Non-invasive real-time monitoring of vineyard soils, berries and leaves with FT-NIR spectroscopy

Miguel Lopo[1], Cláudia A. Teixeira dos Santos[1], Ricardo N.M.J. Páscoa[1], António R. Graça[2], and João A. Lopes[1,3]

[1] REQUIMTE, Laboratório de Química Aplicada, Departamento de Ciências Químicas, Faculdade de Farmácia, Universidade do Porto, 4050-313 Porto, Portugal

[2] SOGRAPE Vinhos S.A., Departamento de Investigação e Desenvolvimento, Aldeia Nova, 4430-852 Avintes, Portugal

[3] Departmento de Farmácia Galénica e Tecnologia Farmacêutica, Faculdade de Farmácia, Universidade de Lisboa, Av. Prof. Gama Pinto, 1649-003 Lisboa, Portugal

Abstract. Production of high quality wines requires a permanent monitoring during the entire winemaking process. A healthy production, ensured by tailor-made strategies that will lead to consumer's satisfaction is of the utmost importance. The influence of the *terroir* characteristics on the features of a wine has always been prone to much debate amongst the wine industry. The composition of grapes is the result of the characteristics of each individual *terroir*. Soil impact on growth of the vineyard, grape variety characteristics and ultimately wine quality is well known. Current strategy for analysing soils (pedology) is based on wet chemistry methods, which are often laborious, expensive, time-consuming and may be of limited use. An efficient, high-throughput analytical method for estimating the impact of soil quality, tillage and thinning on the grapes quality is of paramount importance for the wine industry. Near infrared spectroscopy (NIRS) is a rapid, non-destructive, inexpensive and accurate analysis technique and its use in soil evaluation for discriminating different types of soil as well as soil constituents is rapidly increasing. Results obtained from direct monitoring of four Portuguese vineyards in different locations (wine appellation regions "Alentejo", "Dão", "Douro" and "Vinhos Verdes") using two different portable near-infrared spectrometers are presented. *In-situ* measurements of soils (at different depths), plant leaves and berries were performed on different stages of the ripening period. Spectral analysis was performed with chemometric methods: PCA and PLS-DA. This monitoring approach revealed to be an excellent tool for the support of a vineyard's micro-zoning process.

1. Introduction

There are many complex environmental factors within a vineyard that create the viticultural potential of a *terroir* [1]. The use of the term *terroir* frequently refers to soil, sometimes to a vineyard site as a whole, and often to a complex effect of environmental (and occasionally viticultural and oenological) factors on fruit and wine attributes. It may be defined as "a complex of natural environmental factors, which cannot easily be modified by the producer" [2]. Despite the fact, that the overall body of scientific knowledge about the effect of *terroir* on fruit and wine attributes is relatively limited [3], the *terroir* cannot be viewed in isolation from management and cultivation practices. This is indeed recognized by the International Organization for Vine and Wine (OIV) who considers that *terroir* includes specific soil, topography, climate, landscape characteristics and biodiversity features.

Of all the environmental factors affecting the dynamics and performance of the *terroir*, the different soils in which vines are grown play a major role in the vine behaviour, grape quality and wine sensory characteristics [4]. The conventional pedological soil survey is accomplished by means of extensive field observations (sometimes very subjective) followed by laboratory analysis, which adds valuable information about the soil's properties in question [5]. With most of the conventional laboratory techniques (e.g., wet chemistry), soil analysis is often extremely time consuming and can be prohibitively expensive.

The emergence of precision viticulture, based on a more detailed monitoring and control of the quantity and quality of the productivity in a vineyard, has fostered the need for the development of new technologies, combined with more time- and cost-efficient methodologies.

Near infrared spectroscopy (NIRS) has been used since the 1970s for the routine evaluation of foods and forages in the laboratory [6]. The NIR spectral region is dominated by weak overtones and combinations of vibrational bands from molecular bonds containing hydrogen attached to atoms such as nitrogen, oxygen and carbon. It often allows the analysis of several chemical properties at the same time [7,8]. In NIRS, calibration requires mathematical tools since it requires multivariate regression techniques. These techniques relate NIRS optical measurements at selected wavelengths with reference values previously obtained by conventional methods. The advantages of NIRS are the speed of analysis, simplicity in sample preparation, multiplicity of analysis and absence of using chemicals [6]. This technique can be incorporated in remote sensing devices or used to obtain multispectral images [9]. Spectral measurements can be made in plant leaves, either at-line in the laboratory, in-situ at the vineyard or through remote devices (remote sensing). During the past 20 years NIRS has been successfully applied to the determination of chemical and

physical properties of soil samples, particularly, but also of leaves and grapes.

This type of approach, through precision viticulture, involves the study of phenological phases, plant health and nutritional state, as well as production expectations [10]. With this characterization, an interesting support tool for the cultivation/replantation of new vineyards is provided, focusing on the increase of the quality and yield of produced grapes, leading, ultimately, to a more reliable and appreciated end product.

The purpose of this study was to investigate the potential of NIRS as a rapid and low-cost technique to map vineyards soil, to characterize soils based on direct in-situ measurements of vine leaves, and enabling the estimation of soil type with a superior resolution when compared to currently existing methods. The aim was also to provide the winemaker with further knowledge for planning crops as well as enabling a strategic plantation or replantation of new vineyards.

2. Methodology

2.1. Sample collection and monitoring

Four vineyards, from four different wine appellation regions (regions "Alentejo", "Dão", "Douro" and "Vinhos Verdes") were selected. *Quinta dos Carvalhais* (Mangualde, 40.556721°N–7.787247°W, Dão Wine Region) in the centre of Portugal, *Herdade do Peso* (Vidigueira, 38.141579°N–7.677813°W, Alentejo Wine Region) in the south of Portugal, *Quinta de Azevedo* (Barcelos, 41°34′12.45″N 8°32′25.05″O, Vinho Verde Wine Region) and *Quinta do Mourão*, Lamego, 41°07′23.5″N 7°48′09.5″W, Douro Wine Region) both in the north of Portugal (Fig. 1).

Soils in these vineyards, except *Quinta do Mourão*, were previously characterized by pedology reference methods which justified their selection. Vineyard soils were named according to the international soil classification system [11]. Table 1 compiles the existing types of soils in three of the vineyards (*Herdade do Peso*, *Quinta dos Carvalhais* and *Quinta de Azevedo*) (IUSS 2014). Each vineyard is divided into numbered "blocks", each "block" containing a single variety (*Vitis vinifera* cultivar). Sampling spots were defined according to the mapped soil
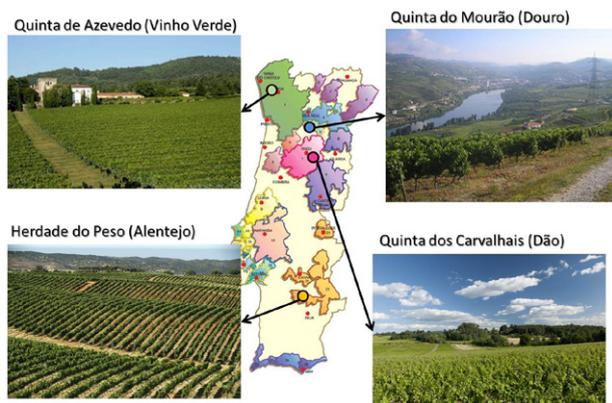
Figure 1. Four different vineyards monitored.

characteristics and grapes varieties. The sampling strategy required digging three 2-m deep holes around each spot (circa 3 m distance between holes). Soil samples were then collected at specific pre-defined depths according to each soil type. However, in many blocks, multiple soil types co-exist, making such a distinction somewhat ambiguous. Depths were chosen according to what was expected to bear more interesting results from a pedology perspective. Furthermore, several monitoring spots were identified with the objective of analysing the variability in leaves spectra of the same grape variety, growing on different soil types. At each defined sampling spot, a total of twenty leaves in five different plants around the defined spot were monitored (four monitored leaves per plant). This analysis was performed shortly after the veraison period.

2.2. Spectra acquisition

Near infrared spectra of soil samples were collected using a portable dispersive NIR spectrometer (model NIR-512, Ocean Optics, Dunedin, FL). It features a temperature-regulated 512-element InGaAs array detector effective in the 900–1700 nm wavelength range providing a spectral resolution less than 5.0 nm (full width at half maximum). Spectra were acquired in diffuse reflectance mode, using a reflectance fibre optical probe (SabIR, ThermoNicolet, Madison, WI) with an irradiation area of 0.03 cm^2. Temperature, integration time and number of scans were set to –4°C, 1.5 s and 40, respectively. Each NIR spectrum was obtained averaging the 40 spectra. Background was measured at the beginning of each analysis using Spectralon (Teflon) as reference. The OOIBase32 software (Ocean Optics, Dunedin, FL) was used for spectrometer configuration, control and data acquisition.

Near infrared leaves spectra were acquired using a FieldSpec 4 Wide-Res (ASD Inc., Boulder, CO) in diffuse reflectance mode spanning the 350–2500 nm range (Fig. 2). The spectral resolution and sampling interval were 3 and 1.4 nm for the 350–1050 nm spectral range and 10 and 2 nm for the 1000–2500 nm range, respectively. The system incorporated a reflectance contact probe (Hi-Brite, ASD Inc., Boulder, CO) with a measurement surface area equivalent to a 10 mm diameter circle, enclosing a halogenous light source (Fig. 2). A background was taken every hour with a Spectralon® disk (ASD Inc., Boulder, CO). Leaves were measured in diffuse reflectance mode directly in the plant with no cleaning process involved. Average size leaves located at one shoot of the first arm were selected.

2.3. Data analysis

Principal components analysis (PCA) [12] and partial least-squares discriminant analysis (PLS-DA) [13] were the selected data analysis methods. The former was used to extract common patterns from soil samples and outlier detection. The latter was used to develop calibration models for soil discrimination purposes. Before application of PCA and PLS-DA, spectral blocks were subjected to mean-centring as this is a requirement of both methods [14]. For PLS-DA models, the data selected for the model were divided to form a calibration (70% of the available samples) and a test (remaining 30%) set [14]. Samples

Table 1. Vineyards soil types.

Soil type		
Vineyard	**FAO-WRBSR**	**FAO-WRBSR 2014 Code**
Quinta de Azevedo	Dystric Cambisol	CM-dy
	Colluvic Cambisol	CM-co
	Eutric Cambisol	CM-eu
Quinta dos Carvalhais	Colluvic Cambisol	CM-co
	Eutric Cambisol	CM-eu
	Lithic Leptosol	LP-li
	Umbrisol (Epileptic)	UM-lep
	Cambic Umbrisol	UM-cm
	Rhodic Luvisol	LV-ro
Herdade do Peso	Calcaric Cambisol	CM-ca
	Hypereutric Cambisol (Epileptic)	CM-je-lep
	Hypereutric Cambisol	CM-je
	Colluvic Cambisol	CM-co
	Hypereutric Cambisol	CM-je
	Dystric Cambisol	CM-dy

Figure 2. FieldSpec 4 Wide-Res (ASD Inc., Boulder, CO) with reflectance contact probe (Hi-Brite, ASD Inc., Boulder, CO).

division was made randomly but ensuring that the same proportion between soil types is present in calibration and test sets to avoid unbalanced classes (soil types) in both sets [15]. The optimal number of latent variables (LVs) was estimated by leave-one-block-out cross-validation (considering blocks of 15 samples) using only the calibration set [12]. The test set is projected onto each optimized PLS-DA model and soil predictions are expressed as confusion matrices. Confusion matrices compare each known sample soil type with the corresponding NIRS prediction and entries are expressed as percentages. The objective of these matrices is not only to estimate the number of correctly predicted samples but also to diagnose which samples are being incorrectly predicted, identifying the most similar soil types in terms of NIRS. All chemometric methods and spectra processing were performed using Matlab version 7.9 (Mathworks, Natick, MA) and the PLS Toolbox version 5.5.1 (Eigenvector Research Inc., Wenatchee, WA).

3. Results

It is interesting to observe the depth effect on NIR signals and it is possible to distinguish deeper samples from corresponding surface or shallower counterparts. An initial spectral analysis was performed to clarify the effect of depth on the different soil types accounted in this work. PCA models for each soil type were developed considering the entire spectral region ($10,000–4000$ cm^{-1}) and samples were analysed in terms of scores. The soil type CM-eu was chosen to illustrate the general behaviour observed. This soil type is present in both *Quinta de Azevedo* and *Quinta dos Carvalhais*, but for practical
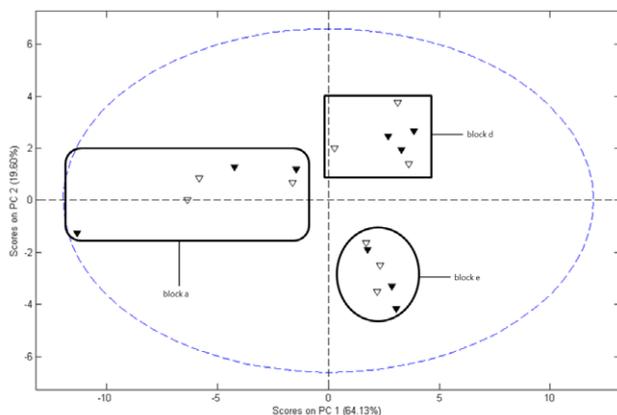
reasons only the results from *Quinta dos Carvalhais*, where CM-eu soil type is present in three different blocks, are shown (Fig. 3). Two depths for this particular type of soil were collected (0–20 cm and 20–75 cm). The PCA model reveals that no substantial difference between the different depths exists. However, a clear separation of samples can be seen, even though samples are of the same type of soil. This separation is not due to sampling depth but with the geographical provenance of samples (from different vineyard blocks). Hence, it is possible to infer that NIRS is able to separate soil samples supposedly of the same soil type appearing in different geographical zones of the same vineyard.

Furthermore, the ability of NIRS to discriminate between soil types was also investigated; PCA models were developed considering each vineyard block separately. These models revealed a good separation of samples according to soil type; *block e* of *Quinta dos Carvalhais* is shown as an example (Fig. 4).
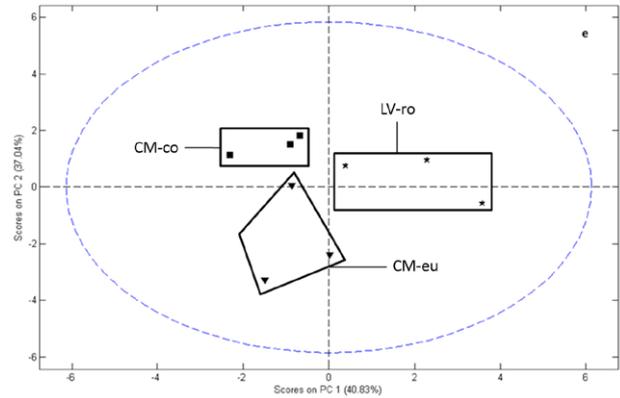
To extend the previous analysis with PCA, supervised PLS-DA models were performed. Overall, 79.3% of correct predictions for the soil type were found for the samples of *Quinta dos Carvalhais* and 72.8% of correct predictions for the samples of *Quinta de Azevedo* (Table 2).

Table 2. Confusion matrix for the PLS-DA soil discrimination model based on the NIRS method (*Quinta de Azevedo*). 72.8% of global correct predictions rate and 10 latent variables (LVs). Values are in %.

Predicted soil types	Real soil types					
	LSS	LSAO	LSAAS	LSAS	LSASQ	Sum
LSS	16.1	0.0	1.6	0.8	0.3	18.8
LSAO	0.4	9.8	0.3	0.0	2.0	12.5
LSAAS	2.8	0.0	16.5	1.3	4.4	25.0
LSAS	0.9	0.0	1.8	11.4	4.6	18.8
LSASQ	0.3	0.0	3.0	2.8	18.9	25.0
Sum	20.5	9.8	23.2	16.2	30.2	100

Figure 4. Score plots obtained from PCA models built on NIR spectra.

The analysis of leaves NIR spectral variability revealed interesting results. PCA models obtained for leaves of the same variety grown in soils with different properties were developed, considering each grapevine variety individually for two of the vineyards (*Quinta dos Carvalhais* and *Herdade do Peso*). The models suggest that this technology is able to discriminate the soil type in which the vine is planted considering only the NIR spectra of the leaves of each vine variety. PCA score plot for variety *Jaen* is shown in Fig. 5.

These promising exploratory results of leaves NIR spectra suggested the use of a supervised classification method with the objective of discriminating different soil types. Once again the supervised PLS-DA method was chosen due to its thorough application and positive results in similar situations. Results revealed that the lowest percentage of correct assignments (92.0%) was obtained for the Syrah variety at *Herdade do Peso* (model encompassing six different soil types). The highest percentage of correct assignments (98.7%) was obtained for *Antão Vaz* variety also at *Herdade do Peso* (Table 3).
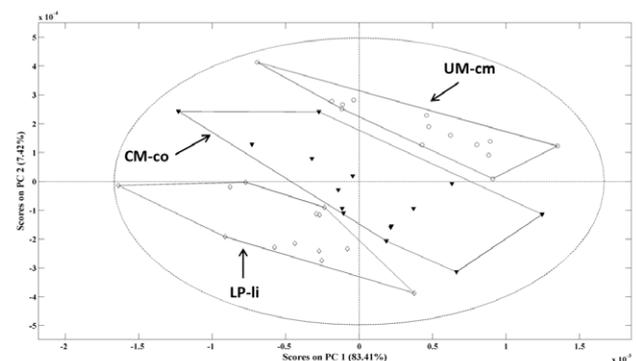

Figure 3. PCA score plot of NIR spectra obtained for CM-eu soil samples in *Quinta dos Carvalhais* (▼= 0–20 cm; ▽= 20–75 cm).

Figure 5. Score plots obtained by PCA of grapevine leaves for the *Jaen* grape variety.

Table 3. Soils discrimination PLS-DA model results for *Herdade do Peso* and *Quinta dos Carvalhais* considering varieties individually and in combination (all varieties model).

Vineyard	Variety	Soils	%[1]	Std[2]	LV[3]
Herdade do Peso	Syrah	5	92.0	4.7	7
	Arinto	3	96.4	5.4	3
	Antão Vaz	2	98.7	3.9	3
	Chardonnay	2	97.0	6.3	3
	All varieties	5	94.2	2.8	10
Quinta dos Carvalhais	Verdelho	4	97.0	4.7	5
	Encruzado	2	95.2	6.1	3
	Jaen	3	97.3	3.8	4
	Alvarinho	2	94.7	7.8	4
	Semillon	3	95.7	6.2	5
	All varieties	6	94.2	3.1	10

[1]% – Percentage of soils correct predictions for the test set.

[2]Std – Standard deviation of soils correct predictions for the test set (obtained by bootstrapping the models 500 times according to Preisner et al. [15]).

[3]LV – Latent variable.

4. Conclusion

This work demonstrated the ability of the high-throughput NIRS analytical method for discriminating between specific vineyard soil types. The proposed method was able to accurately differentiate distinct soil types, it correlates with soils characterized by pedology methods demonstrating very good prediction percentages. Furthermore, a methodology for indirectly scanning vineyard soil types through the in-situ VIS/NIR analysis of grapevine leaves of *Vitis vinifera* species varieties with a portable instrument in diffuse reflectance mode was attempted and demonstrated. Both unsupervised and supervised approaches, using respectively PCA and PLS-DA models generated very accurate soil maps with a very good agreement with previously existent soil maps for the studied vineyards. With these findings, a new support tool to assist on plantation or replantation processes, with the objective of increasing grape quality and yield was explored. Instead of having one grapevine variety in one entire block (as is usual through all vineyards around the world), detailed soil type information will allow the producer, if he so wishes, to define which grapevine varieties are more adjusted to a specific soil type or soil types transition zones.

References

[1] G. Barbeau, R. Morlat, C. Asselin, A. Jacquet, C. Pinard, J. Int. Sci. Vigne Vin, **32**, 69–81 (1998)

[2] V. Carey, MSc. Agric. Thesis, Stellenbosch University (2001)

[3] D. Tesic, D. J. Woolley, E.W. Hewett, D. J. Martin, Aust. J. Grape Wine Res., **8**, 15–26 (2002)

[4] F. Bodin, R. Morlat, Plant Soil, **281**, 37–54 (2006)

[5] E. Ben-Dor, D. Heller and A. Chudnovsky, Soil Sci. Soc. Am. J. **72,** 1113 (2008)

[6] E. R. Deaville, P. C. Flinn, Forage Evaluation in Ruminant Nutrition. CABI Publishing, Wallingford, UK, 301–320 (2000)

[7] I. Murray, Sward Management Handbook. British Grassland Society, UK, 285–312 (1993.)

[8] B.G. Osborne, T. Fearn, P. H. Hindle, Food Analysis. Longman Scientific and Technical, Essex, England, UK (1993)

[9] M. Govender, P. J. Dye, I. M. Weiersbye, E.T.F. Witkowski, F. Ahmed, *Water SA*, **35**, 741–752 (2009)

[10] F. Mazzetto, A. Calcante, A. Mena, A. Vercesi, *Precision Agric*. **11**, 636–649 (2010)

[11] International Union of Soil Science (IUSS) Working Group World Reference Base. FAO, Rome (2014)

[12] T. Naes, T. Isaksson, T. Fearn, T. Davies, NIR Publications, Chichester, UK, 39–54 (2004)

[13] M. Barker, W. Rayens, J. Chemometrics, **17**, 166–173 (2003)

[14] P. Geladi, B. R. Kowalski, Anal. Chim. Acta **185**, 1–17 (1986)

[15] O. Preisner, R. Guiomar, J. Machado, J.C. Menezes, J. A. Lopes, Appl. Environ. Microbiol. **76**, 3538–3544 (2010)