Discriminant Analysis of Geographical Origin of Cork Planks and Stoppers by Near Infrared Spectroscopy

CRISTINA PRADES,[1] ISABEL GÓMEZ-SÁNCHEZ,[2] JUAN GARCÍA-OLMO,[3] AND JOSÉ RAMÓN GONZÁLEZ-ADRADOS[4]

[1]Faculty of Agriculture and Forestry Engineering, University of Córdoba, Córdoba, Spain
[2]PRINUR, S.A., Sevilla, Spain
[3]NIR/MIR Spectroscopy Unit, Central Service for Research Support, University of Córdoba, Córdoba, Spain
[4]INIA-CIFOR, Madrid, Spain

Abstract: *The objective of this study was to assess the potential of visible and near infrared spectroscopy (VIS+NIRS) combined with multivariate analysis for identifying the geographical origin of cork. The study was carried out on cork planks and natural cork stoppers from the most representative cork-producing areas in the world. Two training sets of international and national cork planks were studied. The first set comprised a total of 479 samples from Morocco, Portugal, and Spain, while the second set comprised a total of 179 samples from the Spanish regions of Andalusia, Catalonia, and Extremadura. A training set of 90 cork stoppers from Andalusia and Catalonia was also studied. Original spectroscopic data were obtained for the transverse sections of the cork planks and for the body and top of the cork stoppers by means of a 6500 Foss-NIRSystems SY II spectrophotometer using a fiber optic probe. Remote reflectance was employed in the wavelength range of 400 to 2500 nm. After analyzing the spectroscopic data, discriminant models were obtained by means of partial least square (PLS) with 70% of the samples. The best models were then validated using 30% of the remaining samples. At least 98% of the international cork plank samples and 95% of the national samples were correctly classified in the calibration and validation stage. The best model for the cork stoppers was obtained for the top of the stoppers, with at least 90% of the samples being correctly classified. The results demonstrate the potential of VIS + NIRS technology as a rapid and accurate method for predicting the geographical origin of cork plank and stoppers.*

Keywords NIRS, absorption bands, suberin, lignin, cellulose, *Quercus suber* L, traceability, chain of custody

This study was funded by Project RTA 2006-00119-C02-00. NIR data were obtained using NIRS hardware and software provided by the NIR/MIR Spectroscopy Unit of the SCAI at the University of Córdoba.We are also grateful to Professor Ana Garrido Varo of the Faculty of Agriculture and Forestry Engineering (University of Córdoba, Spain) for her insightful comments on this work.

Address correspondence to Cristina Prades, Faculty of Agriculture and Forestry Engineering, University of Córdoba, Campus de Rabanales, E-14071, Córdoba, Spain. E-mail: ir1prloc@uco.es

Introduction

Cork oak forests are multifunctional ecosystems that fulfill an important environmental, social, and economic function. These forests are used principally for producing cork, a natural product harvested from the bark of the cork oak (*Quercus suber L*) that is employed for stopping wines and other beverages.

Cork is composed of a series of chemical compounds that are divided into structural and nonstructural components. The structural components of cork are suberin (45%–50%), lignin (20%–25%), and holocellulose (15%–20%).[1] These components are responsible for the cell wall lamellated ultrastructrure, which is characteristic of Quercus suber phellem.[2,3] Suberin is a complex polyester based on glycerol and long-chain alpha, omega-diacids and omega-hidroxyacids whose structure at the macromolecular level remain largely unknown.[4] Lignin is a guaiacyl lignin very similar to the lignin of wood from softwoods. The polysaccharides are cellulose and hemicelluloses, with the cellulose accounting for about 10% of total cork weight. The nonstructural components are extractives (8%–24%) and inorganic compounds (1%–2%). Extractives are formed mainly by aliphatic (triterpenes, n-alkanes, n-alkanoids and fatty acids) and simple and polymeric phenols.[1]

Cork tissue is very homogeneous and has an alveolar structure consisting mainly of dead parenchyma cells without meatus. It is crossed by lenticular channels that allow gas exchange of the living tissues inside the tree. However, cork is a very heterogeneous material due to variations in cell size, deformations as a result of growth tensions, the presence of lignified cells, the number and size of the lenticular channels, and the presence of channels filled with a lenticular powder known as earth. Variations in the chemical composition of cork are due to the material inside the channels, which does not contain suberin, has a higher than normal percentage of lignin, and an extractives content that can reach 31.8% compared to a mean percentage of 12.8% cork tissue.[1]

Cork grows in radial rows. Cork cells can be of two types: primary growth cells and secondary growth cells. The size of the two types of cells and the number of each type of cell in the radial row can vary due to the soil characteristics and climate of the region, the age of the tree, and phellogen activity.[5] These variations in growth influence the properties of the cork and determine the thickness of the growth ring, which can vary from 1 to 6 mm depending on the region of origin.[1] Although notable local variations can appear,[6] the largest differences in Spain are observed regionally as mean annual growth can vary from 2,6 mm/year in the north (Catalonia) to 5,4 mm/year in the south (Cádiz).[7]

Cork plank refers to the cork extracted from the cork oak that has not yet been processed, with the exception of boiling in water, after which it is stored and dried under controlled conditions of humidity and temperature. In a subsequent process to manufacture natural stoppers, the cork planks are cut and classified, maintaining the properties of the raw material practically intact. Cork chemical composition is slightly modified by industrial processing.[8,9] Due to its anisotropic structure, cork aspect and properties depend on the radial position and the section considered.[10–12]

More than two million hectares of cork trees are distributed in the Mediterranean area, specifically in southwestern Europe and North Africa. A third of all cork trees can be found in Portugal and a quarter in Spain, while the rest are distributed between Italy, France, Algeria, Tunisia, and Morocco. The cork-producing countries studied here (Portugal, Morocco, and Spain) account for 77% of the total cork oak forest area and more than 75% of cork production.[13] In Spain, the three regions under study (Andalusia, Catalonia, and Extremadura) account for 93% of the cork oak forest area and 90% of cork production.[14]

Cork from sustainable managed forests, which are certified according to criteria established at European Lisbon conference (1998), is now being demanded by many companies. The mechanism for tracking certified material from the forest to the final product through the chain of custody certification,[15] started in the cork sector in 2007, based on the standards established by the Forest Stewardship Council (FSC)[16] and the Programme for the Endorsement of Forest Certification (PEFC). Since that time, an increasing number of companies have become certified all over the world, particularly in countries of Europe and in the United States.

For this reason, it is of interest to develop an objective instrument that can be used to verify the regional origin of cork planks in industry, thereby ensuring the traceability of cork stoppers and improving the chain of custody certification for cork products.

Qualitative analysis by NIRS chemometric techniques has been used to distinguish groups of samples with common characteristics and successfully applied to verify the origin of agricultural products.[17] In the forest sector, the first results concerning NIR spectra and chemical properties of wood were presented 20 years ago. These studies show the suitability of NIRS for the quantitative assessment of the physical, mechanical, and chemical properties of wood such as longitudinal modulus of elasticity, lignin content, cellulose content, wood density, moisture, wood age[18–22] and other parameters of interest for wood.[23]

NIRS has also been used to identify wood species of different genera, wood species of the same genus, and distinguish between different origins of the same species.[24–27]

As regards cork, quantitative analysis by NIRS has been used to measure the visual quality, porosity, and moisture of cork planks with promising results, while qualitative analysis by NIRS has been used to determine their geographical origin at regional level,[28] achieving much better results than when using other techniques such as FT-IR or 13C-NMR.[29–30]

The aim of this study is to evaluate the suitability of NIRS technology for authenticating the geographical origin of cork plank and cork stoppers. To do so, multivariate classification models at the national and international scale were developed using a training set composed of cork planks and stoppers from the most important cork-producing countries and regions in Spain and the world.

Material and Methods

Experimental Material

The cork planks used in the study came from the "CORKASSESS" collection and were supplied by the Center for Forestry Research (CIFOR) of the National Institute of Agricultural Research (INIA) of Spain. The CORKASSES catalog was created in order to cover as much variability as possible in environmental conditions in Quercus suber areas. The main production areas of each country were sampled, with the number of specimens from each population ranging from 5 to 20.[31] A total of 479 cork planks were studied. The cork planks came from Morocco (60 samples), Portugal (240), and Spain (179). In turn, the 179 Spanish samples came from three geographical areas: Catalonia (60 samples), Extremadura (59), and Andalusia (60).

Following standard industry procedure, the samples were boiled in water for one hour at their place of origin. The training set used in this study is representative of the quality classes that are employed commercially. The cork planks were classified following the

traditional method of visual inspection, which combines 4 classes of visual quality (top, medium, commercial, and refuse) and three classes of thickness (<27 mm; 27–40 mm; >40 mm).

In addition to the cork planks, 90 natural cork stoppers measuring 24 mm in diameter and 44 mm in length were used in the study. The stoppers were produced using cork from two geographical areas in Spain: Andalusia (45 stoppers) and Catalonia (45 stoppers). Following a standard process, the stoppers were extracted from the cork planks using a cylindrical blade to cut the cork in an axial direction. The stoppers were then sanded to obtain a smooth surface and the final stopper size. Aside from sanding the stoppers, they did not undergo any other surface treatment. The stoppers were of "top," "medium," and "commercial" quality.

Before obtaining the spectra, the cork plank and cork stopper samples were stabilized in a Dycometal CCK81 climatic chamber for one week at a temperature of 20°C and 65% relative humidity, which is equivalent to a hygroscopic moisture balance of 6%.

Instrumentation and Collection of Spectra

The samples were scanned using a remote reflectance fiber optic probe (NR-6539-A) connected to a Foss-NIRSystems 6500 spectrophotometer equipped with autogain detectors: one for 400 to 1098 nm (known as the VIS region) and another for 1100 to 2500 nm (known as the NIR region).

Reflectance spectra were collected every 2 nm from 400 to 2500 nm. The absorbance data were obtained and stored as log (1/R), where R is the reflectance. All the data were stored using WinISI II software version 1.50 (Infrasoft International, Port Matilda, PA, USA). A reference spectrum was recorded before analyzing each sample.

Four types of spectra were obtained as described below (Figure 1):

1. Cork plank in the transverse section of the cork using a rectangular mask (4 cm^2);
2. Top of the stopper in the transverse section of the cork using a circular mask (3.14 cm^2);
3. Body of the stopper in the tangential section of the cork using a rectangular mask (4 cm^2);
4. Body of the stopper in the radial section of the cork using a rectangular mask (4 cm^2).

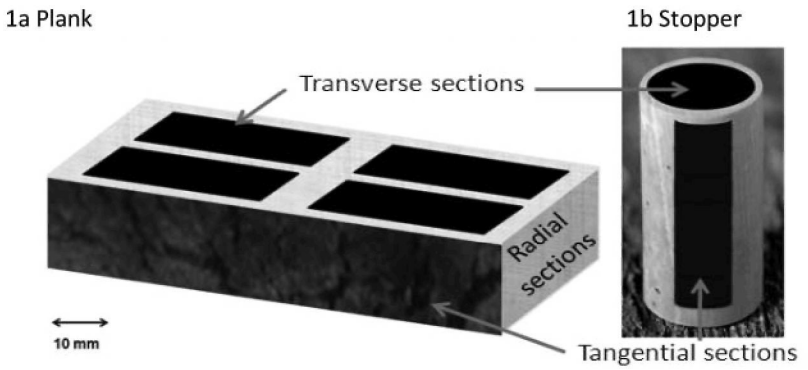

Figure 1. Areas selected for NIR analysis in samples of cork plank (transverse section) (1a), cork stopper tops (transverse section), and cork stopper bodies (radial and tangential sections) (1b).

The spectra for the cork planks were obtained in the transverse section, with number of spectra varying depending on the thickness of the cork. A window size of 10 mm × 40 mm was used to obtain at least two spectra per sample; one on the right side and another on the left side of the section. Since the size of the window did not coincide with the original size of the NR-6539-A probe, a black polystyrene plastic mask was fixed to the window of the probe. Given that the thickness of the samples varied substantially, 2 to 10 spectra were obtained per sample, half on the left side and half on the right side of the sample. These spectra were then averaged to obtain a mean single spectrum, which was used to determine relationship to geographical origin (Figure 1a).

The spectra of the cork stoppers were obtained using two different masks (Figure 1b). A circular mask measuring 20 mm in diameter was used for the stopper tops, while a rectangular mask measuring 10 mm × 40 mm (the same mask used to obtain the spectra of the cork planks) was used for the stopper bodies. Spectra were obtained of both tops of each stopper (transverse section of the cork). Two spectra were also obtained of the stopper bodies, one of the face perpendicular to the growth rings (radial section of the cork) and another of the face parallel to the growth rings (tangential section of the cork). A total of four spectra per stopper were obtained, but could not be combined as the absorbance surfaces were different. The two spectra of the stopper body and the two spectra of the stopper top were averaged to obtain, respectively, the mean spectrum that was used to determine relationship to geographical origin.

Qualitative Chemometric Analysis

The training set of cork plank samples from different countries and regions of Spain and the stoppers from Andalusia and Catalonia were qualitatively analyzed using WinISI II ver.1.50 software. The total population of spectra was divided randomly and proportionally to geographical origin into a calibration and validation training set, with 2/3 and 1/3 of the samples in each set, respectively.

The origin of the samples was determined using classification models developed by discriminant partial least square (DPLS),[32] following the same methodology used in the previous study.[26] Briefly, DPLS correlates spectral variations (X) with defined classes (Y), and maximizes the covariance between the two types of variables, with the purpose of developing a classification model.[33–34] An unknown sample is classified according to the value predicted by the DPLS model, which ideally should be close to the values used to codify the class. In this study, the cork plank was classified into three classes and the cork stoppers into two classes. Class membership was assigned a value of 2 and no class membership a value of 1. The sample was classified in the class with a value closest to 2.

Before developing the classification models, the spectral data were subjected to different signal pre-treatments. Scatter correction was accomplished using the standard normal variate and detrend (SNV+DT) algorithm.[35] The spectra were transformed by using different ranges of the 400–2500 nm spectrum and different combinations of derivative math treatments applied to the spectral data. WinISI derivative math treatments are denoted by a four-digit notation (a, b, c, d), where a is the derivative order, b is the derivative gap, c is the smoothing segment, and d is the second smoothing segment.[36] Moreover, the influence of different wavelength regions was evaluated, and various qualitative models were applied successively on the full spectrum (VIS + NIR; 400–2500 nm), the NIR region (NIR; 1100–2500 nm), and the 400–2200 region, as this is the region recommended by the manufacturer of the optic fiber probe to obtain spectra with a suitable signal-to-noise ratio.

Two classification models for the cork planks and for the stoppers were developed as follows:

1. International-scale models to discriminate between cork plank from Morocco, Portugal, and Spain;
2. National-scale models to discriminate between cork plank from Andalusia, Catalonia, and Extremadura;
3. National-scale models to discriminate between stoppers from Andalusia and Catalonia
   a. Using spectra obtained from stopper tops;
   b. Using spectra obtained from stopper bodies.

The models obtained were evaluated by external validation employing one-third of the samples that were not used in the calibration. The qualitative analysis methods were validated in a traditional manner by studying the percentage of samples that were classified correctly or incorrectly.[34] The number of samples used to calibrate and validate each model is shown in Table 1.

Results and Discussion

Spectral Study

Figure 2 shows the mean spectra obtained for the transverse section of the cork planks and the mean spectra of the stopper tops and bodies in the visible region (Figure 2a) and the NIR region (Figure 2b). The mean spectrum of the cork planks is very close to the mean spectrum of the stopper bodies (in both cases the window measures 4 cm^2). The displacement of the two spectra may be due to the different absorption of the cork in the different sections where the spectra were obtained (transverse for cork plank, radial and tangential for stoppers bodies). The mean spectrum of the stopper tops corresponding to the transverse section of the cork show the same bands, although their intensity, log (1/R), is lower due to the fact that the absorbance surface (3.14 cm^2) is also smaller.

Moreover, minimum differences were found between the spectra obtained for the cork plank and stoppers in the NIR region, with a clear superposition of spectra and differences of around ±2 nanometers (nm) between the absorption peaks of the mean spectra. The absorption peaks corresponding to the —CH groups mostly contribute to the bands around 1215, 1730, 2146, 2310, 2354 nm, while others due to the —OH groups contribute to the bands around 1450 and 1930 nm[28,37] (Figure 2b). Specifically, bands 1215 and 1730 correspond to —CH groups in the second and first overtone, respectively, and are associated to a $—CH_2$ structure. Bands 1450 and 1930 correspond to the first overtone and combination bands of the —OH group, which is mostly present in water. The 2146 band is chiefly associated to the —CH and C=O combination band.[36] The 2308–2500 nm region is assigned to —CH combination bands of asymmetric and symmetric stretch plus one deformation mode in the —CH and $—CH_2$ structures[39] (Table 2).

The bonds in the main compounds of cork were clearly identified in the absorption peaks: suberin (—CH, $—CH_2$, $—CH_2OH$, —COOH); cellulose (—OH, $—CH_2OH$); lignin (OH phenols, $—CH_2—CH_2OH$).

The main components of cork are identified in the literature in the following bands: oil (2310 nm), cellulose (1490, 1780, 1820, 1930, 2335, 2347, 2352, 2488 nm), and lignin (1170, 1410, 1417, 1420, 1440, 1685 nm).[37] Bands 2310, 1930, 2352 and 1440 nm are clearly identified in the raw spectra.

Table 1
Number of samples used in the calibration and external validation of cork planks and stoppers at international and national scale

	Number of samples					
	Cork planks				Stoppers	
	International scale		National scale		National scale	
Geographical origin	Classification	Validation	Classification	Validation	Classification	Validation
Morocco	40	20				
Portugal	160	80				
Spain	119	60				
Andalusia			40	20	30	15
Catalonia			40	20	30	15
Extremadura			39	20		
Total	319	160	119	60	60	30

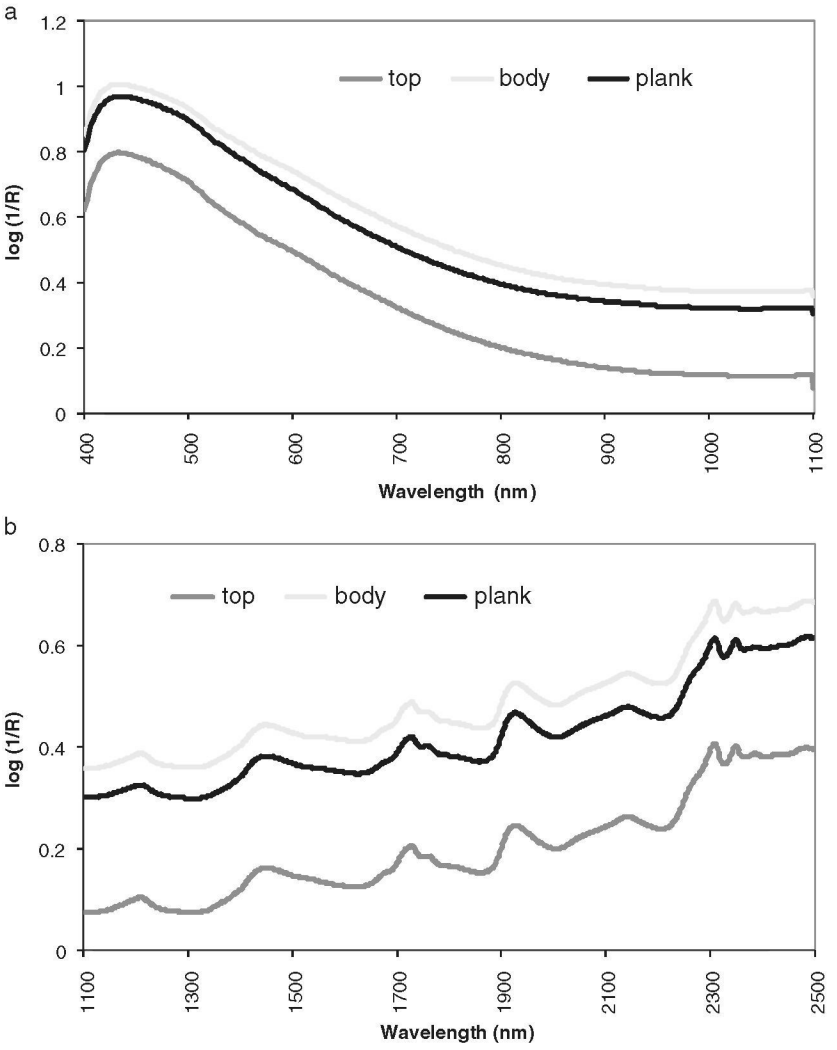

Figure 2. Mean spectrum obtained for cork planks, cork stopper tops, and cork stopper bodies (2a: visible region; 2b: NIR region).

In the visible region (400–1100 nm), the differences between the mean spectra show the same patterns as those described for the NIR region. The spectrum shows maximum absorption at 450 nm and a shoulder between 500 and 600 nm due to the dark brown color of these types of samples.[40]

For a more in-depth spectral analysis, Figure 3 shows the means and standard deviations of the first derivative of the spectra of the cork plank, the stopper tops, and the stopper bodies for all the cases. Separate graphs are shown for the visible region (Figures 3a, 3c, 3e) and the NIR region (Figures 3b, 3d, 3f).

The graphic analysis confirms the homogeneity of the spectra in both regions, whose graphs are very similar. As in the non-derived spectra, the differences between the absorption peaks of the mean values of the first derivative are around $\pm$ 2 nm.

Table 2
Assignments of absorption bands and functional grouping to main compounds of cork [34,35]

Mean absorption band					
Interval (nm)	Peak (nm)	Functional grouping	Structure		Cork compound
1158–1338	1215	—CH	Second overtone	$-CH_2$ group	
1424–1612	1450	—OH	First overtone and combination bands	—OH group	Cellulose
1718–1822	1730	—CH	First overtone	$-CH_2$ group	
1882–1960	1930	—OH	First overtone and combination bands	—OH group	Cellulose
1960–2152	2146	—CH	Combination bands	—CH and C=O	
2308–2500	2310 2354	—CH —CH	Combination bands of asymmetric and symmetric stretch	—CH and $-CH_2$	Suberin Cellulose

The main peaks of the first derivative correspond to bands 1150, 1400, 1690, 1900, 2230, 2300, 2344 and 2436 nm in the three cases, with two marked peaks in band 2300 nm. It can also be observed that the absorption peaks of the first derivative are displaced in relation to the raw spectrum (1150–1215 nm, 1400–1450 nm, 1690–1730 nm, 1900–1930 nm, 2230–2310 nm, 2300–2310 nm, 2344–2354 nm, 2436–2490 nm), showing a parallel relation between both.

The standard deviations with the highest values correspond to bands 1900 and 2310, those with intermediate values to bands 2282, 2330 and 1400 nm, and those with the lowest values to bands 2436, 2230, 1150 and 1690 nm.

As expected, the greatest variability corresponds to the bands with the highest absorbance in the raw spectra, specifically bands 1930 and 2310 nm, followed by bands 1450, 1215 and 1730 nm. Consequently, the greatest variability between spectra could be chiefly due to differences in the content of groups —OH and —CH or C=O.

The highest standard deviation in the visible region occurs for 400 nm and progressively diminishes to 490 nm, reaching another maximum between 500 and 600 nm (Figure 3a, 3c, 3e). From that point onwards, the longer the wavelength the smaller the standard deviation.

Study of the Models

Models were developed for the entire spectral range (400–2500 nm) and using only the NIRS region (1100–2500 nm). Although the results were similar, the best models were obtained in the 400 to 2200 nm region for all cases. With just one exception, the best pre-treatment was (SNV+DT), while the best mathematical treatment was the second

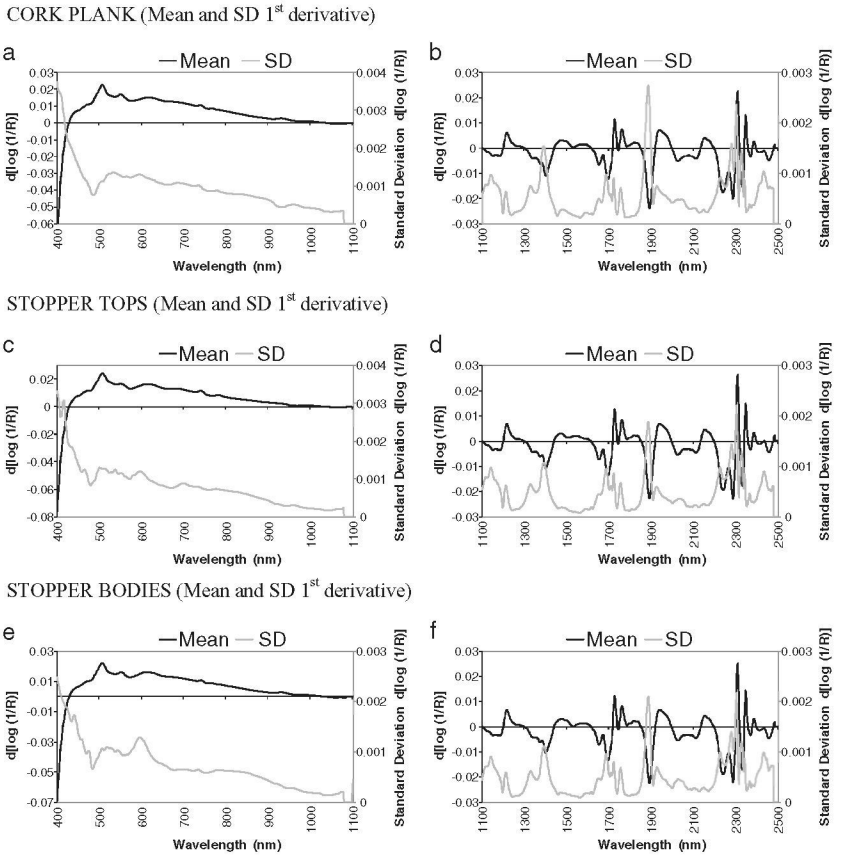

Figure 3. Mean value and standard deviation of the first derivative of the spectra obtained for cork planks (3a: visible region; 3b: NIR region), cork stopper tops (3c: visible; 3d: NIR), and cork stopper bodies (3e: visible; 3f: NIR).

derivative. The classification and evaluation errors were below 5% for the cork plank and 10% for the cork stoppers.

As visible and NIR spectra are highly sensitive to the surface quality and roughness of the samples, they can be classified according to their surface quality and differences in anatomical and structural characteristics rather than geographical origin. To test this possibility, the results obtained for geographical origin, visual quality, and thickness using NIRS analysis were compared considering visual quality and thickness as indicators of surface quality and growth rings, respectively.

The calibration of the visual quality of a training set of 170 planks, in the transversal section, by NIRS quantitative analysis showed a cross validation coefficient of determination (r^2) of 0.45 for 8 visual quality classes and of 0.47 for 3 visual quality classes.[28]

To evaluate the possible influence of ring thickness in the classification, a model was developed for the thickness of the planks, obtaining a classification error of 33.0% (transversal section, derivative 2,5,5,1, spectral range 400–2200 nm and pre-treatment SNV+DT), a larger error than that obtained in the classification by geographical region.

Although the factors analyzed may have an influence on the results of the cork plank classification according to their geographical origin, they do not explain the results obtained.

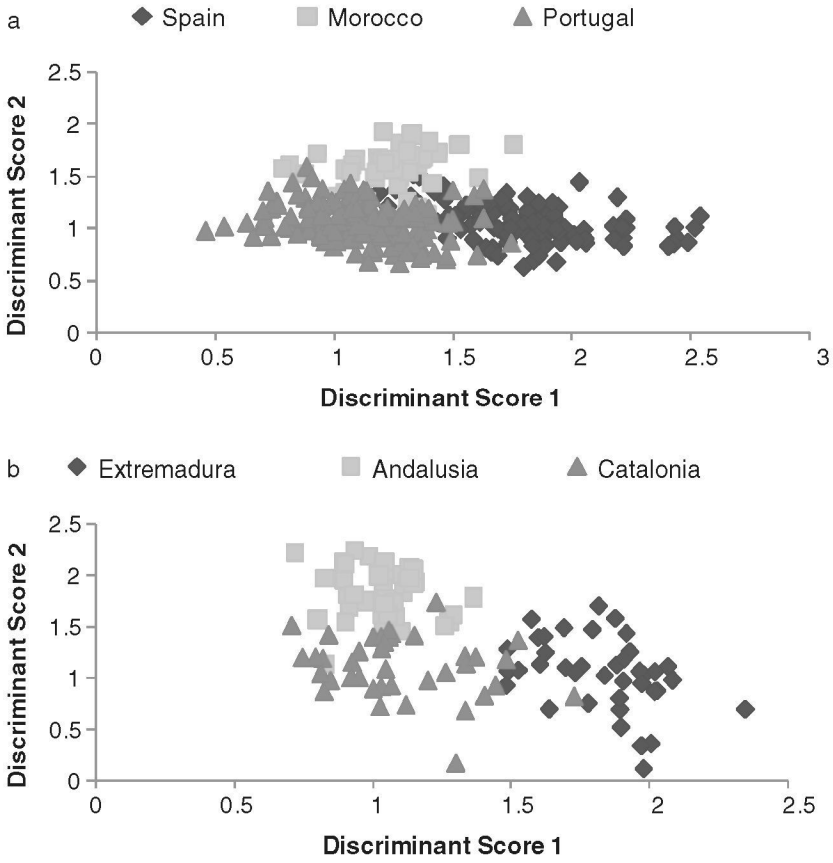

Figure 4. Discriminant PLS plot of spectral populations of cork plank in the models selected at international (4a) and national (4b) scale.

Other factors, or a combination of factors, are detected by NIRS technology, thus permitting the origin of the samples to be classified correctly.

Figure 4 shows the two-dimensional distribution of the spectral populations of cork plank in the selected models at international scale (4a) and national scale (4b) as well as the separation between populations. In what follows, the best models for each case are described in greater detail.

International-scale Models to Discriminate Between Cork Plank from Morocco, Portugal, and Spain. The classification and validation errors in the best international-scale model are equal with a value of 1,9 (Table 3, Figure 4a). This was the only case in which the best model was obtained without pre-treatment (without scatter correction). All of the samples from Portugal, 95% from Spain and 100% from Morocco were classified correctly.

The Moroccan and Portuguese samples have been classified more accurately. The greater variance observed in the sample population from Spain could be due to the less uniform ecological characteristics of the cork oak forests in the country as they are located in northwest (Catalonia), southeast (Extremadura), and southern Spain (Andalusia) and the different forest management practices employed in the three regions; a fact that would also explain the good capacity of the national-scale model to discriminate geographical origin.

Table 3
Models for the classification of cork plank at international and national scale; spectral range 400–2200 nm

Cork plank model	International(*)			National(**)		
Geographical origin	Morocco	Portugal	Spain	Andalusia	Catalonia	Extremadura
Classification						
No. samples misclassified	0/40	0/160	6/119	3/40	2/40	0/39
Classification error (%)						
Partial	0	0	5	7.5	5	0
Total	1.9			4.2		
Validation						
No. samples misclassified	1/20	1/80	1/60	0/20	1/20	0/20
Validation error (%)						
Partial	5	1.3	1.7	0	5	0
Total	1.9			1.7		

Derivative 2,5,5,1 (*)(**); Spectral range 400–2200 (*)(**); Pre-treatment None (*); Pre-treatment SNV+DT(**).

National-scale Models to Discriminate Between Cork Plank from Andalusia, Catalonia, and Extremadura. The best classification model at a national scale was obtained for Extremadura, with 100% of the samples correctly classified, followed by Catalonia (95%) and Andalusia (92.5%) (Table 3, Figure 4b). The total classification error was 4.2%. The results of the external evaluation were better than the results of the classification, with 100% of the samples from Andalusia and Extremadura and 95% of the samples from Catalonia correctly classified. The total validation error was 1.7%.

The correct classification of all the samples from Extremadura, both in the calibration and the evaluation stage, suggests a central geographical position of these samples corresponding to the geographical and ecological characteristics of the cork oak forests from which they come. The discrimination of the Catalonian samples may be related to the growth effect since the cork from this region is characterized by thinner growth rings.

The increase in the error with respect to the international model may also be due to the fact that the training set is smaller, thus reducing variability.

Models for Natural Cork Stoppers. The best calibration model (Table 4) obtained for the cork stoppers using the spectra of the cork tops correctly classified 96.7% of the samples of both origins, with a total classification error of 3.3%. When validating the model, 96.7% of the samples from Andalusia and 86.7% of the samples from Catalonia were classified correctly, with a total validation error of 10%.

The best calibration model (Table 4) obtained for the cork stoppers using the spectra of the cork bodies correctly classified 93.3% of the samples from both regions, with a total classification error of 6,7%. When validating the model, 100% of the samples from Andalusia and 80% of the samples from Catalonia were classified correctly, with a total classification error of 10%.

Table 4
Models for the classification of cork stopper origin; spectral range 400–2200 nm

Stopper model	Stopper tops(*)		Stopper bodies(*)	
Geographical origin	Andalusia	Catalonia	Andalusia	Catalonia
Classification				
No. samples misclassified	1/30	1/30	2/30	2/30
Classification error (%)				
Partial	3.3	3.3	6.67	6.67
Total	3.3		6.67	
Validation				
No. samples misclassified	1/15	2/15	0/15	3/15
Validation error (%)				
Partial	3.3	13.3	0.0	20.0
Total	10.0		10.0	

Derivative 2,5,5,1 (*); Spectral range 400–2200 (*); Pre-treatment SNV+DT (*).

The greater homogeneity of the stoppers that underwent pre-treatment and a continuous selection process may explain the fact that the errors obtained are greater than those found for the cork planks, although the origin of cork stoppers can be determined with total errors of around 10%. As in the above case, the increase in the error may be due to the small sample size and lower variability.

The classification is better when using the spectra of the stopper tops corresponding to the transverse section of the cork, the same section that is used for classifying cork planks.

Discussion

To provide more in-depth information about this analysis and determine the NIRS regions where the DPLS factors of the discriminant models have more weight, Figure 5 shows the discrimination coefficients of the cork plank classification models at international scale.

The NIRS regions (Figure 5b) where the main terms of the equations have a higher value correspond to the following five intervals: 1158–1338 nm, 1424–1612 nm, 1718–1822 nm, 1882–1960 nm, and 1960–2152 nm. These regions include the most important absorption bands with the highest variability for the cork planks and stoppers studied, specifically bands 1200, 1450, 1730, 1930, and 2140 nm (Figure 2).

The values of the discrimination coefficients in the VIS region (Figure 5a) are small until reaching the 700 nm band, when they begin to increase until reaching a maximum of around 1000 nm. The coefficients of the discriminant models on the visible region (Figure 5) are close to zero, thus confirming that the spectral variability observed in the color (Figure 3a, 3c, and 3e) does not contribute to the discriminant power of the model.

In short, the absorptions provide information on the X-H bonds of the organic chemical compounds of cork, where X denotes carbon, oxygen, and nitrogen atoms.[40,41] This suggests that cork possesses some component or characteristic that explains its classification by origin and whose characteristics are reflected in these intervals. The major chemical components of cork (suberin, lignin, cellulose) do not appear to explain this behavior given that their relative proportion is quite homogenous. In the specific case of suberin, the only

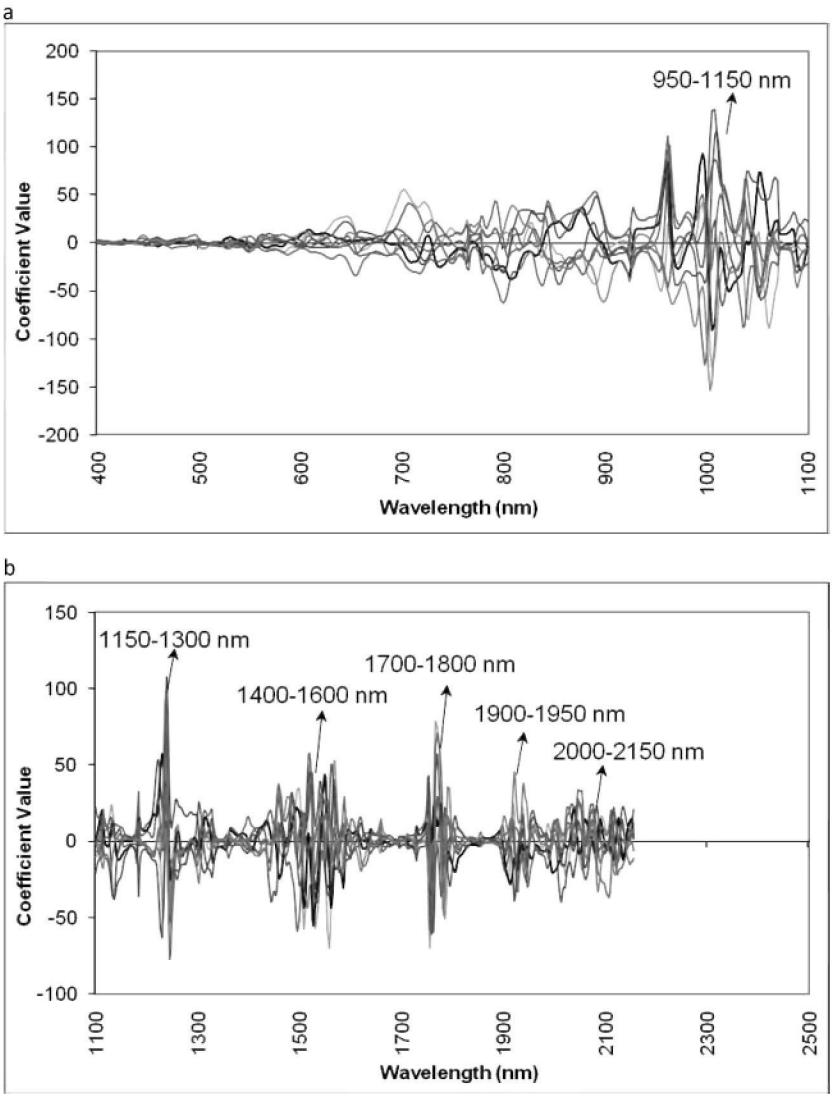

Figure 5. Weights of the bands of the models obtained for cork planks at international scale (5a: visible region; 5b: NIR region).

reference found in the literature does not report significant differences for different sites in Portugal[30] and Spain.[42] In contrast, some minor components such as polyphenols have shown this behavior.[43] A comparative study between the chemical composition of cork and NIRS spectra would be needed to confirm the hypotheses on the relationship between the chemical composition of cork and its geographical origin.

Other factors that undoubtedly influence discrimination have to do with the environmental conditions in which cork oaks grow. Variability across cork oak forests leads to significant variations in the thickness of ring growth and other anatomical and structural characteristics of cork tissue, some of which can be seen on the surface of the cork. It is difficult to explain the results obtained for geographical origin, but they are most likely due to a combination of factors that require further investigation. Surface factors such as visual

quality and thickness do not explain the results obtained for the cork plank classification according to geographical origin.

The models have obtained very satisfactory results in all cases, while the results of the external validation remain within the same range, thus confirming the validity of the equations. Although the classification and evaluation errors were under 10%, they were higher in the adjusted models using fewer samples.

Conclusions

The above results indicate that NIRS technology is capable of differentiating between cork samples from a variety of geographical origins with errors under 5% for cork planks and 10% for cork stoppers. No other method capable of determining the origin of cork has been described in the literature. The models provide good results regardless of the scale used (national or regional) or the industrial processing stage (cork plank or cork stoppers). The main applications of this technique are to improve traceability, an essential process for certifying sustainable forest management and controlling quality and origin in cork stopper manufacturing industries. The implementation of this technique at the industrial scale will require the use of training sets tailored to a specific aim, as well as deeper knowledge of the physical, chemical, and structural characteristics related to the spectral ranges on which all classification models are based.

These results should be confirmed by developing robust models that include more variability such as other geographical origins or samples collected from other harvesting seasons.

References

1. Pereira, H., ed. *Cork: Biology, Production and Uses*; Elsevier: Oxford, UK, 2007.
2. Sen, A.; Miranda, I.; Santos, S.; Graca, J.; Pereira, H. The chemical composition of cork and phloem in the rhytidome of Quercus cerris bark. Industrial Crops and Products **2010**, *31*(2), 417–422.
3. Teixeira, R.T.; Pereira, H. Suberized cell walls of cork from cork oak differ from other species. Microscopy and Microanalysis **2010**, *16*(5), 569–575.
4. Graça, J.; Santos, S. Suberin: A biopolyester of plants' skin. Macromolecular Bioscience **2007**, *7*(2), 128–135.
5. Sánchez-González, M.; Calama, R.; Canellas I.; Montero, G. Variables influencing cork thickness in spanish cork oak forests: A modelling approach. Annals of Forest Science **2007**, *64*(3), 301–312.
6. Caritat, A.; Gutiérrez, E.; Molinas, M. Influence of weather on cork-ring width. Tree Physiology **2000**, *20*(13), 893–900.
7. González-Adrados, J. R.; González Hernández, F.; Calvo, R. La predicción del calibre del corcho al final del turno y su aplicación al muestreo de la producción. Inv. Agr. Sis. y Rec. For. **2000**, *9*(2), 363–374.
8. Conde, E.; Cadahia, E.; García-Vallejo, M. C.; Fernandez de Simon, B.; González-Adrados, J. R. Low molecular weight polyphenols in cork of *Quercus suber*. J. Agric. Food Chem. **1997**, *45*, 2695–2700.
9. Conde, E.; García-Vallejo, M. C.; Cadahia, E. Variability of suberin composition of reproduction cork from *Quercus suber* throughout industrial processing. Holzforschung **1999**, *53*, 56–62.
10. Anjos, O.; Pereira, H.; Rosa, M.E. Tensile properties of cork in the tangential direction: Variation with quality, porosity, density and radial position in the cork plank. Materials and Design **2010**, *31*, 2085–2090.

11. Anjos, O.; Pereira, H.; Rosa, M.E. Characterization of radial bending properties of cork. Eur. J. Wood Prod. **2011a**. In press. DOI: 10.1007/s00107-010-0516-9
12. Anjos, O. Tensile properties of cork in axial stress and influence of porosity, density, quality and radial position in the plank. Eur. J. Wood Prod. **2011**, *69*, 85–91.
13. Barberis, A.; Dettori, S.; Filigueddu, M.R. Management problems in Mediterranean cork oak forest: Post fire recovery. Journal of Arid Environments **2003**, *54*, 565–569.
14. Montero, G.; Cañellas, I. *El Alcornoque: Manual de Reforestación y Cultivo*; Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA), Ediciones Mundi-Prensa: Madrid, Spain, 2003.
15. Wingate, K. G.; McFarlane, P.N. Chain of custody and eco-labelling of forest products: A review of the requirements of the major forest certification schemes. International Forestry Review **2005**, *7*(4), 342–347.
16. Schepers, D. H. Challenges to legitimacy at the Forest Stewardship Council. Journal of Business Ethics. **2010**, *92*(2), 279–290.
17. Roberts, C.A.; Workman, J.; Revves, J.B. *Near-Infrared Spectroscopy in Agriculture*; C.S.A.S.S.S.O. America: Madison, WI, USA, 2007.
18. Schimleck, L.R.; Raymond, C.A.; Beadle, C.L.; Downes, G.M.; Kube, P.D.; French, J. Applications of NIR spectroscopy to forest research. Appita Journal. **2000a**, *53*(6), 458–464.
19. Schimleck, L.; Evans, R.; Ilic, J. Application of near infrared spectroscopy to the extracted wood of a diverse range of species. Iowa Journal. **2003**, *24*(4), 429–438.
20. Yeh, T.F.; Chang, H.M.; Kadla, J.F. Rapid prediction of solid wood lignin content using transmittance near-infrared spectroscopy. Journal of Agricultural and Food Chemistry **2004**, *52*(6), 1435–1439.
21. Jones, P.D.; Schimleck, L.R.; Peter, G.F.; Daniels, R.F.; Clark III, A. Nondestructive estimation of Pinus taeda L. wood properties for samples from a wide range of sites in Georgia. Canadian Journal of Forest Research-Revue Canadienne de Recherche Forestiere **2005**, *35*(1), 85–92.
22. Via, B.K.; So, C.L.; Eckhard, L.G.; Shupe, T.F.; Groom, L.H.; Stine, M. Response of near infrared diffuse reflectance spectra to blue stain and wood age. Journal of near Infrared Spectroscopy **2008**, *16*(1), 71–74.
23. Tsuchikawa, S. A review of recent near infrared research for wood and paper. Applied Spectroscopy Reviews **2007**, *42*(1), 43–71.
24. Brunner, M.; Eugster, R.; Trenka, E.; BerganminStrotz, L. FT-NIR spectroscopy and wood identification. Holzforschung **1996**, *50*(2), 130–134.
25. Schimleck, L.R.; Michell, A.J.; Raymond, C.A. Effect of site on the within-tree variation of wood properties of eucalypts as determined by NIR spectroscopy and multivariate analysis. Appita Journal **2000b**, *53*(4), 318–322.
26. Schimleck, L.R.; Kube, P.D.; Raymond, C.A.; Michell, A.J.; French, J. Extending near infrared reflectance(NIR) pulp yield calibrations to new sites and species. Journal of Wood Chemistry and Technology **2006**, *26*(4), 299–311.
27. Tsuchikawa, S.; Inoue, K.; Noma, J.; Hayashi, K. Application of near-infrared spectroscopy to wood discrimination. Journal of Wood Science **2003**, *49*(1), 29–35.
28. Prades, C.; Garcia-Olmo, J.; Romero-Prieto, T.; García de Ceca, J.L.; Lopez-Luque, R. Methodology for cork plank characterization (Quercus suber L.) by near-infrared spectroscopy and image analysis. Meas. Sci. Technol. **2010**, *21*(6).
29. Lopes, M.H.; Neto, C.P.; Barros, A.S.; Rutledge, D.; Delgadillo, I.; Gil, A.M. Quantitation of aliphatic suberin in Quercus suber L. cork by FTIR spectroscopy and solid-state C-13-NMR spectroscopy. Biopolymers **2000**, *57*(6), 344–351.
30. Lopes, M.H.; Barros, A.S.; Neto, C.P.; Rutledge, D.; Delgadillo, I.; Gil, A.M. Variability of cork from Portuguese Quercus suber studied by solid-state C-13-NMR and FTIR spectroscopies. Biopolymers **2001**, *62*(5), 268–277.
31. Pereira, H. The project CORKASSESS: Field assessment and modelling of cork production and quality. In: H. Pereira (Editor), Cork Oak and Cork—Sobreiro e Cortiça, Centro de Estudos Florestais, 87–92. 1997, Lisbon, Portugal.

32. Kowalski, B., ed. *Chemometrics: Mathematics and Statistics in Chemistry*; D. Reidel Publishing Company: Dordrecht, Holland, 1984.
33. Vandenginste, B.G.M., D.L. Buydens, M.C.; De Jong, S.; Lewi, P.J.; Smeyers-Verbeke, J., ed. *Handbook of Chemometrics and Qualimetrics: Part B*; Elsevier Science BV: Amsterdam, The Netherlands, 1998.
34. Massart, D.L.; Vandergiste, B.G.M.; Buydens, L.M.C.; De Jong, S.; Lewi, P.J.; Smeyers-Verdeke, J., ed. *Handbook of Chemometrics and Qualimetrics, Vol. 2*; Elsevier Science: Amsterdam, 1997.
35. Barnes, R.J.; Dhanoa, M.S.; Lister, S.J. Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. Applied Spectroscopy **1989**, *43*(5), 772–777.
36. Shenk, J.S.; Westerhaus M.O., eds. *Routine Operation, Calibration, Development and Network System Management Manual*; NIRSystems, Inc.: Silver Spring, MD, 1995.
37. Shenk, J.S.; Workman, J.J., Jr.; Westerhaus, M.O. Application of NIR Spectroscopy to Agricultural Products. In: *Handbook of Near-Infrared Analysis*; Burns, D.A., Ed.; NIR Resources: Los Alamos, New Mexico, 2010, 419–474.
38. Alves, A.S.; Rozenberg, P.; Pâques, L.E.; Charpentier, J.P.; Schwanninger, M.; Rodrigues, J. A common near infrared-based partial least squares regression model for the prediction of wood density of *Pinus pinaster* and *Larix* x *eurolepis*. Wood Sci Technol, **2010**, in press. DOI: 10.1007/s00226-010-0383-x
39. Murray, I.; Williams, P.C., Eds. Chemical principles of near - infrared technology. In: *Near Infrared Technology in the Agricultural and Food Industries*; American Association of Cereal Chemists: St. Paul, Minnesota, 1987, 17–34.
40. Williams, P.; Norris, K., Eds. *Near-Infrared Technology in the Agricultural and Food Industries*; American Association of Cereal Chemists: St. Paul, Minnesota, 2001.
41. Shenk, J.S.; Westerhaus, M.O., Eds. *Analysis of Agricultural and Food Proudcts by Near Infrared Reflectance Spectroscopy*; NIRSystems, Inc. Silver Spring, MD, 1995b.
42. Conde, E.; Cadahia, E.; García-Vallejo, M.C.; Gónzalez-Adrados, J.R. Chemical characterization of reproduction cork from Spanish Quercus suber. Journal of Wood Chemistry and Technology **1998a**, *18*(4), 447–469.
43. Conde, E.; Cadahia, E.; Garcia-Vallejo, M.C.; de Simon, B.F. Polyphenolic composition of Quercus suber cork from different Spanish provenances. Journal of Agricultural and Food Chemistry **1998b**, *46*(8), 3166–3171.