Eur J Forest Res (2005) 124: 37–46
DOI 10.1007/s10342-004-0047-1

ORIGINAL PAPER

Piermaria Corona · Sandro Dettori
Maria Rosaria Filigheddu · Federico Maetzke
Roberto Scotti

Site quality evaluation by classification tree: an application to cork quality in Sardinia

Received: 24 May 2004 / Accepted: 22 October 2004 / Published online: 28 January 2005

Abstract Cork harvesting and stopper production represent a major forest industry in Sardinia (Italy). The target of the present investigation was to evaluate the "classification tree" as a tool to discover possible relationships between microsite characteristics and cork quality. Seven main cork oak (*Quercus suber*) producing areas have been identified in Sardinia, for a total of more than 122,000 ha. Sixty-three sample trees, distributed among different geographical locations and microsite conditions, were selected. A soil profile near each sample tree was described, soil samples were collected and analysed. After debarking, cork quality of each sample tree was graded by an independent panel of experts. Microsites where trees had more than 50% of the extracted cork graded in the best quality class, according to the official quality standard in Italy, were labelled as prime microsites, the others as nonprime microsites. Relationships between a binary dummy variable (0 for nonprime microsites, 1 for prime microsites) and site factors were investigated using classification tree analysis to select the relevant variables and to define the classification scheme. Prime quality microsites for cork production proved to be characterised by elevation, soil phosphorus content and sandiness. Results have been compared with those of the more conventional parametric approach by logistic regression. The work demonstrates the advantages of the classification tree method. The model may be appropriate for classifications at landscape and stand mapping levels, where it is possible to sample a number of microsites and to evaluate distributional characteristics of model output, while its precision is only indicative when estimating the prime quality of single microsites.

Keywords *Quercus suber* · Cork quality · Site classification and evaluation · Classification tree · Logistic regression

Introduction

The production of cork stoppers, the most remunerative industrial product from cork oak (*Quercus suber* L.) stands, requires raw material with high elasticity to assure good bottle closure and with limited porosity (Pereira et al. 1996; Vieira Neto 1996), especially if lenticels have a diameter greater than 2 mm (Ferreira et al. 2000).

The quality of raw cork is determined by the interaction of genetic and environmental factors (Natividade 1934). Frequent allogamy in this species leads to the occurrence of very diversified pheno-genotypes: its total genetic diversity is among the highest recorded in oak species (Toumi and Lumaret 1998). Thus, it is possible to discriminate populations using morphological (Garcia-Valdecantos and Catalan 1993; Schirone and Bellarosa 1996), biochemical (Toumi and Lumaret 1998, 2001) and molecular (Bellarosa 2003) descriptors. However, so far the few studies conducted have not been able to quantify a precise correlation between genotype and cork quality (Nóbrega 1997a, b).

On the other hand, empirical experience has shown that cork oak trees that produce good quality cork tend to maintain this standard through successive strippings

P. Corona (✉)
Dipartimento di Scienze dell'Ambiente Forestale e delle sue Risorse, Università della Tuscia,
via San Camillo de Lellis, 01100 Viterbo, Italy
E-mail: piermaria.corona@unitus.it
Tel.: +39-761-357425
Fax: +39-761-357389

S. Dettori · M. R. Filigheddu · R. Scotti
Dipartimento di Economia e Sistemi Arborei,
Università di Sassari,
via de Nicola 9, 07100 Sassari, Italy

F. Maetzke
Dipartimento di Scienze e Tecnologie Ambientali Forestali,
Università di Firenze, via San Bonaventura 13,
50145 Firenze, Italy

throughout their productive life. Researchers and producers are indeed convinced that genetic influence is relevant. For instance, Ferreira et al. (2000) provide statistical grounds for a genetic selection program. As an example from the empirical side, in the year 1962 an important cork producer in Gallura (Sardinia, Italy) planted 800 seedlings obtained from a single mother tree selected for its excellent cork quality: the first stripping of the planted trees confirmed that the quality of the cork, even if virgin, was above average (Asara, personal communication).

It has also been acknowledged that environment influences cork quality (Isasa 1959; Ahmed 1994; Dettori et al. 1996; Courtois and Masson 1999). Particularly, conditions favouring an intense vegetative activity may cause an increase in the diameter of lenticels and a decrease in the commercial value of the cork (Natividade 1950).

Quantitative analysis of the influence of environmental factors on forest stand productivity is a recognised study field. It reflects the needs of field foresters to assess site potential, e.g. in order to decide where to intensify silvicultural care and harvesting facilities. The target of the present investigation was to evaluate the "classification tree" method (Breyman et al. 1984) as a tool to discover possible relationships between microsite characteristics and cork quality as evaluated under current commercial context in Italy.

Conventional studies on relationships between environmental factors and forest production have mainly applied multiple regression and discriminant analysis methods after observing site properties in randomly selected locations. However, such approaches have several weaknesses; refer to Verbyla and Fisher (1989) for a more detailed discussion. Besides statistical shortcomings, a particular problem is that intensive silviculture is often only feasible on the best sites; yet a model that reveals relationships only in the range of poor to good sites may not be useful for identifying site factors that characterise the best sites. Thus, a deliberate attempt was made in the present study to determine prime site occurrence, and to circumvent the above-mentioned statistical problems.

In particular, this study aims to develop a forest site classification procedure to assess prime quality sites for commercial cork production in Sardinia, the main cork-producing region in Italy, which accounts for around 5% of the world cork production and is well known for the good quality of the cork produced.

Materials and methods

In Sardinia, cork oak stands cover more than 122,000 ha distributed in seven recognised main cork-producing areas. Based on the forest map produced by the Cork Research Institute of the regional authority of Sardinia (Stazione Sperimentale del Sughero 1991), area limits were identified defining seven strata. Twenty-four geographical locations were selected, representing typical vegetation, silvicultural systems and cork quality features within each stratum (Table 1, Fig. 1). Sixty-three sample trees were chosen, distributed among different geographical locations and microsite conditions.

Given the high variability of cork quality among trees located even in the same neighbourhood (as observed in Portugal too: e.g. see Costa 1992), it was assumed that each tree (and the associated quality of the produced

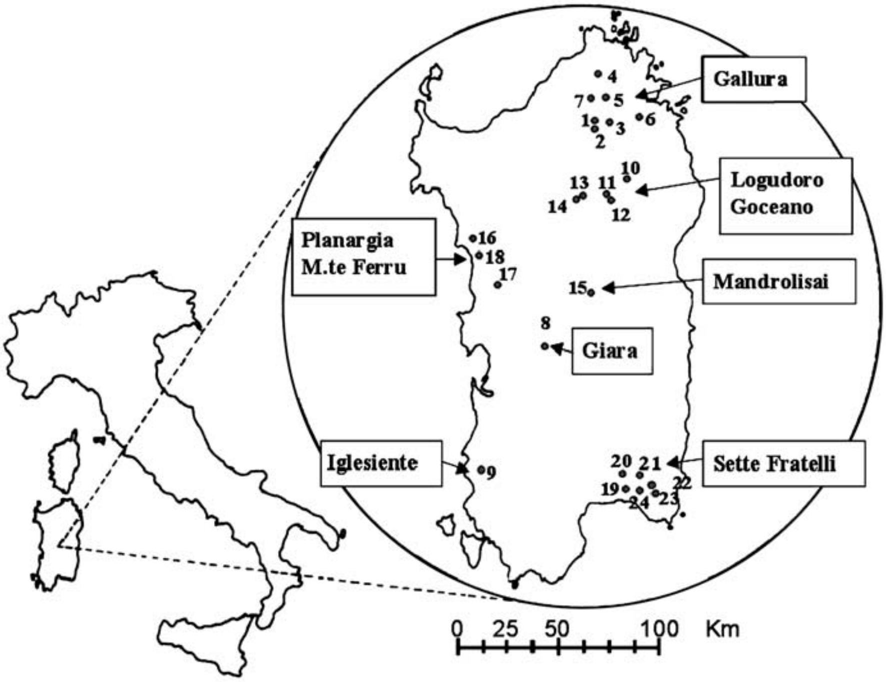


Fig. 1 The study region (Sardinia, Italy) with the main cork-producing areas and the locations where cork oak trees were sampled

Table 1 Main environmental characteristics and number of sample trees from prime and nonprime microsites for the geographical locations where the investigation was carried out

Geographical location	Phytoclimatic belt[a]	Soil class[b]	Sample trees from prime microsites (*n*)	Sample trees from nonprime microsites (*n*)
1	L2w	Dystric Xerochrepts	–	3
2	L2w	Dystric Xerochrepts	–	3
3	L2m	Lithic Xerorthents	–	3
4	L2m	Lithic Xerochrepts	2	1
5	L2w	Dystric Lithic Xerochrepts	–	3
6	L2w	Lithic Xerochrepts	1	2
7	L2m	Lithic Xerochrepts	2	1
8	L2m	Dystric Lithic Xerochrepts	–	3
9	L2m	Lithic Ruptic Xerorthentic Xerochrepts	2	1
10	C2w	Lithic Xerochrepts	3	–
11	C2w	Dystric Lithic Xerochrepts	3	–
12	C2w	Dystric Lithic Xerochrepts	3	–
13	L2c	Lithic Xerochrepts	2	1
14	L2c	Lithic Xerochrepts	2	–
15	C2w	Lithic Ruptic Xerorthentic Xerochrepts	1	2
16	L2w	Dystric Lithic Xerochrepts	1	2
17	L2c	Lithic Ruptic Xerorthentic Xerochrepts	2	1
18	L2w	Dystric Lithic Xerochrepts	–	4
19	L2w	Typic Haploxeralf	–	1
20	L2w	Lithic Xerochrepts	–	2
21	L2m	Dystric Lithic Xerochrepts	–	1
22	L2m	Lithic Xerochrepts	–	2
23	L2w	Lithic Xerorthents	–	1
24	L2w	Dystric Xerochrepts	–	2

[a] Phytoclimatic classification is after Pavari (in Arrigoni 1968). *L2c* Lauretum, second type, cold; *L2m* Lauretum, second type, mild; *L2w* Lauretum, second type, warm; *C2c* Castanetum, second type, warm

[b] Soil classification is according to USDA Soil Taxonomy (1975)

cork) is representative only of its surrounding microsite. A soil profile next to each sample tree was described (FAO 1977), and soil samples were collected and analysed following the official guidelines in Italy (MIRAAF 1994).

Cork was extracted from the sample trees according to conventional harvesting criteria: 1–1.5 m long planks 9–10 years old; debarking up to a minimum stem girth of 60 cm, and/or up to a stem height less than three times the stem girth at breast height.

Cork from each sample tree was separately graded by an independent panel of experts at the Cork Research Institute, according to the official standard method to assess cork quality in Italy (method SSS053, see Stazione Sperimentale del Sughero 2003). The system includes five classes, from the best to the poorest. The microsites where the tree had a proportion of more than 50% of the extracted cork graded in the best quality class were labelled as prime microsites, the others as nonprime microsites.

Analysis procedures

Relationships between a binary dummy variable (0 for nonprime microsites, 1 for prime microsites) and site factors were investigated. Only quantitative factors recorded for all the samples and expressed by variables on ratio/interval scales were examined. Table 2 presents a complete list of the variables, and reports assessment methods and statistics.

Many different multivariate statistical techniques can be used to predict a binary dependent variable like the prime/nonprime microsite dummy. In this study, the classification and regression tree (C&RT) method by Breyman et al. (1984) was adopted. This approach was selected based on different characteristics that are particularly advantageous in this context. It is nonparametric, hence requirements about assumptions are greatly reduced. Each factor is taken into consideration by the model, exploiting the independent information it carries without introducing any compensation effect among predictors. Its response surface is flexible and can easily accommodate complex, non-linear response structures. As experimentally highlighted by Verbyla (1987), insignificant predictor variables are less likely to be included in the C&RT model with respect to conventional stepwise regression models, especially in situations like the one investigated, where sample size is relatively small with respect to the number of predictor variables tested. Distinctively, C&RT is very robust with respect to outliers (i.e. each sample case carries the same weight in classifier development).

Tree development proceeds by identifying at each node the predictor and threshold values that best partition the remaining sample cases into the purest class membership (Verbyla 1987). Tree-growing criteria were

Table 2 Recorded variables, assessment methods, and statistics

Variable	Assessment method	Units	Min	Max	Mean	SD
Elevation a.s.l.	Taken from 1:10,000-scale maps of Sardinia	m	120	830	477	215
Total soil depth	FAO 1977	cm	8	70	35	14
Clay	MIRAAF 1994	%	2.9	25.5	11.7	4.7
Silt	MIRAAF 1994	%	2.4	41.6	14.3	10.0
Sand	MIRAAF 1994	%	43.3	91.3	74.0	12.0
pH (H_2O)	MIRAAF 1994		4.08	6.48	5.12	0.75
Carbon content	MIRAAF 1994	%	0.5	11.2	3.0	2.0
Organic matter	MIRAAF 1994	%	0.8	19.3	5.2	3.5
Total nitrogen	MIRAAF 1994	%	0.05	0.61	0.24	0.13
Phosphorus	MIRAAF 1994	ppm	2	62	30	19
Calcium	MIRAAF 1994	ppm	210	3625	989	720
Magnesium	MIRAAF 1994	ppm	55	720	207	143
Potassium	MIRAAF 1994	ppm	39	434	211	96
Sandiness	Derived variable = (sand–clay)/sand		0.61	0.97	0.83	0.09
Carbon/nitrogen	Derived variable		6	33	12.4	4.5
Calcium/magnesium	Derived variable		0.8	8.5	3.1	1.4
Magnesium/potassium	Derived variable		1.1	11	3.5	2.1
Cork prime quality	Dummy variable assessed by experts	0, 1	0	1	0.40	0.49

set as follows: parental nodes should have at least 15 cases while child nodes should have at least 1, and minimum change in impurity was set to 0.0001.

To understand classification tree advantages, a more conventional logistic regression surface has been computed. The independent variables selected as most effective predictors by C&RT were included in the logistic model, estimating the probability of the event "prime microsite" occurring. This kind of regression model requires far fewer assumptions than discriminant and multiple linear regression analyses. Standard procedures based on maximum likelihood have been adopted to estimate model coefficients.

Results

The classification tree (Fig. 2) discriminated prime cork-quality microsites as a function of elevation (ELEV, m), soil phosphorus content (PHOS, ppm) and sandiness (SAND). The ecobiological corroboration of such quantitative relationships is described in the Discussion

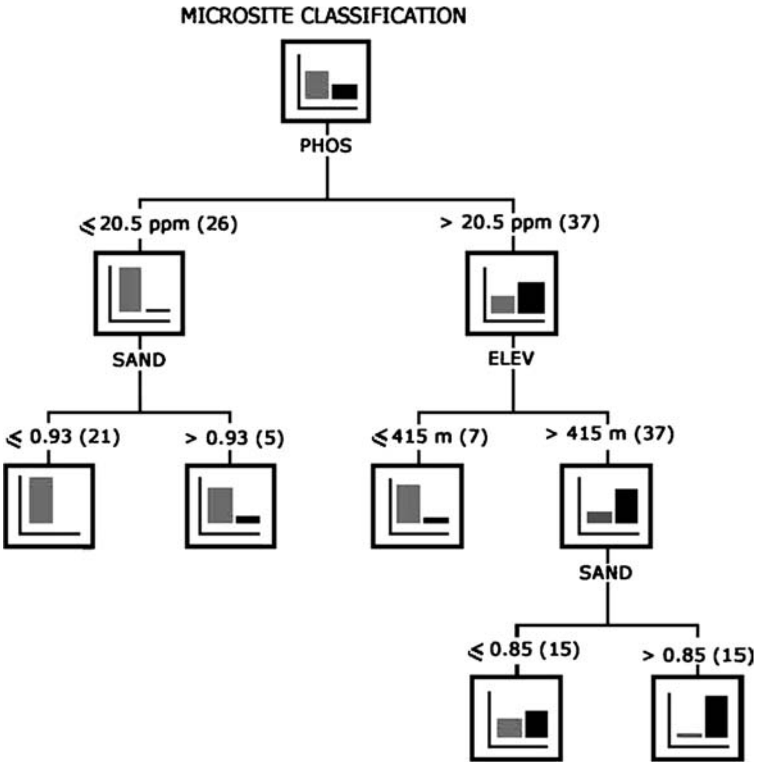


Fig. 2 Classification tree of microsite quality for cork production as a function of elevation (ELEV), soil phosphorus content (PHOS) and sandiness (SAND). *Dark grey boxes* prime microsites, *light grey boxes* nonprime microsites. In *brackets*, the number of sampled microsites corresponding to the node

section. The proportion of the samples correctly classified was quite satisfactory: misclassification risk for the overall tree was only 14.3% (SE = 4.4%); the cross-validated estimate of such a risk (Biggs et al. 1991) was equal to 26%.

The resulting classification tree presents five terminal nodes (hereinafter called "leaves"). In the first three leaves, nonprime microsites prevail; in the fourth, prime sites prevail just slightly (60%); and only in the fifth are they dominant (90%). Thresholds for the last case are: $\mathrm{PHOS} > 20.5$ and $\mathrm{ELEV} > 415$ and $\mathrm{SAND} > 0.854$. In practical terms, microsites characterised by values in the high range for all of these factors present a distinctively higher proportion of prime-quality cork.

Introducing ELEV, PHOS and SAND into a provisional logistic regression model, three outlier cases were identified with standardised residual values higher than 2 and high values of leverage and Cook's distance. After deleting outliers (with 60 observations remaining: 23 for prime microsites, 37 for nonprime microsites), the final parametric model for estimating prime-quality site for cork production was established:

$$\mathrm{Prob(prime\ site)} = \frac{1}{1 + e^{26.8 - 0.0149\mathrm{ELEV} - 0.0953\mathrm{PHOS} - 18.96\mathrm{SAND}}}$$

As shown in Table 3, all model coefficients were significant (based on Wald statistic). According to Nagelkerke R^2 statistic (Nagelkerke 1991), about 80% of the variation in the dependent variable was expressed by the model. Hosmer-Lemeshow (1989) goodness-of-fit test (chi-square value = 3.67, with 8 degrees of freedom, $P = 0.88$) did not reject the null hypothesis of no difference between observed and predicted values. Residuals did not have any significant trend with respect to ELEV, PHOS and SAND. Overall, 86.67% of cases were correctly classified by model calibration, and both the producer's and the user's accuracy were equal, 89% for nonprime microsite and 83% for prime microsite classification.

Coefficients of the logistic regression are characterised by a significant partial correlation with the dependent variable (Table 3), indicating that as any of the selected factors increases in value, so does the likelihood of occurrence of a prime-quality microsite. The variable having the greatest effect on prime-microsite probability estimation was ELEV (partial correlation coefficient equal to 0.32), followed by PHOS and SAND.

Figure 3 displays a graphical representation of the response surfaces for the classification tree and the logistic function. C&RT response is a discrete, step-like surface (non-flat sides should actually be vertical) that has a different structure for PHOS values below or above the 20.5-ppm threshold. The logistic surface, being continuous, is gradually modified by the phosphorus content: left and right graphs display surfaces corresponding to a low and a high value, respectively.

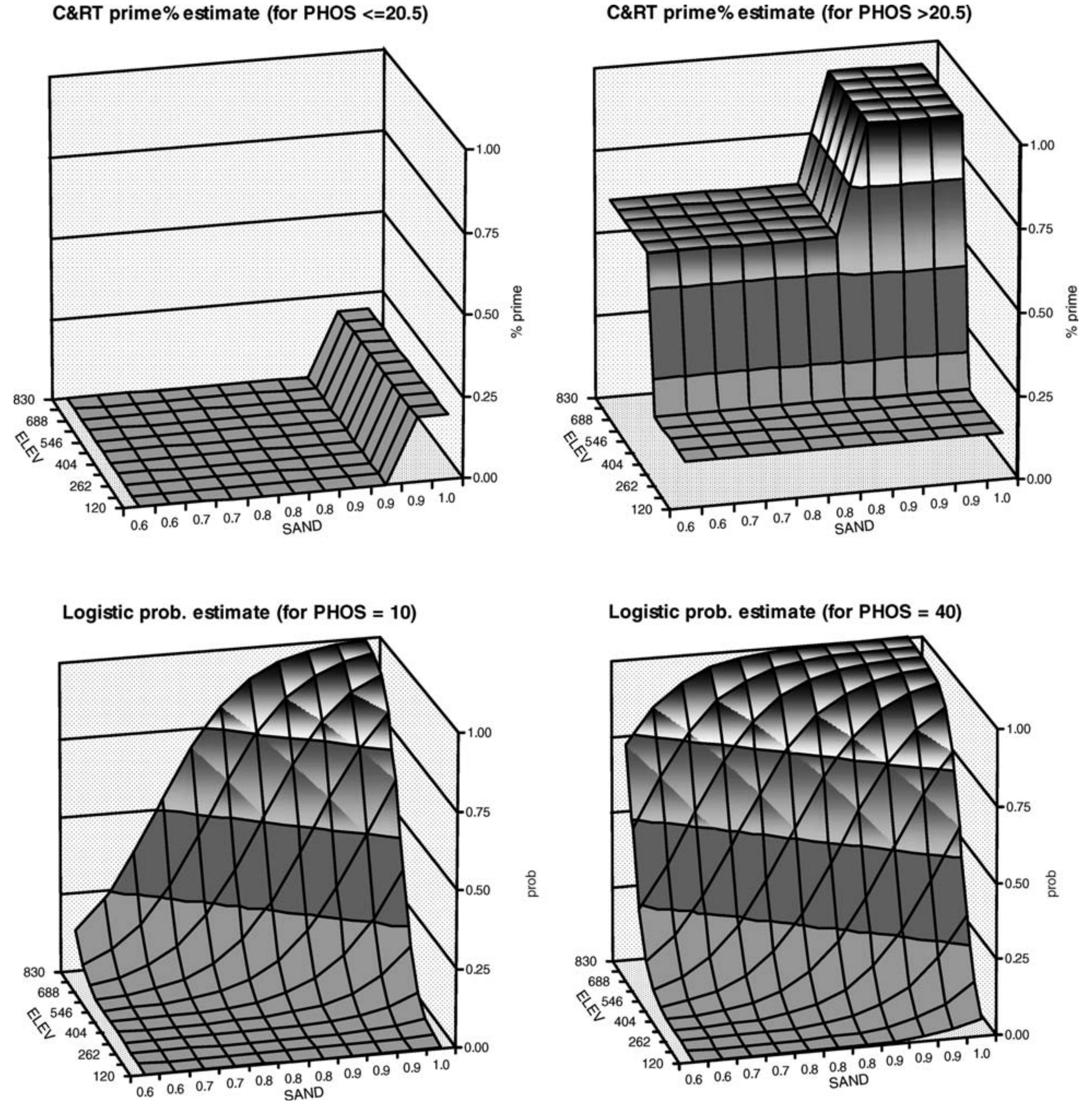
Table 3 Significance of the coefficient estimates of the logistic regression model

	Coefficient value	SE	Significance level	Partial correlation
Constant	−26.7953	10.1829	0.008	
ELEV	0.0149	0.0046	0.001	0.323
PHOS	0.0953	0.0394	0.015	0.220
SAND	18.9582	8.9482	0.034	0.176

To appreciate the meaning of the results, estimations and response surfaces were compared with experimental observations. Considering only the three selected factors, many observations displayed identical predictor values: the 63 microsites fell into 32 distinct groups (Table 4). Eighteen groups included more than one microsite, hence allowing the evaluation of "within" microsite variability; of these, 10 groups were internally homogeneous (all prime or all nonprime), only 1 was ambiguous (composed of a prime and a nonprime microsite), and 7 were mixed though they displayed a prevailing site quality. Groups were classified, according to prevailing quality and internal variability, in categories such as all prime, single prime or mixed prime. In Table 4 all group categories are catalogued, groups are listed, ordered by classification tree output, and individually evaluated. Only microsites falling in leaves five and four of the classification tree were estimated as prime sites. Each estimation was evaluated as "right" or "wrong" by comparing it to the prevailing character of the group. Evaluation of logistic regression was similar; only microsites with probability greater than 0.5 were estimated as prime. The final column compares the two outputs and notes the best one. The group containing two observations with contrasting quality was considered to have "not contributed" to the evaluation. For only one group was the classification tree estimation "wrong": it contains a single-nonprime microsite classified as prime. Logistic regression estimation was "wrong" for three groups, the same single-nonprime group as above and other two mixed-prime groups classified as nonprime.

Figure 4 presents a graphical comparison of experimental observations and response surfaces. The factors considered as predictors define a 3D experimental space. Taking advantage of the threshold of soil phosphorus content singled out by the classification tree, the experimental space has been divided in two slices. Features in each subspace are separately presented in the two graphs, projected on the sandiness–elevation plane.

Data points represent groups of experimental observations and are identified with symbols expressing group category based on prevailing quality and internal variability. Classification tree thresholds divide the plane into subplanes. Each subplane is labelled using the corresponding leaf number and bordered by a distinct broken line, marking the position of the step, in terms of prime microsite proportions, separating adjoining leaves. Logistic regression response surface for $P = 0.5$ is



C&RT prime% estimate (for PHOS <=20.5)

C&RT prime% estimate (for PHOS >20.5)

Logistic prob. estimate (for PHOS = 10)

Logistic prob. estimate (for PHOS = 40)

Fig. 3 Classification tree and logistic regression response surfaces

represented by two parallel continuous lines: one representing the cross-section for phosphorus at the threshold level (20.5 ppm) and the other for maximum and minimum phosphorus levels, respectively, in each graph. Hence, within each subspace, the response surface can be imagined connecting the two lines. Data points above the lines are classified as prime, while classification of points between the lines is not graphically evident: it depends on the phosphorus coordinate of the point.

In the graph for PHOS ≤ 20.5, the SAND threshold marks the step from leaf 1, containing only nonprime sites, to leaf 2, containing a mixed-nonprime group, i.e. the first isolated point where prime quality has been observed in otherwise generally nonprime conditions. All points were classified as nonprime for the logistic regression but the experimental data in this subspace gave no support to surface inclination with respect to the axes. For PHOS > 20.5, the ELEV threshold delimiting the leaf 3 subplane marks the step from the second and last mixed-nonprime group to the mixed-prime groups. Above this threshold, prime microsites prevail. The second, SAND threshold dividing leaf 4 from leaf 5 demonstrates the reduction in internal variability that characterises the upper right part of the graph. The groups containing only prime sites all lie above the logistic regression surface and are all correctly classified. All mixed-prime groups fall within or near the lines.

Table 4 Assessment of the classification tree and the logistic regression estimates versus experimental observations

Microsite characteristics			Microsites (*n*)		Group[a]		Classification tree (C&RT)			Logistic regression			Best result between the compared evaluation approaches
PHOS	SAND	ELEV	Prime	Nonprime	Microsites (*n*)	Category	Leaf	Quality evaluation[b]		Probability estimate	Quality evaluation[b]		
62	0.888	830	3		3	aP	5	PQ	Right	1.00	PQ	Right	–
49	0.858	775	3		3	aP	5	PQ	Right	1.00	PQ	Right	–
50	0.884	700	2		2	aP	5	PQ	Right	0.99	PQ	Right	–
26	0.887	650	3		3	aP	5	PQ	Right	0.90	PQ	Right	–
21	0.887	500	2	1	3	mP	5	PQ	Right	0.37	NPQ	Wrong	C&RT
52	0.880	700	1		1	sP	5	PQ	Right	0.99	PQ	Right	–
41	0.850	550	2	1	3	mP	4	PQ	Right	0.81	PQ	Right	–
31	0.701	790	2	1	3	mP	4	PQ	Right	0.77	PQ	Right	–
61	0.711	540	2	1	3	mP	4	PQ	Right	0.63	PQ	Right	–
37	0.806	470	2	1	3	mP	4	PQ	Right	0.27	NPQ	Wrong	C&RT
54	0.845	700	1	1	2	2c	4	PQ	NC	0.99	PQ	NC	–
27	0.820	630		1	1	sN	4	PQ	Wrong	0.67	PQ	Wrong	–
57	0.967	150		1	1	sN	3	NPQ	Right	0.31	NPQ	Right	–
29	0.914	200		1	1	sN	3	NPQ	Right	0.02	NPQ	Right	–
24	0.876	200		1	1	sN	3	NPQ	Right	0.01	NPQ	Right	–
23	0.837	200		1	1	sN	3	NPQ	Right	0.00	NPQ	Right	–
57	0.651	360	1	2	3	mN	3	NPQ	Right	0.03	NPQ	Right	–
11	0.940	420		1	1	sN	2	NPQ	Right	0.16	NPQ	Right	–
11	0.934	420		1	1	sN	2	NPQ	Right	0.14	NPQ	Right	–
16	0.933	150	1	2	3	mN	2	NPQ	Right	0.00	NPQ	Right	–
17	0.845	420		1	1	sN	1	NPQ	Right	0.05	NPQ	Right	–
6	0.912	350		1	1	sN	1	NPQ	Right	0.02	NPQ	Right	–
19	0.890	234		1	1	sN	1	NPQ	Right	0.01	NPQ	Right	–
12	0.921	234		1	1	sN	1	NPQ	Right	0.01	NPQ	Right	–
19	0.865	234		1	1	sN	1	NPQ	Right	0.01	NPQ	Right	–
19	0.930	130		1	1	sN	1	NPQ	Right	0.00	NPQ	Right	–
19	0.828	360		3	3	aN	1	NPQ	Right	0.02	NPQ	Right	–
10	0.890	320		3	3	aN	1	NPQ	Right	0.01	NPQ	Right	–
2	0.680	590		3	3	aN	1	NPQ	Right	0.01	NPQ	Right	–
8	0.608	638		2	2	aN	1	NPQ	Right	0.01	NPQ	Right	–
11	0.842	300		2	2	aN	1	NPQ	Right	0.00	NPQ	Right	–
20	0.805	120		2	2	aN	1	NPQ	Right	0.00	NPQ	Right	–

[a] Microsites are grouped by selected features and classified in the following categories: *aP* all-prime microsite, *mP* mixed-prime microsite, *sP* single-prime microsite, *2c* two contrasting microsites, *sN* single-nonprime microsite, *mN* mixed-nonprime microsite, *aN* all-nonprime microsite

[b] Quality evaluation is as follows: *PQ* prime-quality microsite; *NPQ* nonprime-quality microsite. The symbol *NC* means that the group did not contribute to the evaluation

Two of them are actually located below the surface and erroneously classified as nonprime by the logistic model, as already shown in Table 4.

Discussion

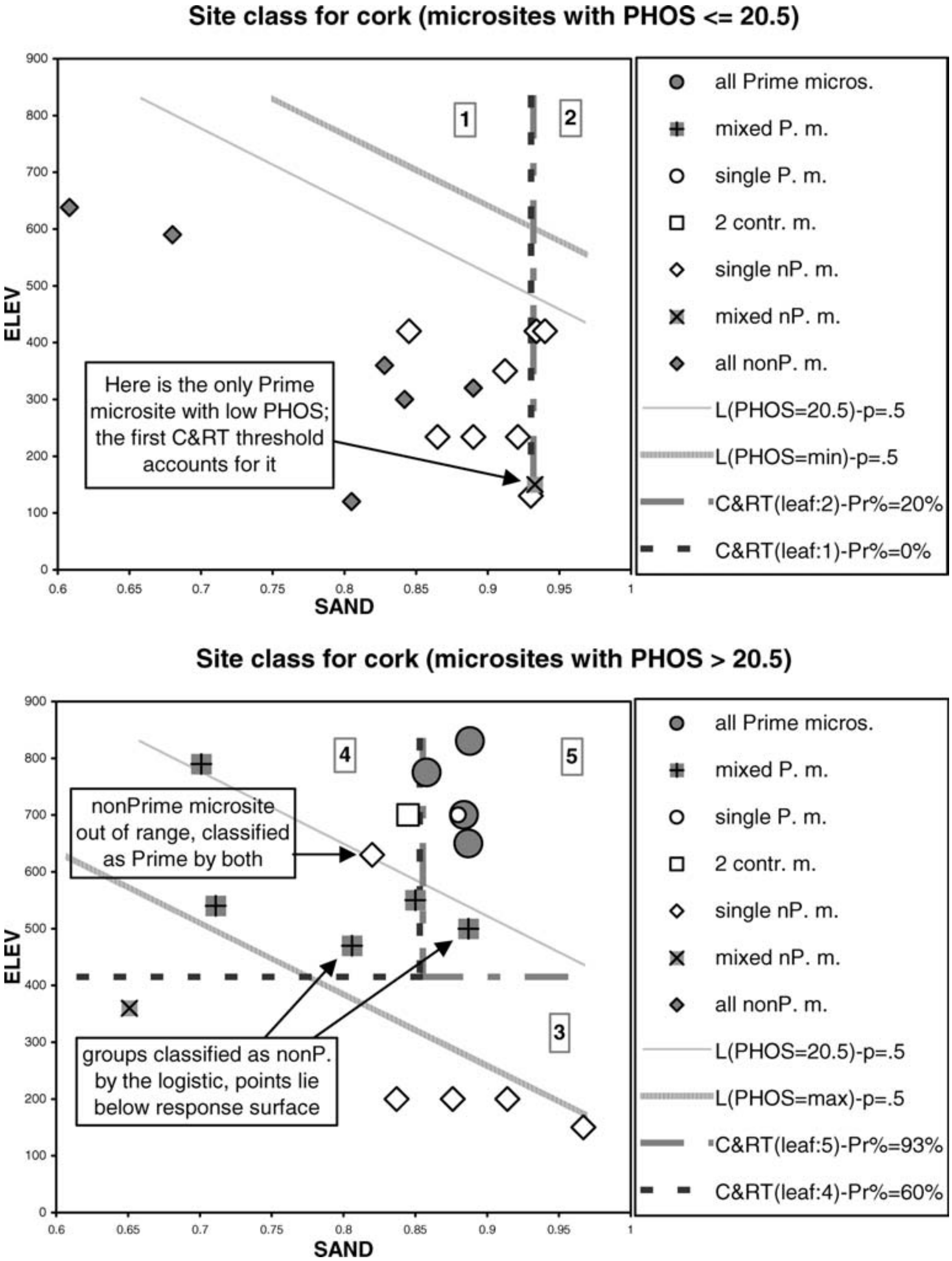
Out of the large number of site factors assessed on each microsite, using the classification tree method, three predictors have been identified as sufficient for subset classification, significantly partitioning site variability with respect to cork quality.

Given the empirical, management-oriented character of the proposed modelling approach, no conclusions should be drawn regarding the ecological influence of the selected factors. Rather, the opposite holds true: it is the ecobiological background knowledge that eventually corroborates the quantitative relationships found.

Elevation is probably related to a significant ecotype differentiation (gradient) of *Q. suber* stands in Sardinia, which has effects on cork quality. In fact, this species has a particular ability to adapt to the environment, such as the ability to adjust its reproductive behaviour. Corti (1955) and Elena Rosselló et al. (1993) observed the existence of two different reproductive strategies related to the time required from pollination to acorn-fall: a short cycle in areas with optimal environmental conditions (subhumid thermo-meso-Mediterranean), and a long cycle (biennial) in areas with harsh climatic conditions close to the northern and eastern limits of the species. In Sardinia (Italy) and Corsica (France), Abeltino et al. (2000), using RAPD's technique for DNA analysis, identified different cork oak populations, probably because of both the geographically isolating barriers of the island orography and the interposition of agricultural areas: in that study, the genetic distance was proportional to the geographical distance.

Already in 1950, Natividade noticed that shape, dimension and number of lenticels had a major influence on cork quality and suggested that the arrangement of

Fig. 4 Graphical comparison of response surfaces and observed values. See text for details



the lenticels was controlled by genetic factors. Ferreira et al. (2000) demonstrated genetic influence on porosity. Environmental conditions which favour or inhibit tree growth, instead, tend to modify lenticel size and overall cork porosity. Natividade (1950) also noticed that, for oaks producing porous cork, as growth rate increases, cork porosity increases, and its quality decreases. Similar findings are reported by Motté (1957) and Isasa (1959). The first author, observing Tunisian cork oak stands, mentioned that stand density reduced wood growth and, to a lesser extent, cork growth of individual trees. Therefore, dense stands had higher production of better quality cork. The second author observed that ecological conditions which favour intense tree growth do not allow the production of high quality cork in Catalonia (Spain). More recently, Gonzales Adrados et al. (1992), analysing cork samples from 278 stands, presented the “Land Suitability Map of Extremadura for Cork Oak and Cork Quality” based on the following ecological factors: temperature, mean annual rainfall, and soil class. Even though the great variability within each geographic area does not allow a model to be expressed, data averages show a trend: cork quality is better in arid, flat and low altitude areas than in cool, humid, steep and high altitude ones. The outcome of that work conflicts somewhat with the conclusions of all the other authors considered. Ahmed (1994), comparing six cork-producing areas in Algeria, concluded that cork quality was influenced by both environment and silvicultural treatments, and was the best when regular growth of suberous tissue occurred. Dettori et al. (1996) found a negative correlation between soil depth and cork quality expressed as a score derived from the sum of 15 physical and mechanical parameters. Courtois and Masson (1999), studying the relationships among cork quality, its mineral composition and the mineral composition of cork oak leaves, found a positive correlation between cork potassium content and cork quality.

Elevation, in the data set analysed here, ranged from 120–830 m a.s.l., offering a good representation of cork production areas in Sardinia. Its positive influence on cork quality could also be related to a reduced growth rate of the suberous tissue and to a greater uniformity of the thickness of the annual growth rings. This in turn is due to lower temperatures and, as a consequence, a shorter growing season. In addition water retention capacity of higher altitude soils, which are shallower and sandier, is frequently lower than the soils of the plains. Hence, greater water stress contributes to reduced tree growth. During summer, water deficits above 350 mm occur even in medium and high hills (Dettori et al. 1997). Summer rain represents only 5–6% of total annual precipitation in the *Lauretum* and *Castanetum* phytoclimatic belts (Arrigoni 1968). The aridity in Sardinian high hills could explain the apparent contradiction with what was reported by Gonzales Adrados et al. (1992) in Extremadura (Spain), where high hill cork oak stands grow in a humid climate.

The investigated sites presented a high variability in the phosphorus content in the soil, while potassium availability was almost always high (Table 2). Although no relationship has been reported in the literature we could access, the study stresses a positive correlation between cork quality and phosphorus, probably due to its role in cellular multiplication and lignification processes (Martin Prevel 1978; Harris et al. 1999). We could not support the positive influence of potassium reported by Courtois and Masson (1999), since the variability for this element was limited.

Conclusions

Reliable operational procedures to assess site quality for cork production in the Mediterranean Basin are lacking. The set of observations analysed confirms that this is quite a complex task to pursue: cork quality variability is high even under identical site conditions, and identification of which factors best characterise a site in terms of cork quality is controversial. However, the work performed highlights the effectiveness of the classification tree approach in factor selection and its efficiency as a classification tool, compared to the logistic regression approach.

Observed microsites were characterised considering 13 primary variables, including elevation, soil depth, texture and chemical properties, and four derived variables, including sandiness expressed as 1−clay:sand. Selected classification tree output used only three of these variables as predictors: elevation, sandiness and phosphorus content of the soil. Results' analysis demonstrated to what extent available observations support the choice. Although interpretation was not always straightforward and literature suggestions were divergent, the fact that microsite quality variability was concentrated within intermediate conditions while extreme leaves of the classification tree (numbers 1 and 5) collected all groups with no variability appears to provide rather strong information.

As a site classifier, the logistic regression function, developed considering previously selected factors, performed just slightly worse than the classification tree. Analysing and comparing the response surfaces demonstrated that the classification tree fits the observed quality data relatively better.

The capability of the model to quantitatively explain the variation in site quality estimation as observed in this study seems quite satisfactory compared to values in similar literature (e.g. Corona et al. 1998). However, 20% of the variation remains unexplained. In addition to the inevitable measurement errors, the main sources of unexplained variation might be among the following: high genetic variation, failure to measure the true causes of prime-quality site probability (missing variables, inconsistent descriptors, etc.), and ecological complexity of the study area (given the environmental heterogeneity of the considered Mediterranean region and the connected, highly variable, synergetic interactions among the soil–environment factors involved). Hence, it may be difficult to extract a single set of predictors that are constantly and strongly related to a quality measure such as the adopted dummy variable.

The current version of the model is clearly a prototype, requiring more refinement and testing (validation). At most, it allows microsite quality classification for cork production within the range of conditions found in the calibration data from Sardinia (ELEV = 120–830 m a.s.l.; PHOS = 2–62 ppm; SAND = 0.61–0.97). Model precision may be only indicative when assessing the potential for prime quality of single microsites, while it may be appropriate for classifications at landscape and stand mapping levels when sampling a number of microsites and evaluating distributional characteristics.

We feel, however, that the results presented are quite remarkable because, although cork quality is closely related to genetic aspects, it is also relevant to have experimentally demonstrated that certain environmental conditions may allow the genotype to express its full potential. This is essential for deciding where to intensify silvicultural care and harvesting facilities or where to prioritize cork oak reforestation.

Acknowledgements The work was carried out with equal contributions by the authors and was financed by a grant from the European Union under the project "Sardegna/Corsica Interreg". The authors wish to thank Gherardo Chirici and the personnel from the Stazione Sperimentale del Sughero (Tempio Pausania, Sassari, Italy) for technical assistance, and Dave Verbyla for the helpful comments provided on an earlier draft.

References

Abeltino PM, Bullitta S, Dettori S, Filigheddu MR, Manchinu M, Meloni MC, Piluzza G (2000) Caratterizzazione e valorizzazione della quercia da sughero (*Quercus suber* L.) dell'area sardo–corsa. In: IV Congresso Nazionale sulla Biodiversità: Germoplasma locale e sua valorizzazione. Agabbio Editore, Alghero, pp 763–766

Ahmed YS (1994) Influence sylvicole dans la production de liège de qualité. In: Simposio Mediterraneo sobre Regeneracion del Monte Alcornocal, Merida, pp 394–401

Arrigoni PV (1968) Fitoclimatologia della Sardegna. Webbia 23:1–100

Bellarosa R (2003) I. Introduction: brief synthesis of the current knowledge on cork oak. In: Varela MC (eds) Evaluation of genetic resources of cork oak for appropriate use in breeding and gene conservation strategies. Handbook of the concerted action on cork oak, reference FAIR 1 CT 95- 0202, 2000. INIA-Estação Florestal Nacional, Portugal

Biggs D, de Ville B, Suen E (1991) A method of choosing multiway partitions for classification and decision trees. J Appl Stat 18:49–62

Breyman L, Freidman JH, Olshen RA, Stone CJ (1984) Classification and regression trees. Wadsworth, Belmont

Corona P, Scotti R, Tarchiani N (1998) Relationship between environmental factors and site index in Douglas-fir plantations in central Italy. For Ecol Manage 110:195–207

Corti R (1955) Ricerche sul ciclo produttivo di specie del genere Quercus della flora italiana. II: contributo alla biologia ed alla sistematica di *Quercus suber* L. e in particolare delle forme a sviluppo biennale della ghianda. Ann Accademia Italiana di Scienze Forestali 4:55–133

Costa A (1992) Modelos de previsão de produsão de cortiça. Secundo Encontro sobre os montados de sobro e Azinho. Direcção Geral das Florestas, Circunscrição Florestal de Évora, Évora, pp 150–159

Courtois M, Masson P (1999) Relations entre la qualité du liège, sa composition minérale et la composition minérale des feuilles du chêne-liège (*Quercus suber* L). Ann Sci For 56:521–527

Dettori S, Falqui A, Filigheddu MR, Manchinu M, Mannoni S (1996) Influenze genetiche e ambientali sulla qualità del sughero sardo. In: 2nd international symposium on cork, Pavia, pp 123–126

Dettori S, Falqui A, Mavuli S, Orrù A, Poddighe D, Todde M (1997) Prime esperienze di coltivazione di ciliegio e noce da legno in Sardegna. Annali Istituto Sperimentale per la Selvicoltura 25/(26):227–241

Elena Rosselló JA, de Rio JM, Garcia Valdecantos JL, Santamaria IG (1993) Ecological aspects of the floral phenology of the cork-oak (*Q suber* L): why do annual and biennial biotypes appear? Ann Sci For 50(Suppl)1:114–121

FAO (1977) Guidelines for soil profile description, 2nd edn. Soil Resources Development and Conservation Service, Land and Water Development Division, Rome

Ferreira A, Lopes F, Pereira H (2000) Caractérisation de la croissance et de la qualité du liège dans une région de production. Ann For Sci 57:187–193

Garcia-Valdecantos JL, Catalan G (1993) Possible criteria for selection of *Quercus suber* plus trees. Ann Sci For 50(Suppl)1:448–451

Gonzáles Adrados JR, Elena Rosselló R, Tella Ferreiro G (1992) Potencialidad del territorio para el Alcornoque en Extremadura. Scientia Gerundendis 18:185–194

Harris RW, Clark JR, Mattheny NP (1999) Arboriculture. Integrated management of landscape trees, shrubs and vines. Prentice-Hall, Upper Saddle River, NJ

Hosmer DW, Lemeshow S (1989) Applied logistic regression. Wiley, New York

Isasa NA (1959) Seccion de ecológia, tipos de montes y suelos forestales. Consideraciones ecológicas sobre la especie *Quercus suber* (alcornoque) en España. An Inst For Invest Exp 11:8

Martin Prevel P (1978) Rôle des elements minéraux chez les végétaux. Fruits 33(7–8):521–529

MIRAAF (1994) Metodi ufficiali di analisi chimica del suolo. ISMEA Edizioni, Roma

Motté M (1957) Bilan de l'etude experimentale de la croissance du chene-liege en Tunisie. Revue Forestiere Francaise 9(7):571–582

Nagelkerke NJD (1991) A note on general definition of the coefficient of determination. Biometrika 78:691–692

Natividade JV (1934) Cortiças. Contribução para o estudo do melhoramento da qualidade. Publicações da Direcção Geral dos Serviços Florestais e Aquícolas I (1), Lisboa

Natividade JV (1950) Subericultura. Ministério da Agricultura, Pescas e Alimentação, Direcção Geral das Florestas, Lisboa

Nóbrega F (1997a) Polimorfismo isoenzimático na caracterização do sobreiro. Silva Lusitana 5(1):29–58

Nóbrega F (1997b) Contribuição para o estudo da relação do polimorfismo isoenzimático com a qualidade da cortiça. Silva Lusitana 5(2):179–191

Pereira H, Lopes F, Graça J (1996) The evaluation of the quality of cork planks by image analysis. Holzforschung 50(2):111–115

Schirone B, Bellarosa R (1996) Una proposta italiana per i descrittori dela quercia da sughero (*Quercus suber* L). Silvanova/CISDAM, De Rerum Natura 1(Suppl):5–13

Stazione Sperimentale del Sughero (1991) Carta Forestale della Sardegna. Edizioni Stazione Sperimentale del Sughero, Tempio Pausania

Stazione Sperimentale del Sughero (2003) Metodo di prova SSS053. Sughero grezzo, Classificazione. SGQ 04/2003-00, Stazione Sperimentale del Sughero, Tempio Pausania

Toumi L, Lumaret R (1998) Allozyme variation in cork oak (*Quercus suber* L.): the role of phylogeography and genetic introgression by other Mediterranean oak species and human activities. Theor Appl Genet 98:647–656

Toumi L, Lumaret R (2001) Allozyme characterisation of four Mediterranean evergreen oak species. Biochem Syst Ecol 29:799–817

USDA (1975) Soil Taxonomy. Soil Survey Staff, Soil Conservation Service. Agriculture handbook vol. 436, Washington DC

Verbyla DL (1987) Classification trees: a new discriminant tool. Can J For Res 17:1150–1152

Verbyla DL, Fisher RF (1989) An alternative approach to conventional soil-site regression modelling. Can J For Res 19:179–184

Vieira Neto P (1996) Anticipating cork stoppers mechanical behaviour according to their visual quality and density. In: 2nd international symposium on cork, Pavia, pp 45–49