



# Assessing the differences of two vineyards soils' by NIR spectroscopy and chemometrics

Sandia Machado<sup>a</sup>, Luisa Barreiros<sup>a,b</sup>, António R. Graça<sup>c</sup>, Manuel Madeira<sup>d</sup>, Ricardo N.M.J. Páscoa<sup>a,\*</sup>, Marcela A. Segundo<sup>a</sup>, João A. Lopes<sup>e</sup>

<sup>a</sup> LAQV, REQUIMTE, Department of Chemical Sciences, Laboratory of Applied Chemistry, Faculty of Pharmacy, University of Porto, Rua de Jorge Viterbo Ferreira n° 228, 4050-313 Porto, Portugal

<sup>b</sup> Escola Superior de Saúde, Instituto Politécnico do Porto, Rua Dr. António Bernardino de Almeida 400, 4200-072, Porto, Portugal

<sup>c</sup> Departamento de Investigação e Desenvolvimento, SOGRAPE Vinhos S.A., Aldeia Nova, 4430-852, Avintes, Portugal

<sup>d</sup> Forest Research Centre (CEF), School of Agriculture (ISA), University of Lisbon (ULisboa), Tapada da Ajuda, 1399-017 Lisboa, Portugal

<sup>e</sup> Research Institute for Medicines (iMed-ULisboa), Faculty of Pharmacy, University of Lisbon, Av. Prof. Gama Pinto, 1649-003, Lisboa, Portugal

## ARTICLE INFO

### Keywords:

Soils  
NIR spectroscopy  
Chemometrics  
PCA  
PLS-DA  
PLS

## ABSTRACT

Soil properties influence greatly the status of vine plants which consequently influences the quality of wine. Therefore, in the context of viticulture management, it is extremely important to assess the physical and chemical parameters of vineyards soils. In this study, the soils of two vineyards were analysed by near-infrared (NIR) spectroscopy and established analytical reference procedures. The main objective of this study was to verify if NIR spectroscopy is a potential tool to discriminate the soils of both vineyards as well as to quantify differences of soil's parameters. For that, a total of eight sampling spots were selected at each vineyard taking into consideration the soil type and sampled at different depths. The data analysis was performed using analysis of variance (ANOVA), principal component analysis (PCA) and partial least squares discriminant analysis (PLS-DA) and partial least squares (PLS) regression. The ANOVA results revealed that 12 out of the 18 parameters analysed through the reference procedures can be considered statistically different ( $p < 0.05$ ). Regarding PCA, the obtained results revealed a clear separation between the scores of both vineyards either considering NIR spectra or the chemical parameters. The PLS-DA model was able to obtain 100 % of correct predictions for the discrimination of both vineyards. PLS regression analysis using NIR spectra revealed  $R^2$  and RER values higher than 0.85 and 10, respectively, for 8 (pH (H<sub>2</sub>O), N, Ca<sup>2+</sup>, Mg<sup>2+</sup>, SB, CEC, ECEC and GSB) of the 18 chemical parameters evaluated. Concluding, these results demonstrate that it is possible to discriminate the soils of the different vineyards through NIR spectroscopy as well as to quantify several chemical parameters through soils NIR spectra in a rapid, accurate, cost-effective, simple and environmentally friendly way when compared to the reference procedures.

## 1. Introduction

In the wine and grape industries, a wide knowledge of soils is extremely important, as it allows a global understanding of the terroir elements that influence the vine performance [1]. This information can help to properly manage the vineyard's environmental system,

\* Corresponding author.

E-mail address: [rnpascoa@ff.up.pt](mailto:rnpascoa@ff.up.pt) (R.N.M.J. Páscoa).

<https://doi.org/10.1016/j.heliyon.2023.e23000>

Received 27 September 2023; Received in revised form 9 November 2023; Accepted 23 November 2023

Available online 28 November 2023

2405-8440/Â© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

increasing the viticultural and oenological quality [2,3]. Nevertheless, the knowledge about the soil effect in the fruit and wine composition is unclear due to the existence of contradictory results possibly due to different ways of interpretation. For example, studies have showed that wine composition reflects the geochemistry of soil [4,5] while other deny this association and other claim that soil only has a significant effect when the nutrient supply is not balanced [6]. Furthermore, other studies show that soil had an impact on berry weight, sugar, and anthocyanin content [7]. However, these effects were not caused directly by the soil chemical properties, but by the water retention associated to its physical properties [8].

Another aspect that may limit the comprehensive understanding of the soils effects is the vast number of properties necessary to access their complete characterization. Commonly, soil analysis involves field observations, highly dependent on the expertise of the technician, and laboratory analysis, which are expensive and time-consuming [9,10]. This leads to the analysis of a restrict number of parameters in a limited number of samples, making it difficult to obtain comprehensive information on soils, especially in terms of soil variability in large regions [11,12].

Currently, it has been common to propose the technique of near-infrared (NIR) spectroscopy as a good alternative for soil analysis, especially for its ability to simultaneously analyze chemical, physical and biological properties [12–15]. Besides being considered a fast, simple, and more economical technique, as it does not require sample processing, it is also considered eco-friendly as it does not require the use of chemicals [13,16]. Also, the portability option avoids the collection and transportation of samples, extending the analysis capacity to a greater number of samples and allowing the access to parameters that can only be estimated *in situ* [13]. Several studies regarding NIR spectroscopy have proved to be effective in accessing extensive information about soils regarding properties such as particle size distribution (texture), color, moisture content, organic matter content, mineralogy, among others [17–22]. In particular, the use of NIR spectroscopy in vineyards soils for its chemical assessment is not new. It started with the determination of cobalt with Salazar and co-workers [23] followed by the *in-situ* application of Cozzolino and co-workers [9] using a portable instrument for the determination of total nitrogen, organic carbon, K, S, P, pH, and electrical conductivity. Good results were obtained in both works fostering the use of NIR spectroscopy. Our research group [24], compared the performance between a benchtop and a portable NIR instruments for the classification of soils, demonstrating unexpectedly equivalent and good results for the purpose with both instruments. The same authors [25] used NIR spectroscopy to discriminate soils using wet and dried samples. The results revealed no significant difference between using wet and dry samples for soil discrimination. More recently, Rodríguez-Pérez and co-workers [26] explored the use of a portable spectroradiometer, covering a range between 350 and 2500 nm, for assessing a total of 12 parameters (such as: pH, CnH, N, P, K, Ca, Mn and Fe). Again, good results were obtained demonstrating the suitability of NIR spectroscopy.

In this article, the soils of two vineyards were analysed using NIR spectroscopy and reference procedures. This was performed to verify if NIR spectroscopy is able to discriminate the soils of both vineyards as well as to quantify the different chemical parameters assessed by the reference procedures using different soil types. For that, a total of eight sampling spots were selected at each vineyard taking into consideration the soil type and sampled at different depths. Since the same variety (Touriga Nacional) is grown in both vineyards, this characterization can be useful to understand/study the contribution of soil characteristics to the differences in the final products. This characterization can be included in a more comprehensive study about the contribution of soil characteristics to the final product differences.

## 2. Materials and methods

### 2.1. Study areas

The study vineyards are located in two Portuguese Denominations of Origin (DO). One is included in the Douro DO (Quinta da Leda, QL) and the other in the Dão DO (Quinta dos Carvalhais, QC). Both vineyards are property of Sogrape Vinhos S.A. (Portugal). The mean annual temperature in QL and QC is about 16 and 14 °C, respectively, and the mean annual rainfall is in the range 40–600 and 800–1000 mm, respectively. The QL vineyard occurs in a landscape made of Cambrian metamorphic formations (schists and phyllites), whereas the QC is in a landscape related to granitic formations (biotitic monzonite granites), that is, they occur in the most common soil parent materials of the Douro and Dão DO. In the former, according to the [27], soils are mostly Skeletic Regosol (Siltic, Humic), showing high content of coarse fragments (570–675 g kg<sup>-1</sup>) and silty loam texture (the silt and clay content ranges 296–375 and 115–160 g kg<sup>-1</sup>, respectively) [28]. In QC, the soils are mostly Eutric Regosol (Loamic, Humic) and show sandy loam texture, and contents of coarse fragments, silt and clay range 258–378, 109–171 and 87–149 g kg<sup>-1</sup>, respectively [29]. Despite to be allocated in the same Reference Soil Group (Regosols) [27]; soils show differences regarding content of coarse fragments and texture. In addition, they show differences regarding the mineralogy of the clay fraction. In the QL, illitic materials (clay micas) are dominant, and kaolinite, vermiculite, and interstratified illite-vermiculite occur at lower extent, whereas in QC, kaolinite is the dominant mineral, associated to low contents of illitic materials, vermiculite and interstratified illite-vermiculite. For vineyards installation, soil ploughing up to 50–60 cm depth was used in the QC, whereas in the QL the ripping practice was followed, given the nature of the parent material, and part of the area was terraced. That is, strong perturbation occurred in soil profile of both vineyards.

### 2.2. Sampling spots

At each vineyard, eight sampling spots (Fig. 1) were selected according to the NDVI information which revealed the zones of lowest and highest vegetal density (Supplementary Material Fig. S1). Besides information of soil mapping for both study areas, samples taken in each sampling spot allowed to classify soils in each sampling location. The soil samples were collected from spots planted with the

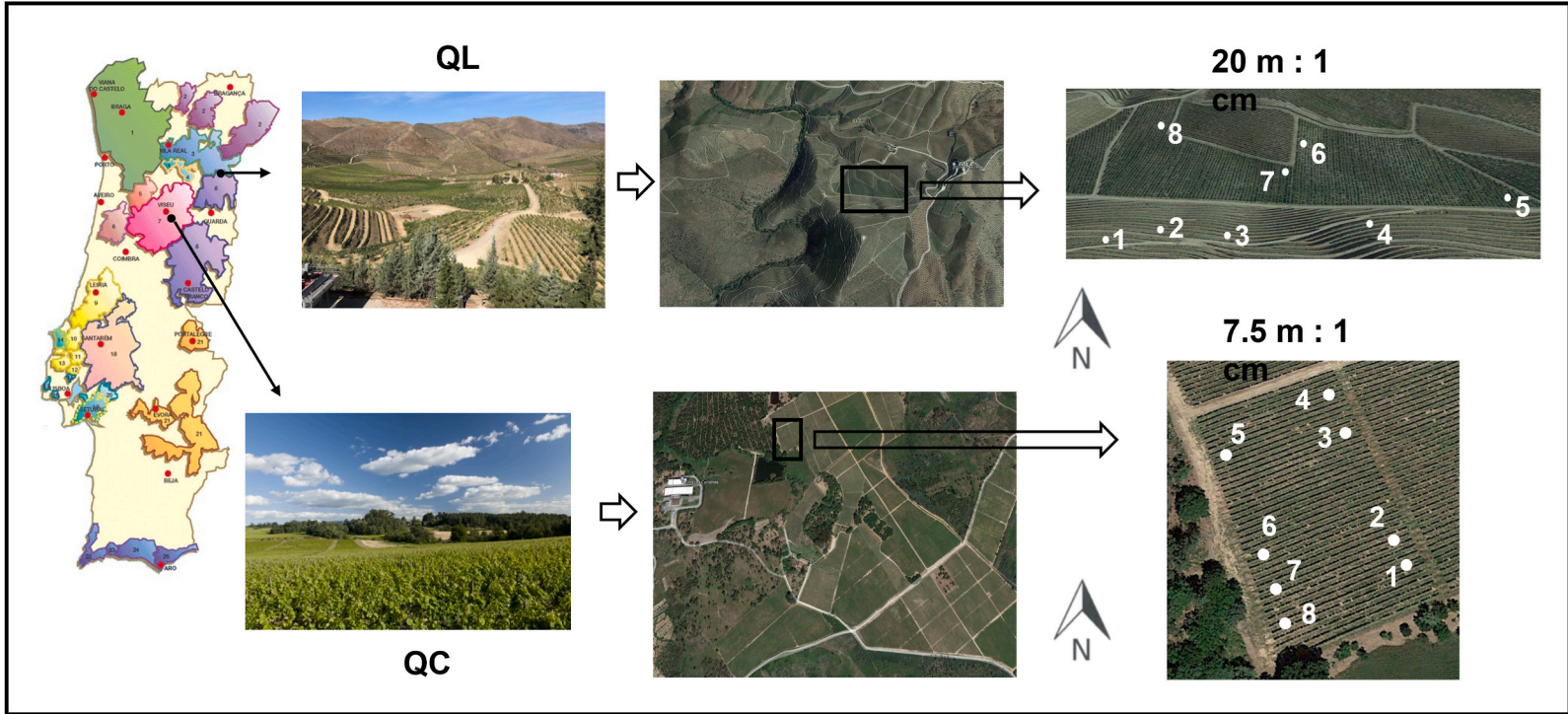


Fig. 1. Diagram of Portuguese wine regions with the location and photographs of both vineyards and the respective sampling spots.

same vinegrape variety (“Touriga Nacional”). Each sampling spot was at the middle distance of two vine rows (Fig. 1), corresponding to the 6th vine from the vineyard edge. At each sampling spot, soil samples were collected by an auger (approximately 500 g), at four different specified depths (Table 1), and then transported to the laboratory for further analysis. Given the soil horizons perturbation for vine installation, soil samples were taken at regular soil depths (0–20, 20–40, 40–60 and 60–80 cm). In some sampling spots, the collection of samples at some depths was not possible due to the high proportion of coarse fragments (QL) or consolidated rock (QC), which resulted in a total collection of 54 samples when considering both vineyards (30 from QC vineyard and 24 from QL vineyard). In QL vineyard the average altitude of the sampling spots is 240 m from sea level and there are some altitude differences between the sampling spots (Table 1). For QC vineyard, the average altitude of the sampling spots is 422 m from sea level, and altitude differences between the sampling spots are negligible (Table 1).

### 2.3. NIR spectral acquisition

The NIR spectra of soil samples were obtained in diffuse reflectance mode using a Fourier-transform near-infrared spectrometer (FTLA 2000, ABB, Canada) equipped with an indium-gallium-arsenide (InGaAs) detector. The spectral acquisition was controlled by the Bomen-Grams software (version 7, ABB, Canada). A portion of approximately 20 g of each soil sample, were transferred into borosilicate flasks for the collection of NIR spectra. Each spectrum was the average of 64 scans acquired with a resolution of  $8\text{ cm}^{-1}$  within a wavenumber range of 10,000 to  $4000\text{ cm}^{-1}$ . Each sample was scanned in triplicate, and the average was considered for further analysis, resulting in a total of 54 spectra. The background correction was performed using a Teflon reference material.

### 2.4. Soil analysis

Soil samples were air-dried and sieved by a 2 mm sieve. Soil properties were determined on the fine soil fraction ( $<2\text{ mm}$ ). Soil pH was determined potentiometrically in distilled water and 1 M KCl (soil:solution ratio 1:2.5). The particulate organic matter was physically separated by a 50 mm sieve, following the methodology described by Bruckert [30]. The organic C in the sieved soil samples and in the particulate organic matter fraction was determined by wet oxidation [31]. Total N was determined using Kjeldhal digestion [32] (Digestion System 40, Kjeltac Auto 1030 Analyzer). Extractable P and K were obtained using the Egnér-Riehm method [33]. Cation exchange capacity (CEC), and exchangeable  $\text{Ca}^{2+}$ ,  $\text{Mg}^{2+}$ ,  $\text{Na}^{+}$  and  $\text{K}^{+}$  were determined by the 1 M  $\text{NH}_4\text{OAc}$  method (at pH 7.0).

**Table 1**

GPS coordinates, altitude, soil type classification and sampling depth of each sampling spot on both study vineyards.

Sampling spots	GPS coordinates	Altitude (m)	Soil type <sup>a</sup>	Sampling depth (cm)
<b>QC vineyard</b>				
1	40°33'33.13"N 7°46'52.93"O	421	Hypereutric Regosol (Loamic, Humic)	0-20; 20-40; 40-60; 60-80
2	40°33'33.41"N 7°46'53.13"O	422	Hypereutric Regosol (Loamic, Humic)	0-20; 20-40; 40-60; 60-80
3	40°33'34.52"N 7°46'53.55"O	424	Eutric Regosol (Loamic, Humic)	0-20; 20-40; 40-60; 60-80
4	40°33'34.85"N 7°46'53.82"O	425	Endoleptic Hypereutric Regosol (Loamic, Ochric)	0-20; 20-40;
5	40°33'33.34"N 7°46'55.17"O	424	Eutric Regosol (Loamic, Ochric)	0-20; 20-40; 40-60; 60-80
6	40°33'33.43"N 7°46'54.80"O	421	Eutric Regosol (Loamic, Humic)	0-20; 20-40; 40-60; 60-80
7	40°33'33.06"N 7°46'54.64"O	421	Hypereutric Regosol (Loamic, Humic)	0-20; 20-40; 40-60; 60-80
8	40°33'32.59"N 7°46'54.38"O	421	Hypereutric Regosol (Loamic, Humic)	0-20; 20-40; 40-60; 60-80 cm
<b>QL vineyard</b>				
1	41°01'08.77"N 7°01'08.80"O	257	Hypereutric Regosol (Siltic, Humic)	0-20; 20-40; 40-60;
2	41°01'09.11"N 7°01'05.91"O	255	Skeletal Hypereutric Regosol (Siltic, Ochric)	0-20; 20-40; 40-60;
3	41°01'09.38"N 7°01'04.54"O	252	Hypereutric Regosol (Siltic, Humic)	0-20; 20-40; 40-60;
4	41°01'09.29"N 7°01'01.38"O	254	Skeletal Hypereutric Regosol (Siltic, Humic)	0-20; 20-40; 40-60;
5	41°01'10.36"N 7°00'58.20"O	246	Skeletal Hypereutric Regosol (Siltic, Humic)	0-20; 20-40; 40-60;
6	41°01'12.70"N 7°01'02.78"O	223	Skeletal Hypereutric Regosol (Siltic, Ochric)	0-20; 20-40; 40-60;
7	41°01'12.08"N 7°01'05.00"O	226	Skeletal Hypereutric Regosol (Siltic, Humic)	0-20; 20-40; 40-60;
8	41°01'14.90"N 7°01'08.05"O	205	Skeletal Hypereutric Regosol (Siltic, Ochric)	0-20; 20-40; 40-60;

<sup>a</sup> According to the World Reference Base for Soil Resources (WRB-FAO, 2015), the classification system used in the soil map of both wine regions.

Exchangeable  $\text{Al}^{3+}$  was obtained by extraction with the KCl 1 M. Ca, Mg, Na, K, and Al were determined by atomic absorption spectrophotometry (AAS; Analyst 300, PerkinElmer), while extractable P was determined by the molybdate-blue method. The sum of exchangeable  $\text{Ca}^{2+}$ ,  $\text{Mg}^{2+}$ ,  $\text{Na}^{+}$  and  $\text{K}^{+}$  is designated as SB. The effective cation exchange capacity (ECEC) was estimated by the sum of SB with extractable  $\text{Al}^{3+}$ . The Ca–Mg–Na–K saturation degree (GSB) was estimated by  $(\text{SB}/\text{ECEC}) \times 100$ , and the Al saturation degree (GSA) as  $(\text{Al}/\text{ECEC}) \times 100$ .

## 2.5. Data analysis

The results of the chemical parameters of soils obtained by reference procedures were firstly analysed by ANOVA to check for statistically significant difference between the chemical parameters in both vineyards. After that, the results of the soil chemical parameters as well as the NIR spectral information were independently analysed through principal component analysis (PCA) [34] and partial least square discriminant analysis (PLS-DA) [35]. And finally, partial least squares (PLS) [36] was applied to identify the relationship between the chemical parameters with the NIR spectral information. Briefly, PCA was used for outliers screening and cluster analysis, PLS-DA was used for developing discrimination models while PLS was used to establish regression models against soil parameters. In PLS, each soil parameter was modelled in separate, which means the PLS-1 algorithm was used. Before the application of any of these chemometric tools, the soils chemical parameters were auto scaled due to their different units and values allowing the variables to have the same impact on the chemometric analysis. Also, all NIR spectra were pre-processed with Savitzky-Golay (using 15 points filter width, second polynomial order and first derivative) followed by standard normal variate (SNV). Before PCA, PLS-DA and PLS matrices were mean centered. The selection of the pre-processing technique was made according to previous works developed by our research team [24,25]. The entire NIR spectral range was used.

For PLS models, the respective data set was divided in two sets, 70 % for calibration and 30 % for validation. This division was made randomly but taking into consideration that the values of the validation set for each soil parameter modelled were within the

**Table 2**  
Soils parameters obtained by the reference procedures.

Parameters	Vineyards	Min	Max	Average	St. Dev.	p-value
pH (H <sub>2</sub> O)	QC	5.50	7.00	6.30	0.40	1.2E-07
	QL	6.90	7.80	7.10	0.20	
pH (1 M KCl)	QC	4.30	5.80	5.10	0.50	9.8E-02
	QL	4.00	5.50	4.80	0.50	
C <sub>org</sub> (g/kg)	QC	7.60	21.60	12.90	4.60	3.5E-01
	QL	6.60	20.30	11.40	4.60	
C <sub>n</sub> H (g/kg)	QC	0.70	7.70	2.40	2.00	2.1E-01
	QL	0.90	10.40	3.40	2.60	
N (g/kg)	QC	0.40	1.30	0.70	0.30	9.7E-01
	QL	0.30	1.50	0.70	0.30	
C/N	QC	9.20	31.90	19.40	5.30	5.8E-02
	QL	12.90	22.60	16.50	2.50	
Ca <sup>2+</sup> (cmol/kg)	QC	1.80	7.80	4.00	1.80	1.8E-08
	QL	12.90	42.80	19.20	7.70	
Mg <sup>2+</sup> (cmol/kg)	QC	0.10	0.70	0.30	0.20	3.3E-10
	QL	2.30	9.80	4.60	1.90	
Na <sup>+</sup> (cmol/kg)	QC	0.01	0.15	0.04	0.04	1.6E-03
	QL	0.03	0.49	0.17	0.14	
K <sup>+</sup> (cmol/kg)	QC	0.16	0.61	0.32	0.12	1.4E-07
	QL	0.01	0.23	0.08	0.07	
SB (cmol/kg)	QC	2.30	9.00	4.70	2.10	5.5E-09
	QL	15.60	52.80	24.00	9.40	
Al <sup>3+</sup> (cmol/kg)	QC	0.10	0.61	0.23	0.13	1.3E-02
	QL	0.07	0.24	0.14	0.04	
CEC (cmol/kg)	QC	4.10	9.60	6.70	1.70	2.3E-06
	QL	7.30	12.60	10.00	1.50	
ECEC (cmol/kg)	QC	2.50	9.20	4.90	2.00	6.2E-09
	QL	15.70	53.00	24.20	9.40	
GSB (%)	QC	42.00	94.00	67.00	17.00	3.4E-10
	QL	174.00	421.00	240.00	73.00	
GSA (%)	QC	1.40	18.80	6.00	4.90	1.4E-04
	QL	0.30	0.90	0.60	0.20	
P (mg/kg)	QC	2.60	174.60	25.90	42.90	6.7E-01
	QL	5.20	61.30	21.10	13.70	
K (mg/kg)	QC	10.30	208.80	119.90	48.80	3.7E-05
	QL	37.20	75.40	59.60	10.30	

Considering a confidence interval of 5 %, 12 parameters can be considered statistically different as they yield p-values below this confidence interval ( $p < 0.05$ ), namely: pH (H<sub>2</sub>O), Ca<sup>2+</sup>, Mg<sup>2+</sup>, Na<sup>+</sup>, K<sup>+</sup>; SB, Al<sup>3+</sup>; CEC, ECEC, GSA, GSB and K. In this sense, considering the parameters analysed by the reference procedures, the soils of both vineyards can be considered statistically different, since the parameters significantly different correspond to 2/3 of the total number of parameters (18).

respective values of the calibration set. The selection of the optimal number of latent variables (LV) was made using the leave-one-sample out cross-validation method and based on a compromise between the lowest root mean square error of cross-validation (RMSECV) and a lowest number of LV. This was performed using only the calibration set. After PLS models optimization using only the calibration set, the validation set was projected to evaluate the respective models' performance. The evaluation of the PLS models' performance was made using root mean square error of calibration (RMSEC), RMSECV, root mean square error of prediction (RMSEP), determination coefficient of prediction ( $R_p^2$ ) and the range error ratio (RER). The RMSEC, RMSECV and RMSEP were calculated according to equation (1).

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (\hat{y}_i - y_i)^2}{N}} \quad (\text{Eq. 1})$$

In Eq. (1),  $N$  represents the number of samples,  $y_i$  represents the experimental result for sample  $i$  and  $\hat{y}_i$  represents the value obtained with the calibration – RMSEC, cross-validation – RMSECV and prediction – RMSEP for each sample. The RER parameter was estimated according to Eq. (2) where  $y_{\max}$  and  $y_{\min}$  represent the maximum and minimum values of the validation set, respectively.

$$RER = \frac{y_{\max} - y_{\min}}{RMSEP} \quad (\text{Eq. 2})$$

For PLS-DA, the samples were classified according to the respective vineyard. Again, the respective data set was divided in two sets, 70 % for calibration and 30 % for validation. This division was made randomly but taking into consideration the same proportion of each vineyard in both sets to prevent unbalanced classes at both sets. The selection of the optimal number of LV was made using the leave-one-sample-out cross-validation procedure based on a compromise between the highest number of correct predictions and a lowest number of LV. Again, this was performed considering only the calibration set. After the estimation of the optimal number of LV, the validation set was projected onto the calibration set to assess the accuracy of the model. The accuracy of the respective model was evaluated through the percentage of correct predictions, expressed in the form of confusion matrices.

All calculations involving PCA, PLS and PLS-DA modelling were performed in Matlab R2014a version 8.3 (MathWorks, USA) using PLS Toolbox version 8.2.1 (Eigenvector Research Inc., USA).

### 3. Results and discussion

#### 3.1. ANOVA of the soil's chemical parameters

The soil parameters obtained by the reference procedures, considering all the sampling spots, were analysed by ANOVA to verify if there are significant differences between both vineyards. Table 2 summarizes the information of the minimum, maximum, average, standard deviation values for each vineyard. The p-value is provided for each parameter in order to evaluate the difference between the two vineyards.

#### 3.2. Exploratory data analysis from soils chemical data

After the ANOVA analysis, a PCA was performed to the soil parameters obtained by the reference procedures to verify the formation of clusters and the presence of outliers, when considering all the sampling spots and horizons. The criterion used for outlier detection was based on the Hotelling's  $T^2$  (weighted sum of squared scores) and Q residuals (sum of squared residuals). Samples that present a score above the confidence limits of 95 % for both statistics was considered an outlier. Taking this criterion into account, it is possible

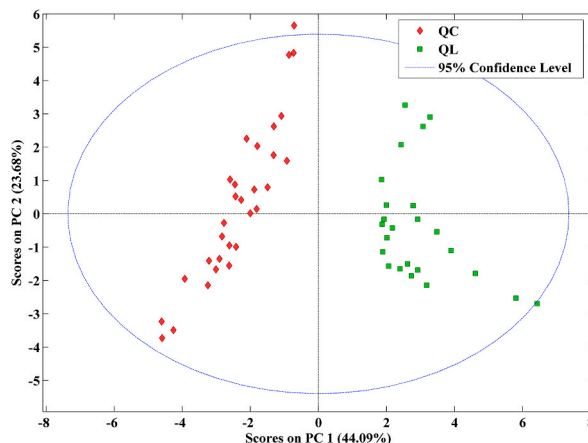


Fig. 2. Score plot of the first two principal components using soils parameters which captured 67.8 % of the total variance (data were auto scaled).

to assume that no outliers were detected. The PCA model was built using two principal components which captured 67.8 % of the total variance and the obtained scores are shown in Fig. 2.

As can be seen, there is a clear separation between the two vineyards, when considered the soil chemical parameters. In fact, the PCA model could be built with just one principal component as the separation seen in the score plot is given by the scores of the first principal component (PC1). The scores of QC vineyard are negative while the scores of QL vineyard are positive. Once more, the PCA analysis using the soils parameters obtained by the reference procedures revealed that soils of both vineyards are different.

The loadings of the PCA model are shown in Fig. 3 to verify which soils parameters had a higher impact on the obtained scores. As the scores of PC1 were capable of clearly separate the soils of both vineyards, just the loadings of this principal component were shown.

As can be seen, the variables (soils parameters) that showed the highest contribution were: pH (H<sub>2</sub>O), C/N, Ca<sup>2+</sup>, Mg<sup>2+</sup>, SB, Al<sup>3+</sup>, CEC, ECEC, GSA and GSB. This makes sense as the obtained results using ANOVA revealed that these parameters were statistically different ( $p < 0.05$ ).

### 3.3. Vineyards soils discrimination analysis from chemical soils data

Subsequently it was decided to use a supervised chemometric tool, namely PLS-DA, for the discrimination of the vineyard's soils using the soils parameters. Fig. 4 shows the results of the application of PLS-DA model using two LVs.

As can be seen, a total of 100 % of correct predictions were obtained for both vineyards, which demonstrates once more that the soils of both vineyards are different considering the parameters analysed. The regression coefficient vector of the PLS-DA model (for LV1 since it is symmetric to LV2) is given in Fig. 5.

The variables (soils parameters) that showed the highest contribution to the respective PLS-DA model were: pH (H<sub>2</sub>O), Mg<sup>2+</sup>, K<sup>+</sup>, CEC, GSB and extractable K which make sense as these are included in the parameters that were considered statistically different ( $p < 0.05$ ) through the ANOVA.

After the application of these chemometric tools, it is evident that the soils of both vineyards can be considered different when the analysed parameters are taken into account.

In order to understand if it is possible to obtain the same results using NIR spectra, the chemometric tools used for the soils chemical parameters will be applied to NIR spectra of soil. All NIR soil samples spectra including soils from different depths were used in the developed PCA, PLS and PLS-DA models because in past study developed by our research group using NIR spectroscopy and soils found that there was no significant influence of the depth where the sample was collected and its NIR spectra [25].

### 3.4. PCA of soils NIR spectra

A PCA was performed using soils NIR spectra to verify the potential formation of clusters and the presence of outliers. No outliers were detected following the criterion above-mentioned. The PCA model was built using five PCs which captured 98.5 % of the total variance and the obtained scores for the first two PCs are shown in Fig. 6.

As can be seen, there is a clear separation between the soils NIR spectra of both vineyards. Just like as obtained for the PCA model considering soils parameters, the PCA model could be built with just one PC as the separation between both vineyards is given by the scores of PC1. The scores of C vineyard are positive while the scores of L vineyard are negative. These results demonstrate that the soils

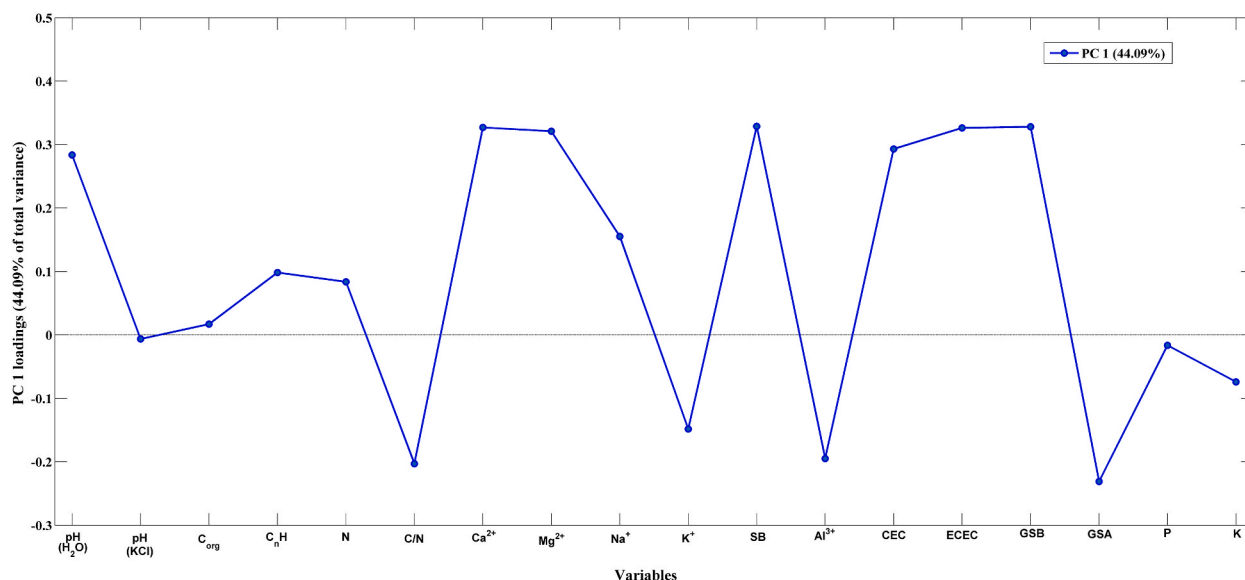


Fig. 3. Loadings of PC1 of the PCA model performed using soils parameters.

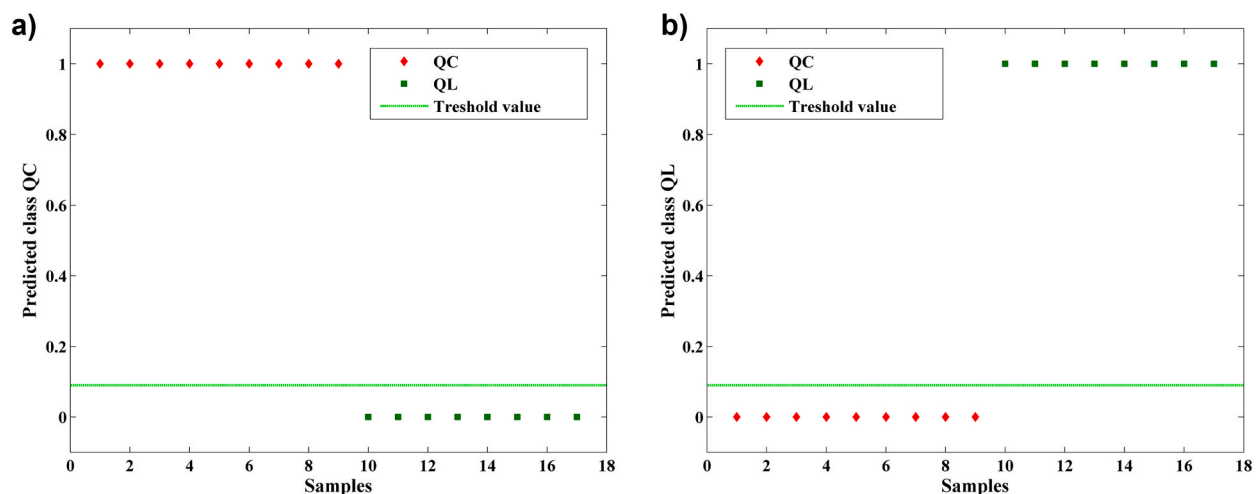


Fig. 4. Class predictions for the 2 LV PLS-DA model calibrated from the soil's chemical parameters for QC (a) and QL (b).

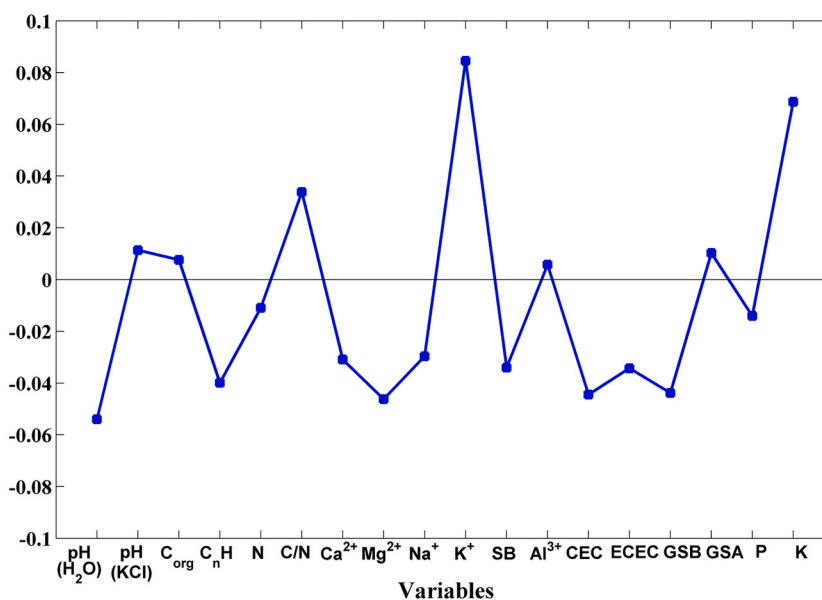


Fig. 5. Regression coefficient vector for the PLS-DA model considering soils parameters pre-processed with auto scaling and using 2 LV.

NIR spectra of both vineyards are different, which is in agreement with the findings when using soils parameters.

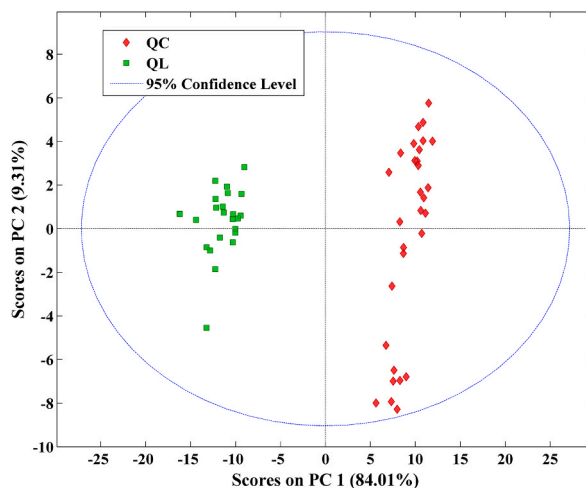
The loadings of the respective PCA model are shown in Fig. 7 to verify which wavenumbers had a higher impact on the obtained scores. Once more, as the scores of PC1 were capable of clear separating the soils of both vineyards, just the loadings of PC1 were shown instead of showing all the loadings of the five PCs.

The wavenumbers that showed the highest contribution to the PCA model were located around 7055, 5260 and 4500  $\text{cm}^{-1}$ . The wavenumbers around 7055  $\text{cm}^{-1}$  can be related with kaolin doublet from clay minerals and water bands [37]. The wavenumbers around 5260  $\text{cm}^{-1}$  can be attributed to O–H water bands while the wavenumbers around 4500  $\text{cm}^{-1}$  can be related with kaolin doublet, smectite and illite from clay minerals as well as to Al–OH bonds [19]. Therefore, this means the soils of both vineyards have different content of these compounds.

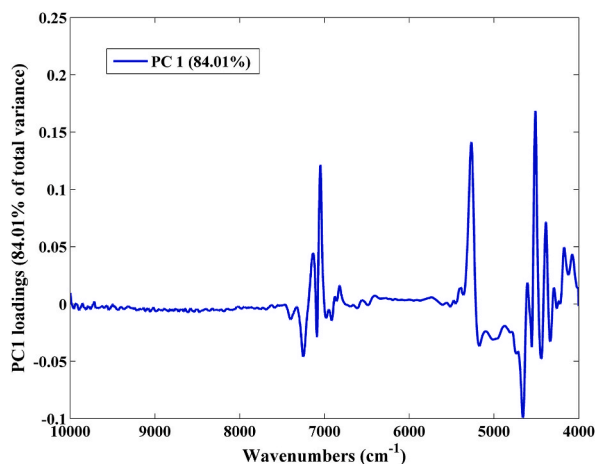
### 3.5. Vineyards soils discrimination analysis from NIR spectra

After the PCA, a supervised chemometric tool, namely PLS-DA, was used for the discrimination of the vineyard's soils using the NIR spectral information. Fig. 8 shows the results of the application of PLS-DA model using two LVs.

A total of 100 % of correct predictions were obtained for both vineyards, which demonstrates again that the soils of both vineyards



**Fig. 6.** Score plot of the first two principal components using NIR soils' spectra which captured 93.3 % of the total variance (spectra were pre-processed with Savitzky-Golay (using 15 points filter width, second polynomial order and first derivative) followed by SNV and then mean centered).



**Fig. 7.** Loadings of PC1 of the PCA model performed using soils NIR spectra.

are different considering NIR spectra. The regression coefficient vector of the PLS-DA model (for LV1 since it is symmetric to LV2) is given in Fig. 9.

The wavenumbers with the highest contribution to this PLS-DA model were located around 7040, 5275 and between 4550 and 4350  $\text{cm}^{-1}$ . These wavenumbers were similar to the ones obtained in the loadings of the PCA performed before using NIR spectra. This result reinforces the previous results that showed that the differences obtained with the use of NIR spectra could be related with kaolin doublet, smectite and illite from clay minerals, Al compounds and the water content. This is in agreement with what was mentioned in subsection 2.1, where in QL vineyard illitic materials are dominant while in QC vineyard, kaolinite is the dominant material. Therefore, it could be stated that the soils NIR spectra of both vineyards are different which is in agreement with the obtained results using the reference procedures. Moreover, the NIR technique provided the same accurate results as the obtained with the reference procedures for soils discrimination with all the advantages of being rapid, cost-effective, without needing any preparation of the sample and environmentally friendly.

### 3.6. Prediction of soils chemical data from NIR spectra

Although the previous results reveal that the soils of both vineyards have significant differences, the PLS models were built considering the soils of both vineyards together. This strategy was performed to obtain more robust PLS models that can be applied to different types of soils, as well as to a greater number of samples. The obtained PLS models' results are shown in Table 3 and Fig. 10 (only the best PLS models were shown).

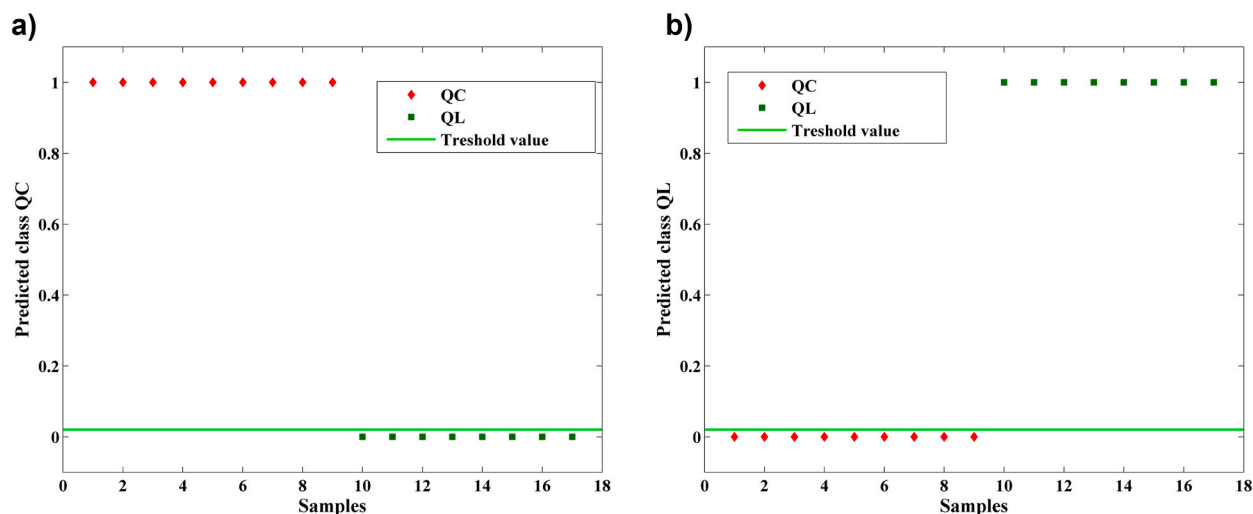


Fig. 8. Class predictions for the 2 LV PLS-DA model calibrated from the soils NIR spectra for QC (a) and QL (b).

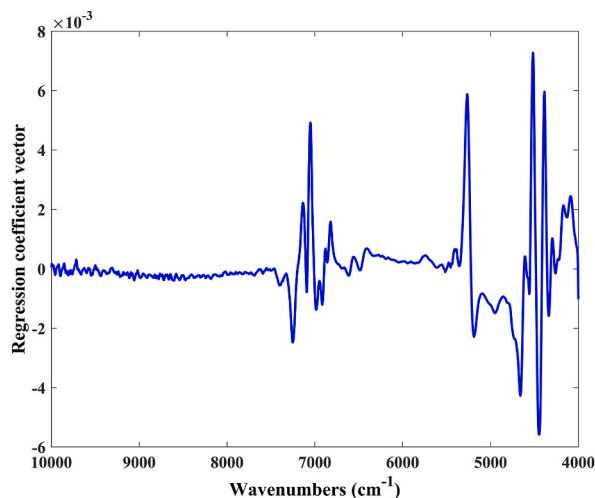


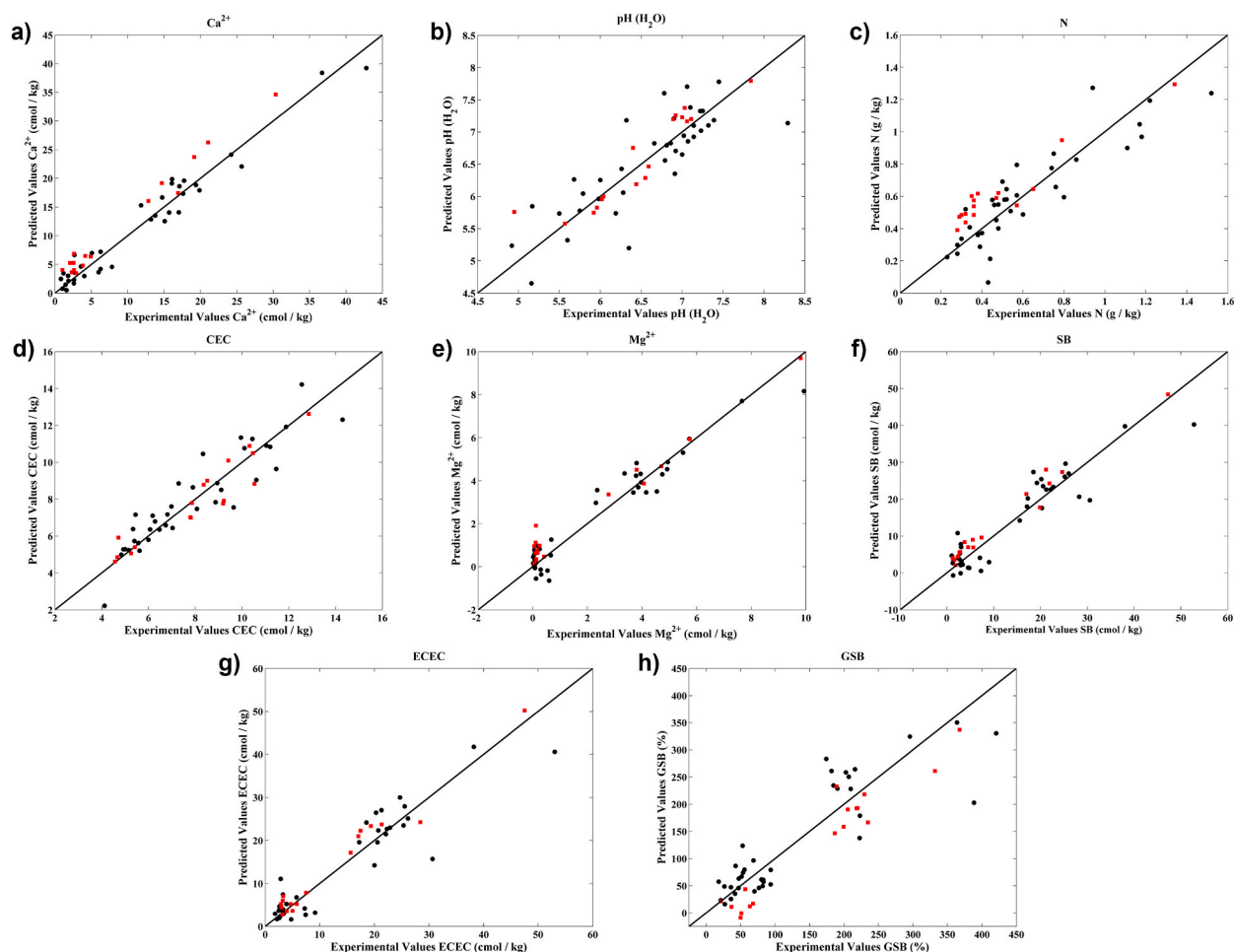
Fig. 9. Regression coefficient vector for the PLS-DA model considering the NIR spectra pre-processed with mean centring and using 2 LV.

The precision of PLS models results demonstrate the capacity of NIR spectroscopy for the quantification of chemical parameters. In detail, the  $R_p^2$  of the best models were all above 0.85 and most of them above 0.9 and the RER values of the best models were all higher than 10, some of which ( $\text{Ca}^{2+}$ ,  $\text{Mg}^{2+}$ , SB, ECEC) were higher than 15 which can be considered as very good models for research quantification [38]. In more detail, for pH parameter, the PLS results obtained in this work ( $R_p^2$  of 0.85 and 0.78 and RMSEP of 0.27 and 0.25 for  $\text{pH}(\text{H}_2\text{O})$  and  $\text{pH}(\text{KCl})$ , respectively) are similar to the ones obtained in several works: Rodríguez-Pérez et al., 2021 [26] ( $R_{CV}^2$  of 0.96 and RMSECV of 0.32 for  $\text{pH}(\text{H}_2\text{O})$ ); Marín-Gonzalez et al., 2013 [39] ( $R_p^2$  of 0.86 and RMSEP of 0.33 for  $\text{pH}(\text{H}_2\text{O})$ ); and de Santana and Daly 2022 [40] ( $R_p^2$  of 0.77 and RMSEP of 0.48 for  $\text{pH}(\text{H}_2\text{O})$  and  $R_p^2$  of 0.72 and RMSEP of 0.47 for  $\text{pH}(\text{CaCl}_2)$ ). For organic carbon ( $C_{org}$ ), our results ( $R_p^2$  of 0.84 and RMSEP of 1.70) are comparable to the ones obtained by Vaudour et al., 2018 [41] ( $R_{CV}^2$  of 0.87 and RMSECV of 1.84). For total N, our results ( $R_p^2$  of 0.90 and RMSEP of 0.10) were similar to the ones obtained by de Santana and Daly, 2022 [40], ( $R_p^2$  of 0.92 and RMSEP of 0.05) and better than those obtained by Rodríguez-Pérez et al., 2021 [26]. For  $\text{Ca}^{2+}$ , the obtained results in this work ( $R_p^2$  of 0.98 and RMSEP of 1.70) are better than to the ones obtained by Rodríguez-Pérez et al., 2021 [26] ( $R_{CV}^2$  of 0.91 and RMSECV 1.96) and Cozzolino and Morón [42] ( $R_{CV}^2$  of 0.90 and RMSECV 2.9). For  $\text{Mg}^{2+}$ , the obtained results ( $R_p^2$  of 0.97 and RMSEP of 0.52) are significantly better than the ones reported by Vaudour et al., 2018 [41] ( $R_{CV}^2$  of 0.54 and RMSECV of 0.018). For SB (sum of exchangeable  $\text{Ca}^{2+}$ ,  $\text{Mg}^{2+}$ ,  $\text{Na}^+$  and  $\text{K}^+$ ) PLS model, the obtained results hereby were very good ( $R_p^2$  of 0.95 and RMSEP of 3.0), significantly better than the ones show in Johnson et al., 2019 [43] ( $R_{CV}^2$  of 0.64 and RMSECV of 6.07). For CEC parameter, again, the obtained results ( $R_p^2$  of 0.89 and RMSEP of 0.77) are equivalent to the ones shown by Vaudour et al., 2018 [41] ( $R_{CV}^2$  of 0.91 and RMSECV of 1.96) and better than the ones reported by de Santana and Daly, 2022 [40], ( $R_p^2$  of 0.79 and RMSEP of 2.62). For eCEC (effective cation exchange capacity), the obtained results were very good ( $R_p^2$  of 0.97 and RMSEP of 2.40), significantly better than the

**Table 3**

PLS calibration models results for the different soils parameters using the entire NIR spectra pre-processed with Savitzky-Golay (using 15 points filter width, second polynomial order and first derivative) followed by SNV and then mean centered.

Parameters	LV	RMSEC	RMSECV	RMSEP	$R_C^2$	$R_P^2$	RER
pH -H <sub>2</sub> O	8	0.19	0.44	0.27	0.94	0.85	11
pH - KCl	6	0.30	0.50	0.25	0.73	0.78	7
Corg	6	1.50	2.20	1.70	0.86	0.84	9
CnH	6	0.79	1.30	1.50	0.85	0.27	4
N	6	0.09	0.14	0.10	0.91	0.90	13
C/N	3	3.70	4.50	2.80	0.42	0.55	6
Ca <sup>2+</sup>	8	1.40	3.50	1.70	0.98	0.98	17
Mg <sup>2+</sup>	5	0.42	0.64	0.52	0.97	0.97	19
Na <sup>+</sup>	2	0.12	0.13	0.09	0.37	0.46	5
K <sup>+</sup>	6	0.06	0.09	0.08	0.75	0.56	5
SB	5	2.20	4.40	3.00	0.97	0.95	16
Al <sup>3+</sup>	6	0.20	0.34	0.21	0.55	0.69	7
CEC	5	0.65	1.00	0.77	0.94	0.89	11
ECEC	7	1.80	4.50	2.40	0.98	0.97	19
GSB	7	19.00	52.00	27.00	0.97	0.93	13
GSA	6	8.70	13.00	11.00	0.68	0.49	5
P	4	13.00	18.00	14.00	0.19	0.07	4
K	5	27.00	38.00	31.00	0.61	0.36	5



**Fig. 10.** Experimental values versus the cross-validation (●) and prediction (■) model estimated obtained for Ca<sup>2+</sup>(a), pH (H<sub>2</sub>O) (b), N (c), CEC (d), Mg<sup>2+</sup>(e), SB (f), ECEC (g) and GSB (h).

one show in Johnson et al., 2019 [43] ( $R_{CV}^2$  of 0.67 and RMSECV of 5.47). For GSB (Ca–Mg–Na–K saturation degree) model, the obtained results hereby were very good with a  $R_p^2$  and RMSEP of 0.93 and 27.0, respectively, and to the best of our knowledge has never been reported.

There were also some important parameters not well predicted such as  $K^+$  and extractable P. For  $K^+$ , the obtained results hereby ( $R_p^2$  of 0.56 and RMSEP of 0.08) were not so accurate but similar to the ones shown by Rodríguez-Pérez et al., 2021 [26] ( $R_{CV}^2$  of 0.68 and RMSECV of 0.08). Regarding extractable P, the obtained results ( $R_p^2$  of 0.07 and RMSEP of 14.0) revealed the lack of accuracy of NIR spectroscopy for its quantification, but again according to other results reported by Johnson et al., 2019 [43] ( $R_{CV}^2$  of 0.16 and RMSECV of 14.6) and Vaudour et al., 2018 [41] ( $R_{CV}^2$  of 0.14 and RMSECV of 0.01), but much worse than the ones reported by Rodríguez-Pérez et al., 2021 [26] ( $R_{CV}^2$  of 0.92 and RMSECV of 5.9).

The analysis of the regression coefficient vectors (Fig. 11) for the best PLS models was also performed to identify the most important wavenumbers for the respective models.

Through the analysis of the regression coefficient vectors it is evident that the spectral regions that showed the highest contribution for all the models were located around 7200, 5200 and within 4650–4200  $cm^{-1}$ . The spectral regions around 7200  $cm^{-1}$  can be attributed to the amount of kaolin doublet from clay minerals, hydroxyl bounds and water bonds [37], the spectral region around 5200  $cm^{-1}$ , can be associated to water bounds while the spectral regions between 4650 and 4200  $cm^{-1}$  can be attributed to kaolin doublet, smectite and illite from clay minerals, Al compounds, Fe–OH and aliphatic compounds [19]. These spectral regions were similar to the ones found for PCA and PLS-DA, indicating that these are the most informative regions of the soils NIR spectra.

Concluding, NIR spectroscopy revealed to be accurate for the quantification of the pH ( $H_2O$ ), N,  $Ca^{2+}$ ,  $Mg^{2+}$ , SB, CEC, ECEC and GSB soils parameters, besides the NIR spectra of both vineyards being different. Although this method has been tested using samples from different location but from the same soil type (Regosol), showing differences regarding texture and mineralogical constitution, the obtained results demonstrated that NIR spectroscopy is able to accurately quantify differences of several chemical parameters in soils in a rapid, accurate, cost-effective, simple and environmentally friendly way. This methodology does not replace the common reference procedures for several other parameters, and when the best precision is required. However, NIR spectroscopy can be used to

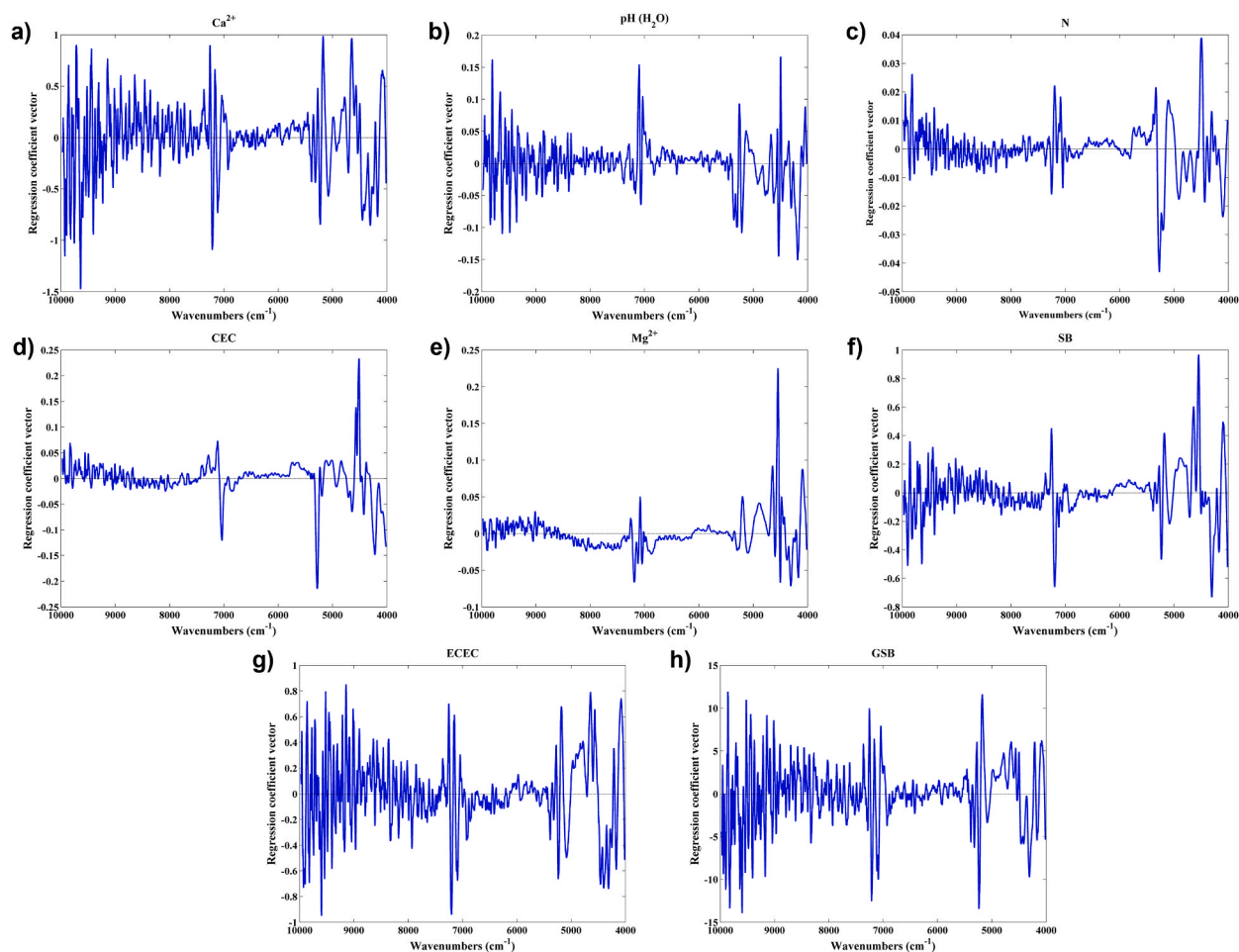


Fig. 11. Regression coefficient vectors for the PLS models for  $Ca^{2+}$  (a), pH ( $H_2O$ ) (b), N (c), CEC (d),  $Mg^{2+}$  (e), SB (f), ECEC (g) and GSB (h).

some extent as a tool that can replace reference procedures if applied in controlled conditions.

#### 4. Conclusions

The present work revealed that the soils of QC and QL vineyards are significantly different from the point of view of the chemical characterization. Collected NIR spectra from both vineyards also revealed significant differences at the spectral level. Regarding the analysis of the soil's chemical parameters, the ANOVA showed that 12 parameters from a total of 18 can be considered statistically different ( $p < 0.05$ ). The PCA displayed a clear separation between the scores of both vineyards in PC1 and the most important variables for this separation were most of the ones identified as statistically different by ANOVA. The PLS-DA yielded 100 % of correct predictions for both vineyards and, once more, the most important variables for this discrimination were some of the ones identified as statistically different by ANOVA. Therefore, through the chemical parameters it is possible to conclude that the soils of both vineyards can be considered different.

Regarding the analysis of soils NIR spectra, the obtained results demonstrated the ability and suitability of NIR spectroscopy to soils analysis. In fact, through the use of NIR spectra it was possible to discriminate both vineyards' soils as well as to quantify several soils parameters, namely pH (H<sub>2</sub>O), N, Ca<sup>2+</sup>, Mg<sup>2+</sup>, SB, CEC, ECEC and GSB with a high accuracy. The PCA using NIR spectra presented a clear separation between the scores of both vineyards in PC1 and pointed to differences in the content of kaolin doublet, smectite and illite from clay minerals as well as to Al–OH bonds. The PLS-DA also yielded 100 % of correct predictions for both vineyards and the spectral regions more important to this discrimination could be related to the compounds abovementioned for PCA as well as to water content. The PLS results revealed R<sub>p</sub><sup>2</sup> and RER values higher than 0.85 and 10, respectively, for 8 (pH (H<sub>2</sub>O), N, Ca<sup>2+</sup>, Mg<sup>2+</sup>, SB, CEC, ECEC and GSB) of the 18 chemical parameters evaluated. Moreover, the analysis of PLS regression coefficient vectors showed that quantification of these compounds can be associated with the amount of kaolin doublet, smectite and illite of clay minerals, hydroxyl bounds, Al compounds, Fe–OH compounds and aliphatic compounds.

Concluding, these results demonstrate that it is possible to discriminate the soils of the different vineyards through NIR spectroscopy as well as to quantify several chemical parameters through soils NIR spectra in a rapid, accurate, cost-effective, simple and environmentally friendly way when compared to the reference procedures. This proves that this methodology can allow the soil analysis in a large-scale that can, therefore, be integrated into a more complete study of the environmental system in order to better understand the contributions of each terroir element in the grapes development and their final characteristics.

#### Data availability statement

Data has not been deposited in publicly repository but will be made available on request.

#### CRedit authorship contribution statement

**Sandia Machado:** Investigation, Conceptualization. **Luisa Barreiros:** Writing - review & editing, Supervision. **António R. Graça:** Resources, Formal analysis. **Manuel Madeira:** Writing - review & editing, Methodology, Investigation, Conceptualization. **Ricardo N. M.J. Páscoa:** Writing - review & editing, Writing - original draft, Validation, Investigation, Data curation, Conceptualization. **Marcela A. Segundo:** Writing - review & editing, Supervision, Formal analysis, Conceptualization. **João A. Lopes:** Writing - review & editing, Validation, Supervision, Project administration, Funding acquisition, Data curation, Conceptualization.

#### Declaration of competing interest

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

#### Acknowledgements

This work received financial support from the European Union (FEDER funds through COMPETE POCI-01-0145-FEDER-016735) and National Funds (FCT, Fundação para a Ciência e a Tecnologia) through project PTDC/AGR-PRO/6817/2014 and also under UIDB/50006/2020 and UIDP/50006/2020. This work was also supported by AgriFood XXI I&D&I project (NORTE-01-0145-FEDER-000041) cofinanced by European Regional Development Fund (ERDF), through the NORTE 2020 (Programa Operacional Regional do Norte 2014/2020). S. Machado thanks FCT and POCH (Programa Operacional Capital Humano) for her PhD grant (SFRH/BD/122730/2016). J. Lopes received additional support from FCT-MCTES. L. Barreiros and R.N.M.J. Páscoa acknowledge funding from FCT through program DL 57/2016 – Norma transitória.

#### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.heliyon.2023.e23000>.

## References

- [1] D. Tesic, et al., Environmental effects on cv Cabernet Sauvignon (*Vitis vinifera* L.) grown in Hawke's Bay, New Zealand.: 1. Phenology and characterisation of viticultural environments, *Aust. J. Grape Wine Res.* 8 (2008) 15–26.
- [2] E.A.C. Costantini, et al., Soil and climate functional characters for grape ripening and wine quality of "Vino Nobile di Montepulciano", First Intl Workshop on Strategies to Optimize Wine Grape Quality (427) (1996) 45–55.
- [3] R.A.V. Rossel, C. Chen, Digitally mapping the information content of visible-near infrared spectra of surficial Australian soils, *Rem. Sens. Environ.* 115 (6) (2011) 1443–1455.
- [4] S. Catarino, et al., Rare earths data for geographical origin assignment of wine: a Portuguese case study, *Bull. OIV* 84 (2011) 333–346.
- [5] J. Greenough, H.P. Longrich, S. Jackson, Element fingerprinting of okanagan valley wines using ICPMS: relationships between wine composition, vineyard, and wine color, *Aust. J. Grape Wine Res.* 3 (2008) 75–83.
- [6] S. Poni, et al., Grapevine quality: a multiple choice issue, *Sci. Hortic.* 234 (2018) 445–462.
- [7] J. Arno, et al., Spatial variability in grape yield and quality influenced by soil and crop nutrition characteristics, *Precis. Agric.* 13 (3) (2012) 393–410.
- [8] C. van Leeuwen, et al., Influence of climate, soil, and cultivar on terroir, *Am. J. Enol. Vitic.* 55 (3) (2004) 207–217.
- [9] D. Cozzolino, et al., In situ measurement of soil chemical composition by near-infrared spectroscopy: a tool toward sustainable vineyard management, *Commun. Soil Sci. Plant Anal.* 44 (10) (2013) 1610–1619.
- [10] A. Stevens, et al., Laboratory, field and airborne spectroscopy for monitoring organic carbon content in agricultural soils, *Geoderma* 144 (1–2) (2008) 395–404.
- [11] R.A.V. Rossel, A.B. McBratney, Soil chemical analytical accuracy and costs: implications from precision agriculture, *Aust. J. Exp. Agric.* 38 (7) (1998) 765–775.
- [12] J.M. Soriano-Disla, et al., The performance of visible, near-, and mid-infrared reflectance spectroscopy for prediction of soil physical, chemical, and biological properties, *Appl. Spectrosc. Rev.* 49 (2) (2014) 139–186.
- [13] E. Deaville, P. Flinn, Near-infrared (NIR) Spectroscopy: an Alternative Approach for the Estimation of Forage Quality and Voluntary Intake, CABI International, Wallingford, UK, 2000, pp. 301–320.
- [14] C.A.T. dos Santos, et al., A review on the applications of portable near-infrared spectrometers in the agro-food industry, *Appl. Spectrosc.* 67 (11) (2013) 1215–1233.
- [15] R.A.V. Rossel, C. Walter, Rapid, quantitative and spatial field measurements of soil pH using an Ion Sensitive Field Effect Transistor, *Geoderma* 119 (1–2) (2004) 9–20.
- [16] J. Reeves, G. McCarty, T. Mimmo, The potential of diffuse reflectance spectroscopy for the determination of carbon inventories in soils, *Environ. Pollut.* 116 (2002) 277–284.
- [17] F. Acevedo, et al., Spent coffee grounds as a renewable source of bioactive compounds, *J. Biobased Mater. Bioenergy* 7 (3) (2013) 420–428.
- [18] J.T. Klopogge, Short introduction to infrared and Raman spectroscopy, Application of Vibrational Spectroscopy to Clay Minerals and Layered Double Hydroxides 13 (2005) 1–7.
- [19] R.A.V. Rossel, T. Behrens, Using data mining to model and interpret soil diffuse reflectance spectra, *Geoderma* 158 (1–2) (2010) 46–54.
- [20] R.A.V. Rossel, et al., Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties, *Geoderma* 131 (1–2) (2006) 59–75.
- [21] K.A. Sudduth, J.W. Hummel, Soil organic-matter, cec, and moisture sensing with a portable Nir spectrophotometer, *Transactions of the Asae* 36 (6) (1993) 1571–1582.
- [22] R.A.V. Rossel, et al., Spatial modeling of a soil fertility index using visible-near-infrared spectra and Terrain attributes, *Soil Sci. Soc. Am. J.* 74 (4) (2010) 1293–1300.
- [23] D.M. Salazar, et al., Visible-near infrared spectroscopy to assess soil contaminated with cobalt, in: International Meeting of Electrical Engineering Research (ENIINVIE), Autonomous Univ Baja Calif (UABC), Campus Ensenada, Ensenada, MEXICO, 2012.
- [24] M. Lopo, et al., Classification of vineyard soils using portable and benchtop near-infrared spectrometers: a comparative study, *Soil Sci. Soc. Am. J.* 80 (3) (2016) 652–661.
- [25] M. Lopo, et al., Near infrared spectroscopy as a tool for intensive mapping of vineyards soil, *Precis. Agric.* 19 (2018) 445–462.
- [26] J.R. Rodriguez-Perez, et al., Estimating soil properties and nutrients by visible and infrared diffuse reflectance spectroscopy to characterize vineyards, *Agronomy-Basel* 11 (10) (2021).
- [27] WRB-FAO, I.W.G., IUSS Working Group WRB, World Reference Base for Soil Resources 2014, update 2015 International soil classification system for naming soils and creating legends for soil maps, in: *World Soil Resources Reports N° 106*. 2015, FAO Rome, 2015.
- [28] Agroconsultores, Coba, Carta de solos, carta do uso actual da terra e carta de aptidão da terra do nordeste de Portugal, Escala 1:100 000. Projecto de desenvolvimento rural integrado de Trás-os-Montes, Trás-os-Montes e Alto Douro University Vila Real, 1991.
- [29] G. Agroconsultores, Elaboração da carta de solos e de aptidão das terras da Zona Interior Centro (Making of the soil map and land suitability of Zona Interior Centro), Instituto de Desenvolvimento Rural e Hidráulica, Lisbon, 2004 (in Portuguese).
- [30] S. Bruckert, Analyse des complexes organo-minéraux de sols, in: *Pédologie 2. Constituants et propriétés du sol*, Masson., Paris, France: Masson, 1979, pp. 185–209.
- [31] L.V.H.J. De Leeenheer, Détermination de la teneur en carbone organique des sols. Étude critique des méthodes titrimétriques, *Pédologie* 8 (1958) 39–77.
- [32] J.M. Bremner, C.S. Mulvaney, *Nitrogen-Total*. *Agronomy* 9, American Society of Agronomy, Madison, Wisconsin, 1982, pp. 595–624.
- [33] H. Egnér, H. Riehm, W.R. Domingo, in: *Untersuchungen über die chemische Bodenanalyse als Grundlage für die Beurteilung des Nährstoffzustandes der Böden. II. Chemische Extraktionsmethoden zur Phosphor- und Kaliumbestimmung* vol. 26, Kungliga Lantbrukshögskolans Annaler, Sweden, 1960, pp. 199–215.
- [34] T. Næs, et al., A user-friendly guide to multivariate calibration and classification, in: *Multivariate Calibration and Classification*, NIR Publications, Chichester, UK, 2004.
- [35] M. Barker, W. Rayens, Partial least squares for discrimination, *J. Chemometr.* 17 (3) (2003) 166–173.
- [36] P. Geladi, B.R. Kowalski, Partial least-squares regression - a Tutorial, *Anal. Chim. Acta* 185 (1986) 1–17.
- [37] J.L. Bishop, et al., Reflectance and emission spectroscopy study of four groups of phyllosilicates: smectites, kaolinite-serpentines, chlorites and micas, *Clay Miner.* 43 (1) (2008) 35–54.
- [38] M.K.D. Rambo, E.P. Amorim, M.M.C. Ferreira, Potential of visible-near infrared spectroscopy combined with chemometrics for analysis of some constituents of coffee and banana residues, *Anal. Chim. Acta* 775 (2013) 41–49.
- [39] O. Marin-Gonzalez, et al., On-line measurement of soil properties without direct spectral response in near infrared spectral range, *Soil Tillage Res.* 132 (2013) 21–29.
- [40] F.B. de Santana, K. Daly, A comparative study of MIR and NIR spectral models using ball-milled and sieved soil for the prediction of a range soil physical and chemical parameters, *Spectrochim. Acta Mol. Biomol. Spectrosc.* (2022) 279.
- [41] E. Vaudour, et al., Predicting Key agronomic soil properties with UV-Vis fluorescence measurements combined with Vis-NIR-SWIR reflectance spectroscopy: a farm-scale study in a Mediterranean viticultural agroecosystem, *Sensors* 18 (4) (2018).
- [42] D. Cozzolino, A. Moron, The potential of near-infrared reflectance spectroscopy to analyse soil chemical and physical characteristics, *J. Agric. Sci.* 140 (2003) 65–71.
- [43] J.M. Johnson, et al., Near-infrared, Mid-infrared or Combined Diffuse Reflectance Spectroscopy for Assessing Soil Fertility in Rice Fields in Sub-saharan Africa, *Geoderma*, 2019, p. 354.