

# Field sampling of cork value before extraction in Portuguese ‘montados’

Alice M. Almeida · Margarida Tomé

Received: 13 August 2009 / Accepted: 23 November 2009 / Published online: 10 December 2009  
© Springer Science+Business Media B.V. 2009

**Abstract** The assessment of cork quality and the estimation of cork value are very important to forest landowners, for management purposes and for cork commercialisation. The Forest Producers Associations have been using a sampling scheme with the objective of estimating cork value (price per unit of weight, usually kg) before extraction, based on the sampling of individual trees along a zigzag transect that covers the entire stand. The sampling error is usually too high, but, from a practical standpoint, it is difficult to increase the sampling intensity if it would imply an increase in sampling costs. The aim of this work was to propose, from data collected in six stands representative of the cork oak stands in Portugal, an alternative sampling methodology with a more efficient precision/cost ratio. Precision and costs of alternative sampling designs based on clusters of different sizes, complemented with analysis of the intracluster correlation coefficient, were studied in order to propose the most adequate sampling strategy. Single-stage cluster sampling with clusters of 5–7 trees guarantees a reasonable sampling error (10–15%) and can be conducted without a large increase in cost.

**Keywords** Cork oak · Cork value · Cork quality · Cork thickness · Single-stage cluster sampling · Intracluster correlation coefficient

## Introduction

Cork oak (*Quercus suber* L.) is one of the most important forest species in Portugal, mostly integrating multifunctional agroforestry systems called ‘montados’. The production of cork is usually combined with cattle grazing, agricultural crops and hunting (Pereira 2007). According to the results of the last National Forest Inventory (data from 2005 to 2006, Tomé et al. 2007), the Portuguese cork oak area (pure stands and mixed stands with cork oak as the dominant species) was estimated as 693,000 ha (21% of the national forest area), located mainly along and south of the Tagus river.

In these stands, the main product is the cork, a thick and continuous layer of suberised cells, produced by the meristematic cork cambium, which makes up the external envelope of the stem and branches. Usually, the first cork layer is removed when the tree attains an overbark perimeter at breast height of 70 cm (virgin cork) and then, periodically, at 9-year intervals (reproduction or mature cork). Cork stripping is done manually using an axe. The stripper makes long cuts on the cork, first horizontal and then vertical, allowing the detachment of large cork pieces—the planks. This

---

A. M. Almeida (✉) · M. Tomé  
Centro de Estudos Florestais, Instituto Superior de  
Agronomia, Universidade Técnica de Lisboa, Tapada da  
Ajuda, 1349-017 Lisbon, Portugal  
e-mail: alicecalmeida@isa.utl.pt

operation is done not only on the main stem, but also on every branch with a perimeter  $>70$  cm, a size that is very common on cork oak trees. The cork cambium is destroyed during the cork stripping process, but the sustainability of production is ensured because the species has the capacity to regenerate a new cambium in the inner part of the bark, immediately below the cells that die as a consequence of cork stripping (Natividade 1950; Pereira and Tomé 2004; Pereira 2007).

Cork production in Portugal is about  $157 \text{ Mg year}^{-1}$  and the country is the first world producer, with 52.5% of the world production (APCOR 2007 cited in APCOR 2008a), and 32% of the world area (DGRF 2006 cited in APCOR 2008a). Cork and cork products exports (mainly cork stoppers) account for 30% of the total of forest products exports, which make up 2.3% of Portuguese exports (APCOR 2008b).

Cork value, and therefore cork price, depends on cork quality, defined according to its ability to produce good quality wine stoppers, undoubtedly the most valuable product made out from cork. Cork quality for cork stoppers depends on three characteristics: (1) total cork thickness after boiling, with a thickness between 32 and 40 mm being the most appropriate for natural cork stoppers; (2) cork porosity, with the best quality associated with low total porosity and small pores; and (3) presence of defects such as cracks or insect galleries that make cork useless for high quality cork stoppers (Natividade 1950; Adrados and Pereira 1996; Pereira and Tomé 2004; Pereira 2007). Estimation of cork value (price per unit of weight, usually kg) before cork extraction and processing is extremely important to forest landowners for cork commercialisation. Since 1992 (DGF 1993), the Portuguese Forest Service suggested a sampling scheme for cork price estimation prior to extraction, based on the sampling of a certain number of individual trees along a zigzag transect, and this scheme began to be implemented by most of the Forest Producers Associations. Some tests, based on the comparison of the price estimated on the basis of this sampling scheme and the real price computed after processing the whole cork from a stand, indicated some bias in the estimation of cork price causing differences that, in some cases, were not acceptable from a practical standpoint (APFC 2000, personal communication). The objective of the study reported here was to propose an alternative scheme that could produce better estimates of cork value

without a significant increase in the cost of the present sampling scheme.

## Methods

### Analysis of the present sampling scheme

The first step towards the improvement of the sampling methodology for the field assessment of cork value was the analysis of the methodology currently used by the Forest Producers Associations. The sampling procedure described here applies to a homogeneous area as, when necessary, the stand is stratified into homogeneous units that are sampled independently. The sampling begins by the planning of a polygonal transect (zigzag) within each homogeneous unit. The transect length, subjectively defined as a function of the size of the area under sampling, is divided by the planned number of sample trees, and the sampling points are systematically defined at equal distances along the transect. The tree closest to each sampling point is selected as a sample tree. A cork piece of  $\sim 20 \text{ cm} \times 20 \text{ cm}$  (usually called “cala”) is taken from each sample tree at 1.30 m. There is some inside-tree variability of cork quality, but it is assumed that the sample taken at 1.30 m is representative of the mean cork value for the tree. This assumption is supported, to some extent, by empirical evidence (Pereira 2001). In the laboratory, the sample is boiled and air dried. Cork thickness is measured before and after boiling. After boiling and drying, each cork sample is classified for cork quality into 1 of 25 classes that combine quality and cork thickness at harvest (Table 1). The cork value of a stand is estimated by averaging the cork values obtained for each sampled tree. No sampling error is computed. This method has some weaknesses from a statistical standpoint, one of them being the way the sample trees are selected that causes an over-sampling of the areas close to the vertices of the polygonal transect and an under-sampling of the areas between them. One of the problems for the definition of a new sampling design is that there is high restriction on the cost that the forest landowner is willing to pay. The current sampling design can be implemented by a team of two persons, one of them specialised in the extraction of the cork samples, during 1 day. The new sampling scheme should not exceed the time needed by the current sampling design.

**Table 1** Cork value for the industrial cork quality classes considered

Caliper class (mm)	Quality class					Industrial use if quality is high
	High (1st/3rd)	Medium (4th)	Medium (5th)	Low (6th)	Refuse	
14–18	46	8	8	8	8	
18–23	46	46	23	8	8	Discs (used in champagne corks)
23–27	46	46	23	8	8	Discs and manufactured corks
27–41	100	65	23	23	8	Natural corks
>41	100	65	23	23	8	Too thick → a lot of refuse

The values are indexed to the class with a higher value (APFC 2007, personal communication)

Data

The study was based on data collected in six stands located in the most important regions of the cork oak distribution area. The stands were selected by the local Forest Producers Associations as representative of the cork oak stands in each region. Most of the stands were pure stands or had cork oak as dominant species.

Each stand was systematically sampled using a systematic grid, applied from a randomly selected coordinate, with the distance between the centre of circular plots estimated as  $\sqrt{10000A/30}$ , where  $A$  was the area of the stand and 30 the planned number of circular plots. The number of plots measured per site is given in Table 2. In Alcácer, Portel and Grândola, some of the planned plots were in fact outside the stand (located in stands dominated by holm oak), while in Aljezur the stand was partially affected by a wildfire between the establishment of the plots and the sampling of the cork.

A large circular plot (radius = 30 m) was established around each plot centre, and the following variables were obtained for each tree inside the plot: distance and azimuth to the plot centre; diameter at breast height (underbark for trees with mature cork and overbark for trees with virgin cork). When a tree with

mature cork was not debarked in the measurement year, cork thickness was measured at two opposite sides at 1.30 m (along the plot radius). Diameter underbark for trees with virgin cork was estimated with the equation developed by Tomé (2004). Table 3 summarises the characteristics of the sampled plots.

A 20 cm × 20 cm cork sample was taken at breast height from trees within 30 m of the plot centre up to maximum of 20 trees, the closest to the plot centre (Table 2). All the samples were boiled and dried, and the cork thickness was measured before and after boiling. After boiling and drying, each cork sample was classified for cork quality into 1 of 25 classes that combine quality and cork thickness at harvest (Table 1).

Statistical analysis

The objective of the sampling for cork value (price per unit of weight, usually kg) estimation is to estimate the population parameter:

$$Q = \sum_{k=1}^{n_k} Q_k p_k = \frac{\sum_{i=1}^{N_t} Q_i}{N_t},$$

where  $Q_k$  is the cork value for cork quality of class  $k$ ,  $p_k$  is the true proportion of cork sampling units in

**Table 2** Characterisation of the sampling undertaken in each stand

Stand	Area (ha)	Number of plots	Number of trees per cluster			$T$ (°C)	$P$ (mm)
			Min	Mean	Max		
Coruche	158	22	5	12	17	16	657
Mora	85	31	5	17	22	16	650
Alcácer	72	30	5	13	20	16	587
Aljezur	87	14	1	13	22	16	556
Portel	48	26	1	10	18	17	644
Grândola	55	20	6	15	20	17	668

**Table 3** Characterisation of the six stands that were used in the study

Stand	N (ha <sup>-1</sup> )		Gu (m <sup>2</sup> ha <sup>-1</sup> )		du (cm)		Cork value	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Coruche	60	25	4.28	1.89	26.15	14.79	41.23	26.06
Mora	81	23	5.96	1.52	29.34	8.58	38.57	25.50
Alcácer	99	32	11.33	1.90	35.38	14.44	36.65	26.35
Aljezur	96	48	4.81	3.37	23.11	10.26	29.15	24.84
Portel	45	24	4.77	2.70	32.20	18.12	33.84	27.96
Grândola	99	52	5.46	2.67	23.71	11.79	29.81	22.85

*Note:* Other species are not included  
*N* number of trees per ha,  
*Gu* basal area under bark,  
*du* diameter at breast height under bark, *SD* standard deviation

cork quality class *k*, *n<sub>k</sub>* is the number of cork quality classes, *Q<sub>i</sub>* is the cork value for tree *i* and *N<sub>t</sub>* is the total number of trees in the stand.

Table 1 shows the cork values that were used for each cork quality class. The values are indexed according to the value of the most valuable cork quality class and were provided by one of the Forest Producers Association involved in the study (APFC 2007, personal communication).

The data analysis was carried out in order to define the most appropriate sampling design to estimate the parameter *Q* and to minimise costs (time needed for field sampling implementation) for a pre-fixed percent error. The high variability of cork value among trees within the same stand (Ferreira et al. 2000; Corona et al. 2005) suggested the use of cluster sampling (theory presented in, for example, Cochran 1977; Loetsch and Haller 1973; Vries 1986; Schreuder et al. 1993; Shiver and Borders 1996).

Consider that a group of neighbouring trees is a cluster or primary sampling unit and that each tree is a secondary sampling unit. The following notations will be used:

- N*—number of clusters,
- n*—number of clusters sampled,
- M*—number of trees per cluster,
- Q<sub>t<sub>i</sub></sub>*—total of values for cluster *i*,
- Q<sub>ij</sub>*—value for cork of tree *j* in cluster *i*,

$$\overline{Q_t} = \frac{\sum_{i=1}^N Q_{t_i}}{N}$$

population mean for the primary units, and

$$\overline{\overline{Q}} = \overline{Q} = \frac{\sum_{i=1}^N \sum_{j=1}^M Q_{ij}}{NM} = \frac{\sum_{i=1}^N Q_{t_i}}{NM}$$

population mean for the secondary units.

The objective of the sampling was to obtain an estimate of the mean value per tree  $\frac{\hat{Q}}{Q}$ , the respective variance  $s_{\frac{\hat{Q}}{Q}}^2$  and the percent sampling error  $E_{\frac{\hat{Q}}{Q}}\%$ , denoted by %*E*.

The unbiased estimates of the means for the primary and secondary units are given by:

$$\frac{\hat{Q}_t}{Q_t} = \frac{\sum_{i=1}^n \sum_{j=1}^M Q_{ij}}{n} = \frac{\sum_{i=1}^n Q_{t_i}}{n}$$

estimator of population mean for the primary unit, and

$$\frac{\hat{Q}}{Q} = \frac{\sum_{i=1}^n \sum_{j=1}^M Q_{ij}}{nM} = \frac{\sum_{i=1}^n Q_{t_i}}{nM}$$

estimator of population mean for the secondary unit.

The estimators of the respective variances can be easily found by:

$$s_{\frac{\hat{Q}_t}{Q_t}}^2 = \frac{N - n}{N} \frac{1}{n} \frac{\sum_{i=1}^n \left( Q_{t_i} - \frac{\hat{Q}_t}{Q_t} \right)^2}{(n - 1)}$$

estimator for variance of  $\frac{\hat{Q}_t}{Q_t}$ , and

$$s_{\frac{\hat{Q}}{Q}}^2 = \text{var} \left( \frac{1}{M} \frac{\hat{Q}}{Q_t} \right) = \frac{1}{M^2} \frac{N - n}{N} \frac{1}{n} \frac{\sum_{i=1}^n \left( Q_{t_i} - \frac{\hat{Q}}{Q} \right)^2}{(n - 1)}$$

estimator for variance of  $\frac{\hat{Q}}{Q}$ .

The number of degrees of freedom is (*n* - 1).

If population size (*N*) is very large in relation to the number of clusters sampled (*n*), as is the case in this study (*N* - *n*)/*N* → 1 and the expressions for the

variance of the means simplify. For the variance of the mean for the secondary units:

$$s_{\frac{\wedge}{Q}}^2 = \frac{1}{M^2 n} \frac{\sum_{i=1}^n \left( Q_{t_i} - \frac{\wedge}{Q} \right)^2}{(n - 1)}.$$

Percent sampling error can be computed as:

$$\%E = \frac{E_{\frac{\wedge}{Q}}(\%) }{\frac{\wedge}{Q}} = \frac{s_{\frac{\wedge}{Q}} t_{\alpha/2}}{\frac{\wedge}{Q}} 100,$$

where  $s_{\frac{\wedge}{Q}}$  is the standard error of the mean and  $t_{\alpha/2}$  is the  $t$ -Student with  $(n - 1)$  degrees of freedom ( $\alpha = 0.05$ ).

The definition of the more appropriate sampling design implies the determination of the number and size of clusters ( $n$  and  $M$ , respectively) that lead to the minimum cost for a certain pre-fixed percent error.

The existence of tree coordinates allowed the simulation of several sampling designs (combination of  $n$  clusters with  $M$  trees per cluster) with different cluster sizes and the comparison of the associated percent errors and sampling costs. The results of these analyses were complemented by the intracluster correlation coefficient of each one of the stands studied.

#### Simulation of sampling designs with different cluster sizes

Tree coordinates were registered in each plot. Therefore, using the data from the  $M$  trees per cluster, it was possible to simulate samplings with clusters of 1, to  $M$  trees and, from them, to estimate the sample size needed to obtain a pre-fixed percent error.

For a sampling with clusters of  $M$  trees, the number of clusters needed to obtain an  $E\%$  error can be estimated from:

$$n = \frac{s_{\frac{\wedge}{Q}}^2 z_{\alpha/2}^2}{\%E/100Q} \quad \text{with} \quad s_{\frac{\wedge}{Q}}^2 = \frac{\sum_{i=1}^n \left( Q_{t_i} - \overline{Q} \right)^2}{M^2(n - 1)}.$$

The number of clusters needed to obtain pre-fixed errors of 10, 15 and 20% were computed for sampling with different cluster sizes, as well as the respective costs.

#### Analysis of intracluster correlation coefficient

The concept of the intracluster correlation coefficient, which is related with the structure of the population,

plays a prominent role in this research. The variance of the estimator for the population mean per secondary unit can be expressed as a function of the intracluster correlation  $\rho$  and the variance of the secondary units  $\sigma_Q^2$  (Vries 1986):

$$s_{\frac{\wedge}{Q}}^2 \doteq \frac{1}{nM} \sigma_Q^2 (1 + (M - 1)\rho) \quad (\text{for } n \ll N),$$

where

$$\sigma_Q^2 = \frac{\sum_{i=1}^N \sum_{j=1}^M \left( Q_{ij} - \overline{Q} \right)^2}{NM}$$

is the population variance of the secondary units.

We may then conclude that the precision of a single-stage cluster sampling depends on the following factors:

1. The number of clusters ( $n$ ) in the sample, if sample size is increased  $k$  times, then  $\%E$  becomes  $\sqrt{k}$  times smaller.
2. The cluster size ( $M$ ), in a random forest  $\rho = 0$  and  $\%E$  is inversely proportional to  $\sqrt{M}$ . If  $\rho > 0$  (patchy forest) and constant with cluster size, we have for clusters  $k$  times larger that  $\frac{1}{kM}\{1 + (kM - 1)\rho\} < \frac{1}{M}\{1 + (M - 1)\rho\}$ . Therefore, for clusters  $k$  times larger, the  $\%E$  also decreases for a fixed  $n$ . However, for any  $\rho > 0$ , the rate of decrease is smaller than for  $\rho = 0$ . In practice,  $\rho$  tends to decrease with increasing cluster size, therefore,  $\%E$  decreases slightly more than for a constant  $\rho$ , but still less strongly than for  $\rho = 0$ .
3. The variance of the secondary units ( $\sigma_Q^2$ ), this parameter is small in homogeneous stands and large in heterogeneous ones. In general, it will increase with the size of the forest area, as there is more variation over large areas.
4. The intracluster correlation coefficient ( $\rho$ ), if  $\rho = 0$  (random forest), then  $s_{\frac{\wedge}{Q}}^2 = \frac{\sigma_Q^2}{nM}$  which

equals the variance of the mean of a simple random sampling of size  $nM$ , drawn from a population of  $nM$  secondary units. In other words, if  $\rho = 0$ , there is no statistical difference between sampling with a few large clusters and with many small clusters. If  $\rho > 0$ , we can easily see that the variance of the estimator of the mean per secondary unit in single-stage cluster sampling is always larger than the corresponding

**Table 4** Analysis of variance for a single-stage cluster sampling

Source of variation	Sum of squares	df	Mean squares	F-test
Between PU's	$M \sum_{i=1}^n (Q_i - \bar{Q})^2$	$n - 1$	$s_b^2$	$M\sigma_b^2 + \sigma_w^2$
Within PU's	$\sum_{i=1}^n \sum_{j=1}^M (Q_{ij} - \bar{Q}_i)^2$	$n(M - 1)$	$s_p^2$	$\sigma_w^2$
Total	$\sum_{i=1}^n \sum_{j=1}^M (Q_{ij} - \bar{Q})^2$	$nM - 1$		

PU primary unit, df degrees of freedom

variance in simple random sampling with the same sample bulk.

It is easy to conclude that the estimation of the intracluster correlation coefficient is an important step for the definition of the optimum sampling design. It can be estimated on the basis of the analysis of variance due to the clusters (Table 4). An approximately unbiased estimator of  $\rho$  is (Vries 1986):

$$\hat{\rho} = \frac{s_b^2 - s_y^2}{(M - 1)s_y^2} \approx \frac{s_b^2 - s_p^2}{s_b^2 + (M - 1)s_p^2},$$

where  $s_y^2$  is an estimator of the variance of the population (secondary units) and  $s_b^2$  and  $s_p^2$  are the between and within primary units mean squares, respectively.

The problem of determining the optimal sampling design (optimum cluster size and number of clusters) can also be analysed using the intracluster correlation coefficient. As a starting point, it is possible to identify the number of clusters ( $n$ ) that gives as good results as sampling with  $n_t$  trees, if the population consists of  $N_t$  trees. The two methods must be equally precise, so the variances of the estimated population totals with each one of the methods must be equal:

$$N_t^2 \frac{\sigma_Q^2}{n_t} = (NM)^2 \frac{\sigma_Q^2}{nM} \{1 + (M - 1)\rho\}.$$

From this requirement, it follows that the number of clusters of  $M$  trees must satisfy:

$$n = \frac{n_t}{M} \{1 + (M - 1)\rho\}.$$

Note that

$$\frac{1 + (M - 1)}{M} = 1 \rightarrow \frac{1 + (M - 1)\rho}{M} \leq 1.$$

Therefore, as  $0 \leq \rho < 1 \rightarrow n \leq n_t$ .

For a certain  $\rho$ , the greater is the value of  $M$ , the greater will be the difference between  $n$  and  $n_t$ . On the

other hand, for a certain  $M$ , the closer the value of  $\rho$  to 1, the smaller will be the difference between  $n$  and  $n_t$ . This means that in populations with small intra-cluster correlations coefficients the single-stage cluster sampling is more efficient than random sampling.

### Estimation of total sampling time

For the analysis of alternative sampling designs, there is the need to analyse the costs associated with each sampling design. The cost has two important components: the cost of travelling between plot centres and the cost of extracting a cork sampling unit from one tree (including the time to travel between trees belonging to the same cluster). Here, the costs will be expressed by the time needed to perform a field sampling.

The cost of travelling between plot centres can be estimated with a method similar to the one that was proposed by Zeide (1980). In a stand of area  $A$ , assuming that the area is approximately square and that the plots are placed according to a square grid, the mean distance between plots will be

$$\text{dist} = \sqrt{\frac{10000A}{n}} (m).$$

If the walking speed is estimated as  $K \text{ km h}^{-1}$ , the time needed for travelling ( $t_1$ ) in a sampling of  $n$  clusters is equal to:

$$t_1 = n \frac{\text{dist}}{1,000K} = \frac{1}{10K} \sqrt{An} \text{ (hours)}.$$

On the other hand, if the time to extract a cork sampling unit is estimated in  $X$  min, the total time to sample  $n$  plots ( $t_2$ ) can be estimated by:

$$t_2 = nM \frac{X}{60} \text{ (hours)}.$$

The total time required for the sampling ( $t$ ) is then:

$$t = t_1 + t_2 = \frac{1}{10K} \sqrt{An} + nM \frac{X}{60} \text{ (hours)}.$$

Total sampling costs were estimated for each stand and, for comparison purposes, standardised for an area of 100 ha.

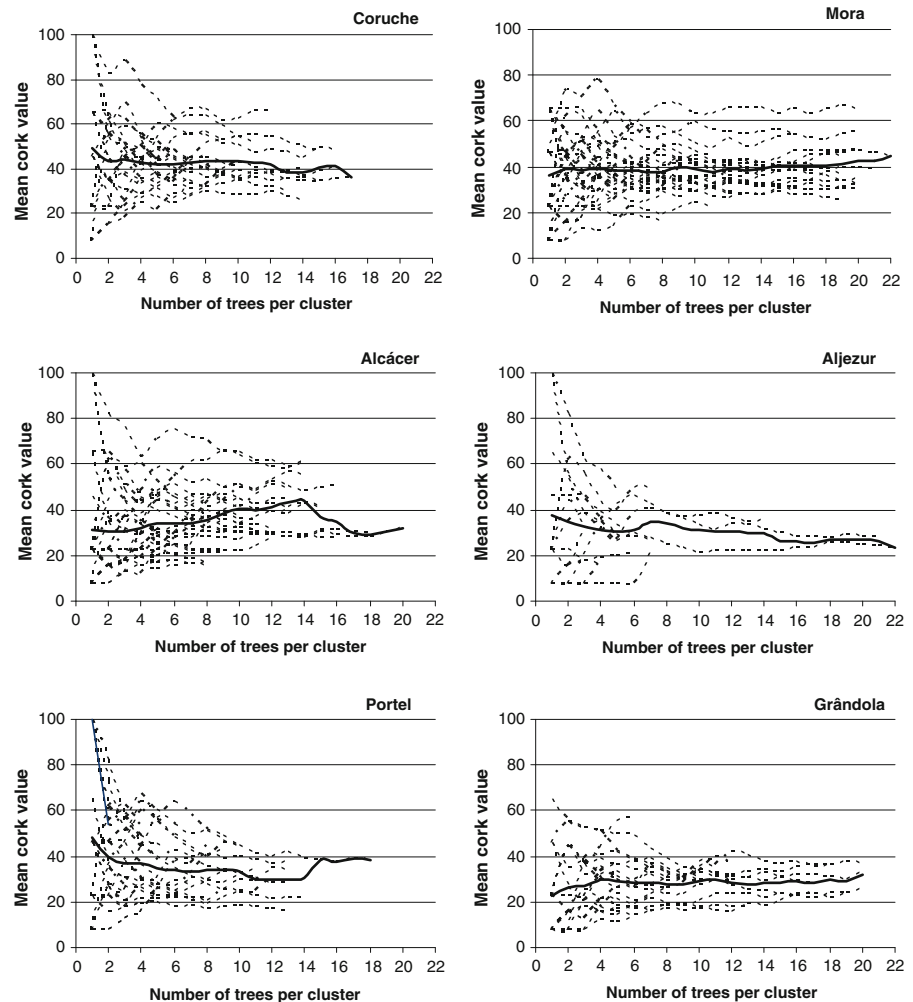
**Results**

Figure 1 shows the mean cork value per cluster for clusters of different sizes and the respective average for each cluster size, for the six stands. The stands do not show exactly the same pattern, but there is a slight decrease in variability until the number of trees per cluster is close to 7–8. From then on, the mean cork value per cluster is more or less stable. However, the most important conclusion that can be drawn from this figure is the high variability that can be observed between clusters from the same stand.

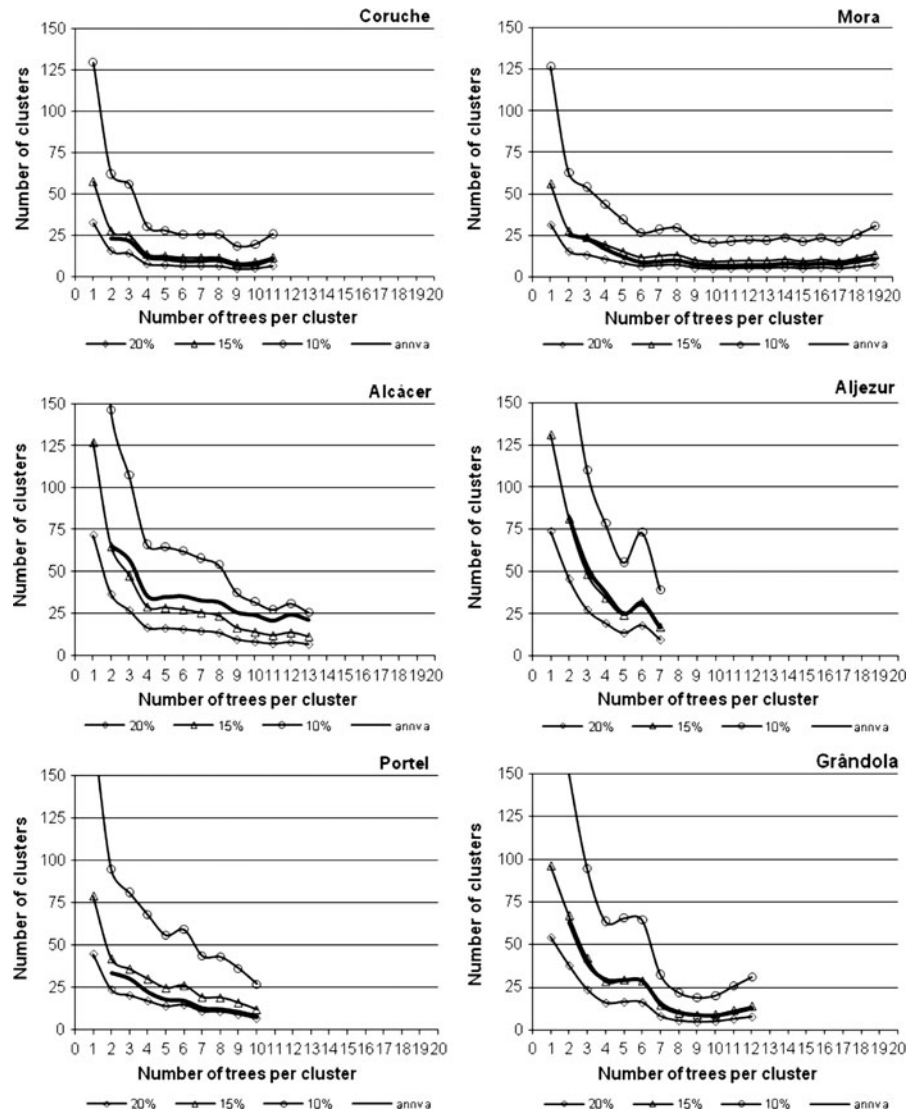
The mean cork value is different for the six stands, being lower in Grândola and Aljezur and higher in Coruche and Mora. The mean cork value is more or less independent from cluster size, except in Aljezur where it decreases with cluster size.

Sample sizes for different cluster sizes and different percent sampling errors are shown in Fig. 2. For every percent error considered, sample size decreases sharply with cluster size until a size of 5–7 trees, depending on the stand. From then on, decrease is less marked and, from a cluster size of 9–10 trees, there is no apparent decrease in the sample size needed to achieve a certain precision. For a fixed total number of trees sampled, the sampling error increases with the size of the cluster (Fig. 3), although this dependence is not strong for small cluster sizes. With the exception of Portel, the results show that, for

**Fig. 1** Mean cork value per cluster (*plot*) for clusters with different sizes in the six stands



**Fig. 2** Sample sizes needed to achieve pre-fixed percent sampling errors with clusters of different sizes for the six stands. Sample sizes based on the intracluster correlation coefficient (ANOVA) for a sampling error of 15%



clusters with a size smaller than five, there is not a big difference in the sampling error for different sampling designs if the total number of trees sampled is the same. This conclusion is important for the cost of the sampling, as it is preferable to sample less clusters (due to the decrease in the cost of travelling between clusters) even if they are slightly bigger. For a cluster size of 5–7 trees, a reasonable error rate (10–15%) only holds if more than 100 trees in total are sampled (Fig. 3).

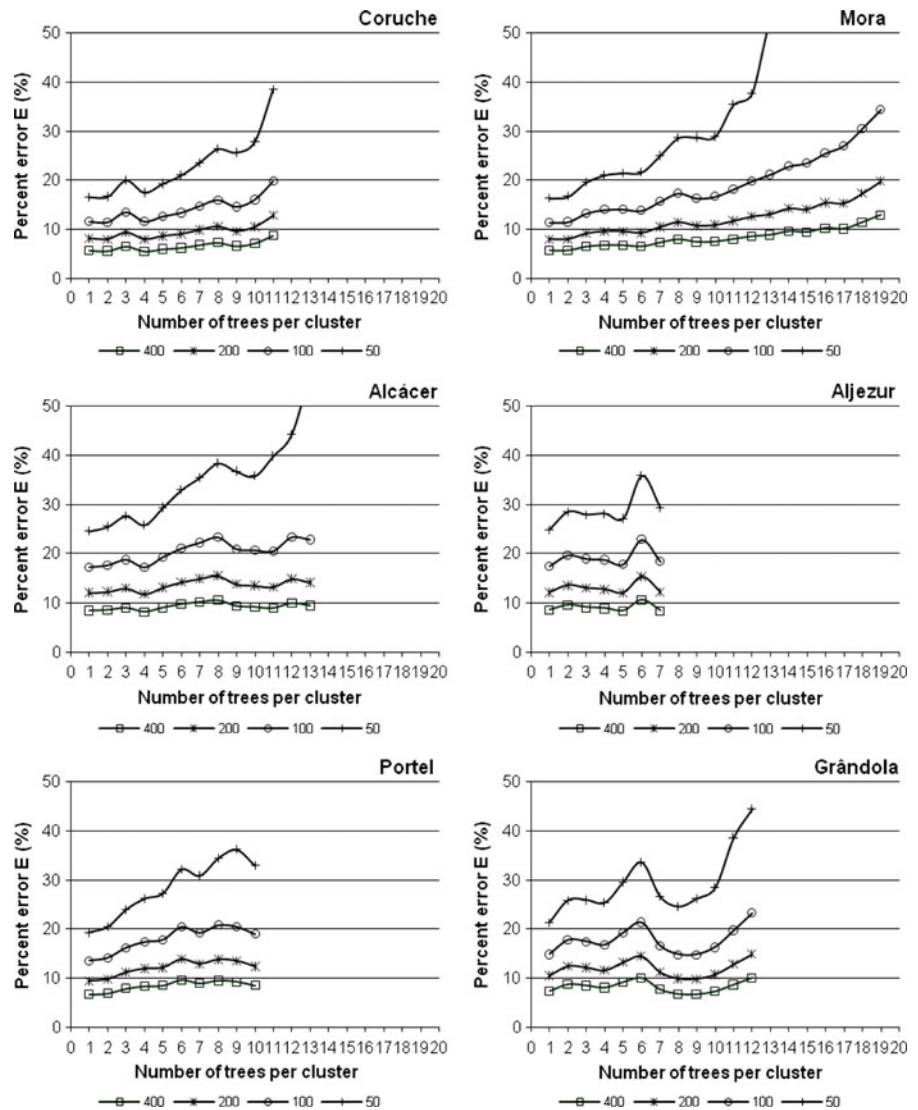
The intracluster correlation coefficient for the six stands is shown in Fig. 4. With the exception of Aljezur and Grândola, the intracluster correlation coefficient is close to zero irrespective of cluster size,

indicating the random distribution of cork quality. Therefore, in these stands, the single-stage cluster sampling is more efficient than random sampling. Even for Aljezur and Portel, the intracluster correlation is not very high, and it tends to zero with increasing cluster size. Figure 2 shows, for a sampling error of 15%, the sample sizes estimated by the analysis of the intracluster correlation coefficient. Sample sizes estimated by this method are slightly bigger in Alcácer and lower in Portel stands, but the pattern shown is quite similar.

Total time needed for field sampling implementation is shown in Fig. 5 for a sampling error of 15%. It was computed using the sample size estimated by



**Fig. 3** Percent sampling error for a fixed total number of trees sampled (400, 200, 100 and 50) in clusters with different sizes for the six stands



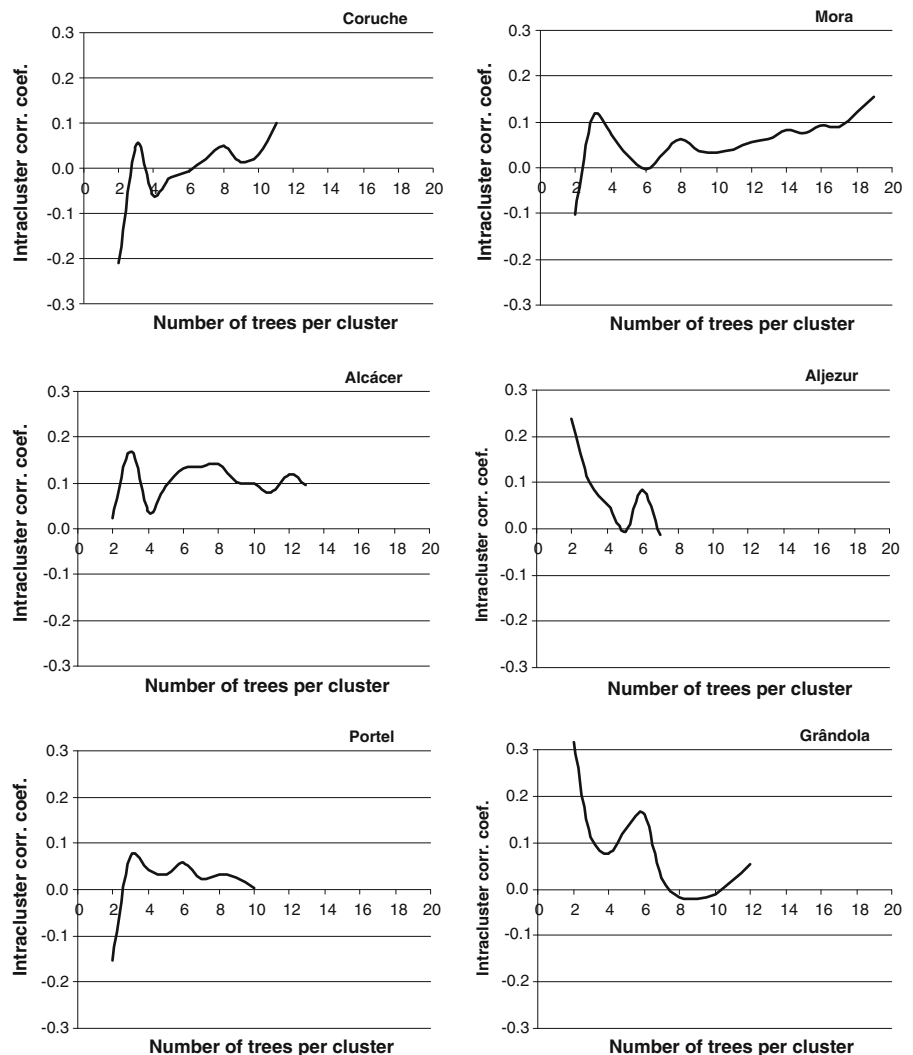
simulation of samplings with different cluster sizes. Time needed for travelling ( $t_1$ ) decreases with the number of trees per cluster, as the total number of clusters needed to achieve the pre-fixed %E decreases. On the contrary, the time needed to sample trees ( $t_2$ ) increases with the number of trees per cluster. The total time needed to achieve a pre-fixed %E depends on the stand (Fig. 5). However, two general conclusions can be drawn: (1) the use of cluster sampling implies a decrease in the time needed to implement field sampling if the size of the cluster is small, but this advantage is lost if the size of the cluster is too big; and (2) there is a tendency to show a minimum for a cluster size around 5–7 trees.

Total time needed for field sampling for a sampling error of 10% was also computed. Results were similar to those shown in Fig. 5 but with higher total sampling times. Generally, when %E decreased from 15 to 10%, total sampling time increased by an average of 49.5%.

**Discussion**

The objective of the research reported here was to propose a sampling scheme that could be generally used for forest landowners to estimate cork value in the field before extraction. The proposed method

**Fig. 4** Intracluster correlation coefficient for the six stands



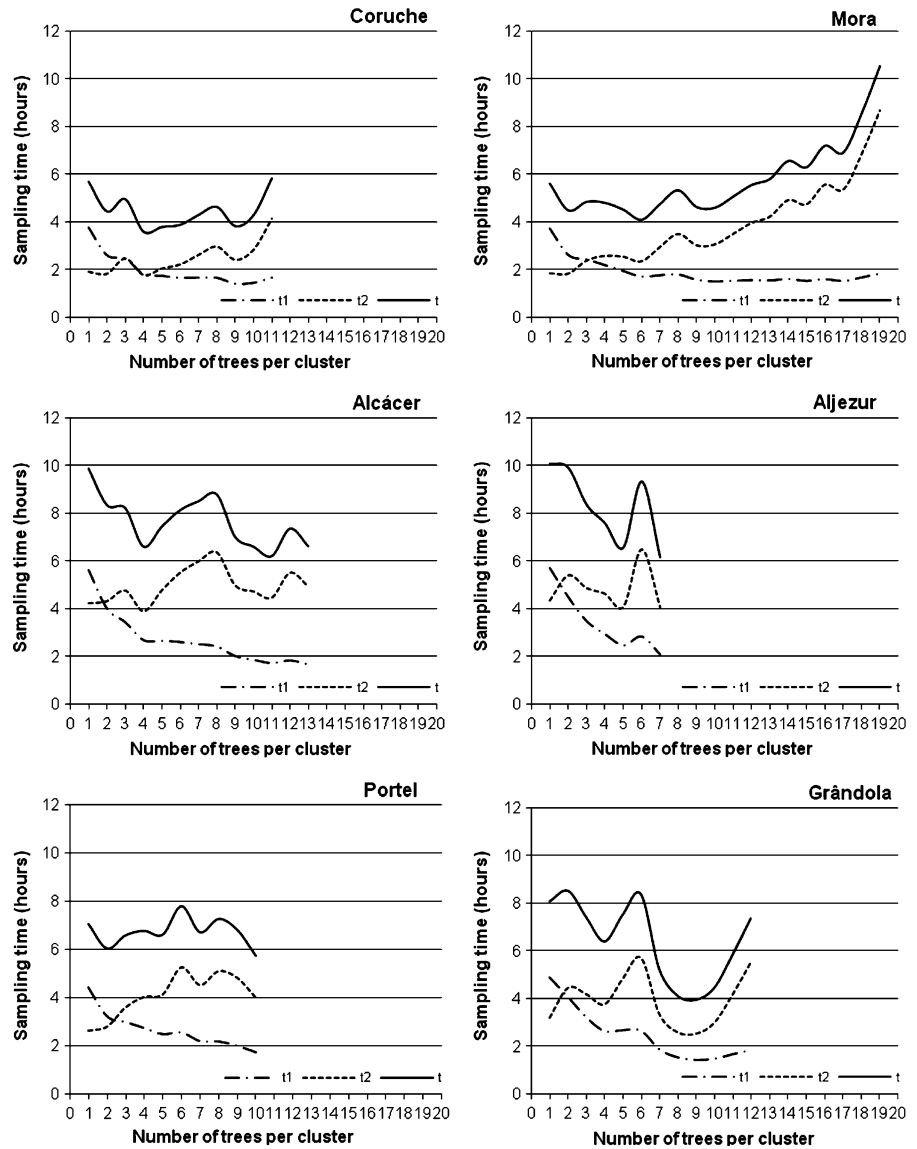
should provide a reasonable sampling error (10–15%), but it must be undertaken within 1 day in order to have a cost compatible with the method previously in use, which provided unreasonable results.

The literature is lacking studies for optimising sampling schemes for cork value estimation. However, general sampling theory, namely single-stage cluster sampling, could be used in order to find an appropriate sampling scheme. It was found that, relative to cork value, most of the stands analysed have a spatial pattern very close to random, indicating that a single-stage cluster sampling is more efficient than a random sampling with the same bulk. In order to find the optimum number and

size of clusters, different samplings were simulated and the corresponding total sampling times were estimated.

Total sampling time—the time to travel between clusters and the time to sample the trees that are part of the selected clusters—depends on the characteristics of each particular stand, but we concluded that using clusters of 5–7 trees was close to the optimum solution. The analysis of total sampling time did not consider that the single-stage cluster sampling could incorporate car transportation, which is very useful as the cork samples are not easily carried. Therefore, in operational samplings, a reduction in the estimated sampling times is expected.

**Fig. 5** Total and partial time estimated for the field sampling for different clusters sizes ( $t$ , total time;  $t_1$ , time needed for travelling;  $t_2$ , time needed to sample trees)



**Conclusions**

The research presented here produced some important conclusions, common to all the studied stands:

1. The spatial distribution of cork value (or cork quality) was approximately random, as shown by the intracluster correlation coefficient, and indicated that single-stage cluster sampling is more efficient than random sampling of the same bulk;
2. Single-stage cluster sampling with a cluster size of 5–7 trees guarantees a reasonable sampling error (10–15%);
3. Sampling errors around 15% can be obtained with ~20–30 primary units (clusters);
4. It is possible to sample one homogeneous stand, with a reasonable sampling error ( $\leq 15\%$ ), with a total sampling time compatible with 1 day of work;
5. The recommendation for forest landowners is to use clusters of five trees whose number depends on the information available before the sampling takes place:
  - a. if the stand has been measured in previous cork extractions, the number of primary units can be estimated from the formula

$$n = \frac{s_Q^2 z_{\alpha/2}^2}{\%E/100 \bar{Q}}$$

where  $\bar{Q}$  and  $s_Q^2$  are estimated from the previous sampling.

- b. if the stand was never measured, the number of primary units selected should be between 20 and 30, based on cork quality variability assessed by an expert.

From the results presented here, we propose a methodology that can be generally applicable for the estimation of cork value in the field before cork is extracted, which will be very important for forest landowners.

**Acknowledgments** The research reported here is part of the PhD of the first author supported by grant SFRH/BD/17733/2004 of the Portuguese Foundation for Science and Technology (FCT) and was supported by the CarbWoodCork project (POCI/AGR/57279/2004): “Simulation of the effect of different forest management strategies and climate change on wood/cork and carbon sequestration for the most important species of Portuguese forest”. Data were collected under the project SUBER-DEMO (AGRO No 81 2001/2004): “Demonstração da gestão de montados de sobre apoiada em inventário florestal e modelos de crescimento e produção”. We acknowledge unpublished cork price data provided by the Coruche Forest Producers Association (APFC).

## References

- Adrados JRG, Pereira H (1996) Classification of defects in cork planks using image analysis. *Wood Sci Technol* 30:207–215
- Apcor (2008a) Sector da cortiça em números 2008. Available <http://www.apcor.pt>. Accessed 27 Oct 2008
- Apcor (2008b) Notícias Apcor—Julho, Agosto e Setembro 2008. Available <http://www.apcor.pt>. Accessed 27 Oct 2008
- Cochran WG (1977) *Sampling techniques*, 3rd edn. Wiley, New York
- Corona P, Dettori S, Filigheddu MR, Maetzke F, Scotti R (2005) Site quality evaluation by classification tree: an application to cork quality in Sardinia. *Eur J For Res* 124:37–46
- DGF (1993) Método DGF 5—Cortiça de reprodução crua, em árvores dispostas ao acaso. Colheita de amostras para estimativa da sua repartição percentual, após cozedura, por calibres e qualidades. Divisão de Estudos, Lisboa
- 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
- Loetsch F, Haller KE (1973) *Forest inventory*, vol 1. BLV Verlagsgesellschaft München, Bern, Wien (trad. inglês de E. F. Brünig)
- Natividade JV (1950) *Subericultura*. Ministério da Economia, Direcção Geral dos Serviços Florestais e Aquícolas, Lisboa
- Pereira H (2001) Relatório final do projecto Corkassess. Centro de estudos Florestais. ISA, Lisboa
- Pereira H (2007) *Cork: biology, production and uses*. Elsevier, Amsterdam
- Pereira H, Tomé M (2004) Cork oak and cork. In: Burley J, Evans J, Yongquist JA (eds) *Encyclopedia of forest sciences*. Elsevier, Amsterdam, pp 613–620
- Schreuder HT, Gregoire TG, Wood GB (1993) *Sampling methods for multiresource forest inventory*. Wiley, New York
- Shiver BD, Borders BE (1996) *Sampling techniques for forest resource inventory*. Wiley, New York
- Tomé M (2004) Modelo de crescimento e produção para a gestão do montado de sobre em Portugal. Relatório final do projecto POCTI/AGR/35172/99. Publicações GIMREF RFP 1/2004. Universidade Técnica da Lisboa, Instituto Superior de Agronomia, Centro de Estudos Florestais, Lisboa
- Tomé M, Barreiro S, Cortiçada A, Paulo JA, Meyer A, Ramos T, Malico P (2007) Inventário florestal 2005–2006. Áreas, volumes e biomassas dos povoamentos florestais. Resultados Nacionais e por NUT's II e III. Publicações GIMREF. RT 5/2007. Universidade Técnica de Lisboa. Instituto Superior de Agronomia, Centro de Estudos Florestais, Lisboa
- Vries PG (1986) *Sampling theory for forest inventory. A teach-yourself course*. Springer, Dordrecht
- Zeide B (1980) Plot size optimisation. *For Sci* 26(2):251–257