plants (Haboudane et al. 2002). N shortage will reduce leaf Chl content and consequently the reflectance in the VIS region will increase (Daughtry et al. 2000). Therefore, the reflectance peak at 550 nm could indicate a low Chl content and consequently a low LNC, matching the results obtained considering 1.4% as a critical threshold.

Systematic wavelength combinations and pre-processing

The systematic combination of wavelengths from all spectral ranges showed that the determination coefficients increased as the number of combined wavelengths increased. The application of pre-processing decreased considerably the noise generated by the measurements obtained from the spectro-radiometer, also improving the results. As in the case of the PLSR models, second derivative combined with SNV and/or SM produced the best determination coefficients. These results agree with the ones from other publications which demonstrated the significant and positive effects of using the whole spectral range and pre-processing to estimate LNC in olive trees. Ferwerda and Skidmore (2007) carried out a stepwise regression, with a limit of four predictors, between leaf chemical composition (N, K, Na, P, Ca and Mg) of four woody plant species (olive tree, mopane, heather and willow) using different datasets: raw reflectance spectra, derivative spectra and band depth. The model based on derivative spectra offered the highest prediction accuracy. Rotbart et al. (2013) obtained the most robust prediction models applying the first derivative to the logarithm of the reflectance reciprocal [D_1(log(1/R))] with the PLSR method.

VI to estimate LNC

Reflectance-based indices formulated using wavelengths associated with known absorption features showed lower determination coefficients with LNC than indices derived from the systematic wavelength combinations carried out in this study. The wavelengths combined in the VIs with higher determination coefficients were located mainly in the VIS region (350 nm–750 nm) and in the SWIR region (1100 nm–2500 nm).

In the present study, wavelengths located in the blue range (395 nm, 397 nm and 412 nm) were selected. They could be directly related to the absorption pits of the Chl_a at 430 nm (Curran 1989) and indirectly related to the N content (Haboudane et al. 2002). Organic components (e.g., cellulose, lignin, proteins, oil, sugar or starch) absorb radiation strongly in the blue spectral domain as a result of stretching and bending vibrations of the strong molecular bonds between hydrogen atoms and the atoms of carbon, nitrogen and oxygen (Osborne 2000). Other studies showed that blue bands were sensitive to estimate LNC at later growing stage of the plant (Cohen and Alchanatis 2018). Therefore, 395 nm, 397 nm and 412 nm wavelengths also could be related to the absorption region of these organic components in the blue region and their reactions with the N content. This could justify bio-physically why these wavelengths were selected to optimize the VIs.

Wavelengths at the red domain of the VIS region (e.g., 646 nm, 648 nm and 652 nm) appeared in several of the proposed VIs. These wavelenghs could be related to the absorption pits of the Chl_b and Chl_a at 640 nm and 660 nm (Curran 1989). This link is related to the strong influence of N on Chl production and functioning. Therefore, N could be indirectly estimated using wavelengths within the red domain (Haboudane et al. 2002). The 750 nm wavelength is in the limit between red (VIS) and red edge (NIR) domain. Several authors selected this wavelength to estimate Chl_{ab} content from hyperspectral reflectance (Zarco-Tejada et al. 2001; Wu et al. 2008). Therefore, it also could be related to N content.

The 1128 nm, 1275 nm, 1615 nm, 1659 nm, 1685 nm, 1749 nm and 2209 nm wavelengths centered at SWIR region and selected to optimize VIs could be related to N, protein, oil, lignin, sugar, cellulose, starch and water content (Curran 1989). The spectral range 1500 nm–1600 nm is a region dominated by absorption features due to N-H bond stretching located at 1510 nm (Curran 1989) and this bond is related to the amount of N present in the protein (Ferwerda 2005). However, the reflectance within the SWIR region is sensitive to foliar water content. Ceccato et al. (2002) and Ferwerda (2005) identified that the indices which included the 1770 nm absorption feature were least affected by foliar water content.

Authors like Zarco-Tejada et al. (2004), Ferwerda (2005), Ferwerda and Skidmore (2007) or Gómez-Casero et al. (2007) selected similar wavelengths to the ones considered in the present study to estimate N content in olive trees. Zarco-Tejada et al. (2004) combined 750 nm and 710 nm wavelengths in a simple ratio formulation and used the 550 nm, 670 nm and 700 nm wavelengths to calculate MCARI (Daughtry et al. 2000) and TCARI (Haboudane et al. 2002) indices. Ferwerda (2005) obtained their best determination coefficients combining wavelengths from the VIS (627 nm - 640 nm and 690 nm - 700 nm) and from the SWIR (1516 nm - 1580 nm, 1770 nm, 1803 nm and 2196 nm) regions in the NDI formulation. Ferwerda and Skidmore (2007) obtained their best determination coefficients combining the 733 nm, 1203 nm, 1792 nm and 955 nm wavelengths through a stepwise regression routine and limiting the number of predictors to four. Gómez-Casero et al. (2007) selected different wavelengths within the VIS-NIR region depending on the type of fertilization applied. For soil application, selected wavelengths were within the blue (400 nm, 420 nm and 470 nm), red (610 nm, 630 nm and 720 nm) and NIR domain (750 nm–900 nm). For soil and foliar application, selected wavelengths were within the blue (420 nm and 490 nm), green (510 nm), red (690 nm and 710 nm) and NIR domain (890 nm).

In the present study, the best determination coefficient and the lowest RMSE/$RMSE_{cv}$ were obtained by using the DCNI index formulation (R^2=0.72). Camino et al. (2018) estimated LNC for durum and bread wheat using this index with the formulation developed by Chen et al. (2010), and reported a determination coefficient of R^2=0.56, with wavelengths within the red domain (670 nm, 700 nm and 720 nm). In this work, TCARI and MCARI indices reported R^2=0.64 and R^2=0.53, respectively. Zarco-Tejada et al. (2004) applied these indices to estimate LNC in olive trees using wavelengths within the VIS region (550 nm, 670 nm and 700 nm), according to the formulation of Daughtry et al. (2000) and Haboudane et al. (2002). The determination coefficients were lower in the case of the TCARI index (R^2= 0.60) and higher in the case of MCARI index (R^2=0.64). SR and NDI indices also presented similar determination coefficients (R^2=0.56 in both cases), although the selected wavelengths were different. SR index combined VIS and SWIR regions (648 nm and 1615 nm) and NDI combined wavelengths at the blue and red domain (397 nm and 648 nm). Therefore, based on the whole previous discussion and indices, and in agreement with Herrmann et al. (2010), the SWIR region is more sensitive to N content than the VNIR region. Additionally, indices based on combination between SWIR and VIS wavelengths are better predictors of LNC in olive trees than the ones that only use the VIS region.

PLSR models

The largest determination coefficient was obtained with D_2R dataset and 7 LV (R^2=0.98). D_2R+SNV+SM produced lower determination coefficient (R^2=0.93), but the difference between calibration and validation coefficients, as well as the number of LV, was lower. These results indicated that this model is more robust, which presented a RMSE of 0.04% and a $RMSE_{cv}$ of 0.10%. Rotbart et al. (2013) estimated LNC in olive trees using the whole spectral range and PLSR models, applying different pre-processing. The results were similar to those reported in this paper, i.e. the largest determination coefficient was obtained with pre-processing data (R^2=0.91), with a number of 7 LV, and RMSE and $RMSE_{cv}$ of 0.05% and 0.07%, respectively.

In agreement with other authors that reported that PLSR methods have greater potential than spectral indices for deriving N content in crops such as winter wheat (Hansen and Schjoerring 2003; Li et al. 2014) or maize (Quemada et al. 2014), in this study PLSR models (D_1R, D_1R+SNV, D_1R+SNV+SM, D_2R, D_2R+SNV and D_2R+SNV+SM) produced higher determination coefficients (R^2) than VI. However, the results obtained with the PLSR method show higher difference between R^2 and R^2_{cv} than the VI method, indicating a possible overfitting of the models. But the differences between RMSE and $RMSE_{cv}$ were lower than 20% (0.20) in all cases. Therefore, according to Shao et al. (2007), this result indicates a suitable number of LV selected for each pre-processing and, accordingly, the models did not present overfitting.

Taking into account the uncertainties presented by both methods, it is difficult to determine which method yielded better fits between the estimation and the LNC measured. R^2_{cv} is an indicator of the prediction capacity of the model. Therefore, since DCNI and TCARI vegetation indices presented higher R^2_{cv} and lower $RMSE_{cv}$ than PLSR models in all pre-processed datasets, in the case of this study, VI are more suitable than PLSR models to estimate LNC in olive trees without fertilization treatments.

Conclusions

The present study demonstrated that hyperspectral data are useful to estimate leaf N content in olive trees without fertilization treatments. The VI and PLSR models considered by using the full spectrum (350–2500 nm) produced larger determination coefficients than the ones from spectral indices formulated using wavelengths associated with known absorption features. Applying pre-processing to the spectral data, the noise generated by the measurement of the spectro-radiometer was reduced and the correlations between leaf N content and reflectance data were improved, particularly when the reflectance spectra were expressed as the second derivative and combined with SM and/or SNV pre-processing. VIs that produced the largest determination coefficients were DCNI and TCARI. The wavelengths most sensitive to N variation used to define VIs were selected from the VIS and SWIR spectral regions, which relate to chlorophyll a+b and N absorption features. Therefore, implementing the SWIR region to estimate leaf N content in olive trees improved predictions. PLSR models yielded higher accuracy than VI, although the uncertainties associated with the noise of the hyperspectral data were higher. These methods would allow accurate fertilization plans depending on the olive tree requirements. Further research is necessary to gain knowledge about the temporal variation of the leaf N content in olive trees to offer to the farmers accurate information about the nutritional status of the plants at each phenological stage.

Acknowledgments

The authors wish to thank the GeoAmbiental Research Group of the University of Extremadura (Spain) for providing the spectro-radiometer used in the spectral data collection. In addition, the authors would like to thank the three anonymous referees for helping to improve both the readability and the content of this paper.

This research has been financially supported by Junta de Extremadura, Spain (projects GR18090 and GR18108), European Union (European Regional Development Funds), and NotAnts S.L.U. through the project AA-16-0091-1.

References

Akdemir, B., Saglam, C., Belliturk, K., Makaraci, A., Urusan, A., & Atar, E. (2018). Effect of spatial variability on fertiliser requirement of olive orchard cultivated for oil production. *Journal of Environmental Protection and Ecology, 19*(1), 319-329.

Alchanatis, V., Cohen, Y., Levi, A., Ostrovsky, V., Schmilovitch, Z., Dag, A., et al. (2009). Estimating olive leaf mineral content using spectral reflectance. In U. Yermiyahu, A. Dag, A. Ben-Gal (Eds.), *International Symposium for Olive*

Irrigation and Oil Quality (pp. O23). Leuven, Belgium: International Society for Horticultural Science.

Atzberger, C., Guérif, M., Baret, F., & Werner, W. (2010). Comparative analysis of three chemometric techniques for the spectroradiometric assessment of canopy chlorophyll content in winter wheat. *Computers and Electronics in Agriculture, 73*(2), 165-173.

Barnes, R. J., Dhanoa, M. S., & Lister, S. J. (1989). Standard Normal Variate Transformation and De-Trending of Near-Infrared Diffuse Reflectance Spectra. *Applied Spectroscopy, 43*(5), 772-777.

Barnes, R. J., Dhanoa, M. S., & Lister, S. J. (1993). Correction to the Description of Standard Normal Variate (SNV) and De-Trend (DT) Transformations in Practical Spectroscopy with Applications in Food and Beverage Analysis—2nd Edition. *Journal of Near Infrared Spectroscopy, 1*(3), 185-186.

Barranco, D., Fernández-Escobar, R., & Rallo, L. (1997). *El cultivo del olivo (Olive cultivation*) (1ª ed.). Madrid, Spain: Mundi-Prensa Libros.

Bi, Y., Yuan, K., Xiao, W., Wu, J., Shi, C., Xia, J., et al. (2016). A local pre-processing method for near-infrared spectra, combined with spectral segmentation and standard normal variate transformation. *Analytica Chimica Acta, 909*, 30-40.

Camino, C., González-Dugo, V., Hernández, P., Sillero, J., & Zarco-Tejada, P. J. (2018). Improved nitrogen retrievals with airborne-derived fluorescence and plant traits quantified from VNIR-SWIR hyperspectral imagery in the context of precision agriculture. *International Journal of Applied Earth Observation and Geoinformation, 70*, 105-117.

CAMO (2011). CAMO Software: Unscrambler. Online resource: https://www.camo.com. Accessed 15 May 2020.

Ceccato, P., Gobron, N., Flasse, S., Pinty, B., & Tarantola, S. (2002). Designing a spectral index to estimate vegetation water content from remote sensing data: Part 1: Theoretical approach. *Remote Sensing of Environment, 82*(2-3), 188-197.

Chen, P., Haboudane, D., Tremblay, N., Wang, J., Vigneault, P., & Li, B. (2010). New spectral indicator assessing the efficiency of crop nitrogen treatment in corn and wheat. *Remote Sensing of Environment, 114*(9), 1987-1997.

Cilia, C., Panigada, C., Rossini, M., Meroni, M., Busetto, L., Amaducci, S., et al. (2014). Nitrogen status assessment for variable rate fertilization in maize through hyperspectral imagery. *Remote Sensing, 6*(7), 6549-6565.

Clevers, J. G., Kooistra, L., & van den Brande, M. M. (2017). Using Sentinel-2 data for retrieving LAI and leaf and canopy chlorophyll content of a potato crop. *Remote Sensing, 9*(5), 405.

Cohen, Y. & Alchanatis. V. (2018). Spectral and Spatial Methods for Hyperspectral and Thermal Image-Analysis to Estimate Biophysical and Biochemical Properties of Agricultural Crops. In: P.S. Thenkabail, G.J. Lyon, A. Huete (Eds.), *Hyperspectral Remote Sensing of Vegetation: Biophysical and Biochemical Characterization and Plant Species Studies* (pp. 73–102). Boca Raton, USA: CRC Press, Taylor and Francis group.

Curran, P. J. (1989). Remote sensing of foliar chemistry. *Remote Sensing of Environment, 30*(3), 271-278.

Curran, P. J., Dungan, J. L., & Peterson, D. L. (2001). Estimating the foliar biochemical concentration of leaves with reflectance spectrometry: testing the Kokaly and Clark methodologies. *Remote Sensing of Environment, 76*(3), 349-359.

Daughtry, C., Walthall, C., Kim, M., De Colstoun, E. B., & McMurtrey Iii, J. (2000). Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. *Remote Sensing of Environment, 74*(2), 229-239.

EUROSTAT, European Statistical System (2017): Agriculture data. Online resource: http://ec.europa.eu/eurostat/data/database. Accessed 15 May 2020.

FAOSTAT, Food and Agriculture Organization of the United Nations (2017): Food and agriculture data. Online resource: http://www.fao.org/faostat/en/#data. Accessed 15 May 2020.

Fernández-Escobar, R., Marin, L., Sánchez-Zamora, M., García-Novelo, J., Molina-Soria, C., & Parra, M. (2009). Long-term effects of N fertilization on cropping and growth of olive trees and on N accumulation in soil profile. *European Journal of Agronomy, 31*(4), 223-232.

Fernández-Escobar, R., Moreno, R., & García-Creus, M. (1999). Seasonal changes of mineral nutrients in olive leaves during the alternate-bearing cycle. *Scientia Horticulturae, 82*(1-2), 25-45.

Ferwerda, J. G. (2005). *Charting the quality of forage: measuring and mapping the variation of chemical components in foliage with hyperspectral remote sensing*. Wageningen University, Netherlands, PhD Thesis.

Ferwerda, J. G., & Skidmore, A. K. (2007). Can nutrient status of four woody plant species be predicted using field spectrometry? *ISPRS Journal of Photogrammetry and Remote Sensing, 62*(6), 406-414.

Gómez-Casero, M. T., López-Granados, F., Peña-Barragán, J. M., Jurado-Expósito, M., García-Torres, L., & Fernández-Escobar, R. (2007). Assessing nitrogen and potassium deficiencies in olive orchards through discriminant analysis of hyperspectral data. *Journal of the American Society for Horticultural Science, 132*(5), 611-618.

Haboudane, D., Miller, J. R., Tremblay, N., Zarco-Tejada, P. J., & Dextraze, L. (2002). Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sensing of Environment, 81*(2–3), 416-426.

Hansen, P. M., & Schjoerring, J. K. (2003). Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. *Remote Sensing of Environment, 86*(4), 542-553.

Herrmann, I., Karnieli, A., Bonfil, D. J., Cohen, Y., & Alchanatis, V. (2010). SWIR-based spectral indices for assessing nitrogen content in potato fields. *International Journal of Remote Sensing, 31*(19), 5127-5143.

Hruschka, W. R. (1987). Data analysis: Wavelength selection methods. In P. C. Williams, & K. H. Norris (Eds.), *Near Infrared Technology in Agricultural and Food Industries* (pp. 35-55). St. Paul, MN, USA: American association of Cereal Chemists, Inc.

Im, J., & Jensen, J. R. (2008). Hyperspectral remote sensing of vegetation. *Geography Compass, 2*(6), 1943-1961.

Li, F., Mistele, B., Hu, Y., Chen, X., & Schmidhalter, U. (2014). Reflectance estimation of canopy nitrogen content in winter wheat using optimised hyperspectral spectral indices and partial least squares regression. *European Journal of Agronomy, 52*, 198-209.

Liang, S. (2005). *Quantitative remote sensing of land surfaces* (Vol. 30). New York, USA: John Wiley & Sons.

López-Granados, F., Jurado-Expósito, M., Alamo, S., & García-Torres, L. (2004). Leaf nutrient spatial variability and site-specific fertilization maps within olive (Olea europapaea L.) orchards. *European Journal of Agronomy, 21*(2), 209-222.

Miphokasap, P., & Wannasiri, W. (2018). Estimations of Nitrogen Concentration in Sugarcane Using Hyperspectral Imagery. *Sustainability, 10*(4), 1266.

Moros, J., Garrigues, S., & de la Guardia, M. (2010). Vibrational spectroscopy provides a green tool for multi-component analysis. *Trends in Analytical Chemistry*, 29(7), 578-591.

Muik, B., Lendl, B., Molina-Díaz, A., Pérez-Villarejo, L., & Ayora-Cañada, M. J. (2004). Determination of oil and water content in olive pomace using near infrared and Raman spectrometry. A comparative study. *Analytical and Bioanalytical Chemistry, 379*(1), 35-41.

Ninyerola, M., Roure, J. M., & Pons X. (2005). *Atlas climático digital de la Península Ibérica: metodología y aplicaciones en bioclimatología y geobotánica (Digital climate Atlas of the Iberian Peninsula: methodology and applications in* *bioclimatology and geobotany). Barcelona, Spain:* Universidad Autónoma de Barcelona.

Osborne, B. G. (2006). Near-infrared spectroscopy in food analysis. In R.A. Meyers (Ed.), *Encyclopedia of Analytical Chemistry: Applications, Theory and Instrumentation, 5*, (pp. 1-14). Chichester, UK: John Wiley & Sons.

Pimstein, A., Eitel, J. U., Long, D. S., Mufradi, I., Karnieli, A., & Bonfil, D. J. (2009). A spectral index to monitor the head-emergence of wheat in semi-arid conditions. *Field crops research, 111*(3), 218-225.

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 crops research, 121*(1), 125-135.

Puckett, L. J., Tesoriero, A. J., & Dubrovsky, N. M. (2011). Nitrogen Contamination of Surficial Aquifers—A Growing Legacy. *Environmental Science & Technology, 45*(3), 839-844.

Quemada, M., Gabriel, J. L., & Zarco-Tejada, P. (2014). Airborne hyperspectral images and ground-level optical sensors as assessment tools for maize nitrogen fertilization. *Remote Sensing, 6*(4), 2940-2962.

R Development Core Team (2008). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Online resource: http://www.r-project.org. Accessed 15 May 2020.

Rotbart, N., Schmilovitch, Z., Cohen, Y., Alchanatis, V., Erel, R., Ignat, T., et al. (2013). Estimating olive leaf nitrogen concentration using visible and near-infrared spectral reflectance. *Biosystems Engineering, 114*(4), 426-434.

Savitzky, A., & Golay, M. J. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry, 36*(8), 1627-1639.

Schulze, E. D., De Vries, W., Hauhs, M., Rosen, K., Rasmussen, L., Tamm, C.-O., et al. (1989). Critical loads for nitrogen deposition on forest ecosystems. *Water, Air, and Soil Pollution, 48*(3-4), 451-456.

Serrano, L., Peñuelas, J., & Ustin, S. L. (2002). Remote sensing of nitrogen and lignin in Mediterranean vegetation from AVIRIS data: Decomposing biochemical from structural signals. *Remote Sensing of Environment, 81*(2), 355-364.

Shao, Y., He, Y., Gómez, A. H., Pereir, A. G., Qiu, Z., Zhang, Y. (2007). Visible/near infrared spectrometric technique for nondestructive assessment of tomato 'Heatwave' (Lycopersicum esculentum) quality characteristics. *Journal of Food Engineering*, *81(4)*, 672-678.

Singh, S. K., Houx III, J. H., Maw, M. J., & Fritschi, F. B. (2017). Assessment of growth, leaf N concentration and chlorophyll content of sweet sorghum using canopy reflectance. *Field crops research, 209*, 47-57.

Wold, H. (1966). Estimation of principal components and related models by iterative least squares. In P.R. Krishnaiah (Ed.), *Multivariate Analysis* (pp. 391-420). New York, USA: Academic Press.

Wold, S., Sjöström, M., & Eriksson, L. (2001). PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems, 58*(2), 109-130.

Wu, C., Niu, Z., Tang, Q., & Huang, W. (2008). Estimating chlorophyll content from hyperspectral vegetation indices: Modeling and validation. *Agricultural and Forest Meteorology, 148*(8–9), 1230-1241.

Zarco-Tejada, P. J., Miller, J. R., Morales, A., Berjón, A., & Agüera, J. (2004). Hyperspectral indices and model simulation for chlorophyll estimation in open-canopy tree crops. *Remote Sensing of Environment, 90*(4), 463-476.

Zarco-Tejada, P. J., Miller, J. R., Noland, T. L., Mohammed, G. H., & Sampson, P. H. (2001). Scaling-up and model inversion methods with narrowband optical indices for chlorophyll content estimation in closed forest canopies with hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing, 39*(7), 1491-1507.