Site quality assessment of degraded *Quercus frainetto* stands in central Greece

Kyriaiki Kitikidou (1), Elias Milios (1), Elias Tsirekis (1), Elias Pipinis (2), Athanasios Stampoulidis (1)

The potential yield of a site is measured by site index, which is defined as the dominant height of a stand at a base age. A site index model for site quality assessment of *Quercus frainetto* (Hungarian oak) stands in central Greece was developed using a base age of 50 years. Data were collected from 39 temporary sample plots of 10 x 10 m. Linear regression models widely used in site index studies were fitted to height-age data. The adjusted coefficient of determination ($R^2_{\text{adj}}$), root mean square error (RMSE), bias, coefficient of determination for the prediction ($R^2_{\text{pred}}$) and residual plots were used for the choice of the best-fitting model. The best model was $H = -0.231 + 0.251A - 0.001A^2$, where $H$ is the predicted height at age $A$. The guide curve method was adopted in constructing the sites curves, with the chosen model as the guide curve. Based on this curve, the study area was divided into three site quality classes (I to III), with class I representing the best and class III the poorest. Also, the presence of a Simpson’s paradox in these analyses is discussed. The results showed that a 50-year-old stand in the study area attained an average dominant height of about 11, 8 and 6 m on site quality classes I, II and III, respectively. The Hungarian oak stands of the present study can be considered very low productivity stands.

Keywords: Guide Curve Method, Hungarian Oak, Simpson’s Paradox, Site Quality

Introduction

Site quality assessment is the evaluation of the natural productive capacity of a forest site for a tree species. Site quality assessment is very important in forest management, because a site could support one species excellently while supporting poorly other species. The oldest and most widely used technique for evaluating site quality or productivity is the site index. It is very important in forest management, because a site could support one species excellently while supporting poorly other species.

Site index is the average total height of dominant and co-dominant trees at a specified reference or base age (Clutter et al. 1983). It is scarcely affected by various densities and species composition, relatively stable under various thinning intensities and is strongly correlated with volume. The principal uses of site index are: (1) to estimate the height at any given age from site index; (2) to estimate site index from height at any given age; (3) to stratify forest land into productive classes; and (4) to provide entry to yield tables (Wang & Payandeh 1995). However, the most common objective of site index investigations is to determine the height development pattern that the stand is expected to follow throughout its life-cycle (Clutter et al. 1983).

Site index models are essential quantitative tools in forest management (Teshome & Pety 2000). Most of the techniques used to fit site index curves could be considered as special cases of three general development models, namely: (i) the guide curve; (ii) the difference model; and (iii) the parameter prediction method (Clutter et al. 1983). Among the various techniques for developing site index curves, the guide curve method has become popular in those cases where repeated measures through time on the same permanent plots are unavailable (Clutter et al. 1983, Nanag & Nunifum 1990).

Oak forests covers 22.6% of the total forested area in Greece (Ministry of Agriculture 1992). According to Christensen (1997), Hungarian oak (*Quercus frainetto* Ten. syn. *Q. conferta* Kit.) is an important timber species in most mainland Greece, but not in the drier low elevation area in the southern part of the country. Moreover, Bergmeier & Dimpopulos (2008) refer that in mainland Greece *Q. frainetto* is the species that most frequently dominates in comparison with the other deciduous oaks.

The aim of this work was to obtain site index curves for site quality assessment of degraded Hungarian oak stands in central Greece to be used in the management of current and future Hungarian oak stands.

Materials and Methods

Study area

The study was conducted in degraded Hungarian oak stands located in an area of approximately 290 ha (of which 260 are forested) in central Greece (38º 53 ‘N, 22º 03’ E - Fig. 1). Data were collected at an elevation ranging from 650 to 750 m a.s.l. The substratum is flysch and the soil texture is loamy sand, clay-loam to loam (Xirogiannis 2001). According to the closest meteorological station (390 m a.s.l.), the mean annual rainfall is 873 mm and the mean annual air temperature is 15.16 °C (Xirogiannis 2001).

In the past, tree branches in some locations of the area were being cut by residents of nearby villages for livestock feeding, while other places were under agricultural cultivation. Nowadays, the main disturbance is grazing; degree of grazing is low to absent, and only in a small part of the area is severe. Silvicultural treatments have not been applied in the area, mainly because of the lack of adequate forest roads in most of the study area.

Experimental design

Thirty-nine plots 10 x 10 m were established randomly in the forested part of the study area, where the degree of grazing is low to absent, and the main canopy trees were considered approximately even-aged (Fig. 1a, Fig. 1b). In each plot, the height of all trees was measured. The tallest dominant tree was cut in each plot. As a result, the 39 sample trees were among the 100 tallest trees per hectare (100 x plot area = 1 ha), so they can be used for the site index estimation (Beekhuis 1966). Many of the sample trees are possibly sprout origin. Cross sectional discs were cut from each felled tree at the ...
ground level (≤0 m), at the breast height (1.3 m) and at 1 m intervals. The last disc was collected from the 4-cm bole diameter. Annual growth rings were measured along the mean radius on each cross-sectional disc using the LINTAB® system (RinnTech, Heidelberg, Germany) and the software package TSAP-Win® (Rinn 2003). As for stem analysis, tree height at each age was calculated using a modified version of Carmeans’ algorithm (Carmean 1972, Newberry 1991).

Development of growth equations

Little information is available in the scientific literature on site index for degraded Hungarian oak stands in central Greece; thus the first step was the choice of the appropriate base age. As recommended by several authors (Curtis et al. 1974, Trousdell et al. 1974, Teshome & Petty 2000), the appropriate base age for site index determination was set to 50 years based on the culmination age of mean annual increment. Among the various techniques for developing site index equations, the guide curve method was adopted in this study since data was obtained from temporary sample plots (Clutter et al. 1983, Nanang & Nunifum 1999). Three linear equations commonly used in site index studies were selected from the forestry literature (Carmean 1972, Clutter et al. 1983, Teshome & Petty 2000) and used for data fitting (eqn. 1 to 3):

\[ H = b_0 + b_1 A + b_2 A^2 + e \]
\[ H = b_0 + b_1 A + b_2 A^{1/2} + e \]
\[ H = b_0 + b_1 \ln(A) + e \]

where \( H \) is the tree height at the age \( A \), \( b_0 \), \( b_1 \), \( b_2 \) are the regression coefficients to be estimated, and \( e \) is the error. The above equations were then compared and the best fitting model was chosen. Numerous criteria for selecting the best model are available; however, since the dependent variable in the above equations was the same, the highest adjusted coefficient of determination \( \left( R^2_{adj} \right) \) and the lowest absolute and relative root mean square error (RMSE) were considered appropriate criteria for the selection of the best model. Examination of the residuals was also carried out for assessing the best models by calculating absolute and relative biases (Vanclay 1994, Soares et al. 1995, Gadow & Hui 1998). Selection criteria were (eqn. 4 to 8):

\[ R^2_{adj} = 1 - \frac{n - 1}{n - p} \sum_{i=1}^{n} \left( H_i - \hat{H}_i \right)^2 \]
\[ RMSE = \sqrt{\frac{\sum_{i=1}^{n} \left( H_i - \hat{H}_i \right)^2}{n-1}} \]
\[ RMSE\% = \left( \frac{\sum_{i=1}^{n} \left( H_i - \hat{H}_i \right)}{\sum_{i=1}^{n} H_i} \right) \times 100 \]
\[ bias = \frac{\sum_{i=1}^{n} \left( H_i - \hat{H}_i \right)}{n} \]
\[ bias\% = \left( \frac{\sum_{i=1}^{n} \left( H_i - \hat{H}_i \right)}{\sum_{i=1}^{n} \hat{H}_i} \right) \times 100 \]

where \( H_i \) is the observed height value for the \( i \)-th observation, \( \hat{H}_i \) is the estimated height value, \( \hat{H} \) is the mean of the observed values, \( n \) is the number of observations, and \( p \) is the number of regression coefficients to be estimated.

The predictive performances of the studied models were also evaluated using prediction errors or PRESS (PRediction Sum of Squares) residuals. These residuals were calculated by omitting each observation in turn from the dataset, fitting the model to the remaining observations, predicting the response without the omitted observation and comparing the prediction with the observed value (eqn. 9):

\[ H_i - \hat{H}_{i,-i} = e_{i,-i} \]

where \( H_i \) is the observed height value and \( \hat{H}_{i,-i} \) is the prediction value obtained by omitting the \( i \)-th observation \((i = 1, 2, ..., n)\). Each candidate equation has \( n \) PRESS residuals associated with it and the PRESS statistics is defined as (Myers 1986 - eqn. 10):

\[ PRESS = \sum_{i=1}^{n} \left| H_i - \hat{H}_{i,-i} \right|^2 = \sum_{i=1}^{n} \left| e_{i,-i} \right|^2 \]

The accuracy of the predictions obtained with the different candidate equations was assessed by computing the coefficient of determination for the prediction (eqn. 11):
The interpretation of this coefficient of determination is similar to that described for other similar coefficient, i.e., a value close to 1 indicates high predictive performances for the model considered.

Stands were stratified into three site quality classes, with site quality decreasing from class I to III. Using the predicted height growth trend described by the chosen model (guide curve), anamorphic site index curves were fitted to pass through three different site index classes of I to III. This was accomplished by holding the shape parameters constant and varying the asymptote parameter as necessary to achieve the required dominant height when stand age (4) equals the base age ($A_0 = 50$ years). This approach has been successfully used in previous studies (Clutter et al. 1983, Teshome & Petty 2000).

Results and Discussion

A summary of the height characteristics of the 39 sample trees used for site index modeling are reported in Tab. 1. Mean total height ranged from 7.2 m at 38 years to 10.86 m at 59 years.

The parameters and statistics of the regression models considered (eqn. 1 to 3) are presented in Tab. 2. The initial observation was that all the models had high $R^2_{adj}$ and low RMSE and bias. All parameters included in the models were statistically significant, indicating that none of the models was overparameterized. As expected, residuals did not display any significant departure from the normal distribution (Fig. 2). A comparison of the three linear models revealed similar fitting performances, though the model of eqn. 1 showed the lowest RMSE and bias (Tab. 2) and therefore it was selected for the construction of the site index curves for Hungarian oak stands in the study area. The predicting curve from eqn. 1 was within the observed variation of the dominant height-age data (Fig. 3). The site index curves for the study area showed that a 50-years-old Hungarian oak stand attained an average dominant height of about 11, 8 and 6 m on site classes I, II and III, respectively (Fig. 4).

Simpson’s paradox (known also as the Yuille-Simpson effect, the reversal paradox, or the amalgamation paradox) is an apparent paradox in which the results of separate groups seem reversed when the groups are combined (Cohen 1986). This paradox is primarily an issue of aggregated statistical analysis where separate groups are analyzed together. In other words, it occurs when a variable has an effect in the same direction within subsets of the data, but when the subsets are combined, the effect vanishes or reverses in direction. The variable used to separate these groups/subsets is called “confounder” (Wagner 1998, Abramson et al. 1992). Cohen (1986) describes the uncertainty principle associated with Simpson’s paradox arising when two populations are stratified. The stratification process is undertaken to increase homogeneity among the strata. However, stratification may reverse the apparent rank ordering of the two populations means.

In our study, the confounder is the site index, and the Simpson’s paradox concerns the total height estimated in contrast to the base age (50 years). Even though models’ parameters are correctly estimated, it seems that the estimated height at the age of 50 is lower than the true height measured at an age below 50, and higher than the true height measured at an age above 50.

Such apparent paradox may also be attributable to a King Kong effect (Makridakis et al. 1998). The King Kong effect was named from a theoretical study on the relationship of weight to height in a sample of gorillas in which one extreme value (i.e., King Kong) is added. The addition of such extreme value to the sample of gorillas skews the height and weight data and affects their correlation. As a result, a weight/height relationship greater

---

**Table 1** - Summary statistics of the sample trees (dominant trees) used in site index modeling. (SE): standard error of the mean.

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>Count</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>35</td>
<td>1</td>
<td>9.5</td>
<td>9.5</td>
<td>9.5</td>
<td>-</td>
</tr>
<tr>
<td>38</td>
<td>1</td>
<td>7.2</td>
<td>7.2</td>
<td>7.2</td>
<td>-</td>
</tr>
<tr>
<td>43</td>
<td>1</td>
<td>7.5</td>
<td>7.5</td>
<td>7.5</td>
<td>-</td>
</tr>
<tr>
<td>47</td>
<td>1</td>
<td>8.7</td>
<td>8.7</td>
<td>8.7</td>
<td>-</td>
</tr>
<tr>
<td>50</td>
<td>3</td>
<td>9.2</td>
<td>9.5</td>
<td>9.25</td>
<td>0.252</td>
</tr>
<tr>
<td>51</td>
<td>1</td>
<td>7.2</td>
<td>7.2</td>
<td>7.2</td>
<td>-</td>
</tr>
<tr>
<td>52</td>
<td>1</td>
<td>7.7</td>
<td>7.7</td>
<td>7.7</td>
<td>-</td>
</tr>
<tr>
<td>54</td>
<td>7</td>
<td>7.2</td>
<td>11.5</td>
<td>9.31</td>
<td>0.489</td>
</tr>
<tr>
<td>55</td>
<td>1</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>-</td>
</tr>
<tr>
<td>57</td>
<td>7</td>
<td>6.8</td>
<td>10.5</td>
<td>8.8</td>
<td>0.492</td>
</tr>
<tr>
<td>58</td>
<td>3</td>
<td>6.5</td>
<td>10.5</td>
<td>7.97</td>
<td>1.272</td>
</tr>
<tr>
<td>59</td>
<td>5</td>
<td>7</td>
<td>12.5</td>
<td>10.86</td>
<td>1.002</td>
</tr>
<tr>
<td>60</td>
<td>1</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>-</td>
</tr>
<tr>
<td>63</td>
<td>1</td>
<td>9.2</td>
<td>9.2</td>
<td>9.2</td>
<td>-</td>
</tr>
<tr>
<td>64</td>
<td>2</td>
<td>10.2</td>
<td>10.5</td>
<td>10.35</td>
<td>0.15</td>
</tr>
<tr>
<td>66</td>
<td>3</td>
<td>9.6</td>
<td>11.5</td>
<td>10.33</td>
<td>0.59</td>
</tr>
</tbody>
</table>

---

**Table 2** - Estimated parameters and statistics for the three models fitted to temporary plot data for Hungarian oak stands.

<table>
<thead>
<tr>
<th>Model</th>
<th>Coeff.</th>
<th>Parameter estimates</th>
<th>$R^2_{adj}$</th>
<th>RMSE</th>
<th>% bias</th>
<th>bias %</th>
<th>$R^2_{pr}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$b_0$</td>
<td>-0.231</td>
<td>0.799</td>
<td>1.414</td>
<td>26.093</td>
<td>-2.94E-13</td>
<td>0.799</td>
</tr>
<tr>
<td></td>
<td>$b_1$</td>
<td>0.251</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$b_2$</td>
<td>-0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>$b_0$</td>
<td>0.919</td>
<td>0.788</td>
<td>1.451</td>
<td>26.784</td>
<td>1.77E-12</td>
<td>0.788</td>
</tr>
<tr>
<td></td>
<td>$b_1$</td>
<td>0.161</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$b_2$</td>
<td>-1.404</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>$b_0$</td>
<td>-3.656</td>
<td>0.700</td>
<td>1.727</td>
<td>31.874</td>
<td>1.32E-12</td>
<td>0.700</td>
</tr>
<tr>
<td></td>
<td>$b_1$</td>
<td>2.949</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

Fig. 2 - Residual Q-Q plots for the three regression models considered in this study.
than that existing before adding the extreme value is obtained (Makridakis et al. 1998). In our study, the site (arbitrarily defined from the general tree development) could be the King Kong value. Adding this variable a smoother height growth potential of the tree is obtained, i.e., a tree with an estimated height at base age which classifies it in a higher site quality, actually belongs to a lower site quality.

The absence of permanent plots necessitated the use of temporary plot data of different ages. This limitation permitted only the use of the guide curve method for the development of site index models. Therefore, it was assumed that the distribution of plots with respect to site quality was the same across all ages (Alder 1980). The failure of this assumption would lead to the fitted guide curve being biased away from the true average height growth trend (Walters et al. 1989). It appears in Fig. 3 that the assumption was not violated, as the fitted guide curve did not appear to be biased away from the true average height growth trend of the species under investigation.

The guide curve method was used to generate anamorphic site index curves. These curves are a family of lines with constant slope but various intercepts, achieved by holding the shape parameters constant and varying the asymptote parameter to obtain the required dominant height value when stand age (4) equals the index age (4x). The base age of 50 years used for the development of site index curves in the present study seems appropriate, since base age corresponds to the period of completion of rapid growth, i.e., when mean annual increment is culminated (Trousdell et al. 1974).

Fig. 4 presents the site index curves developed for the species under investigation. Areas of good site quality are those characterized by a high height growth (Clutter et al. 1983). The best site quality class was site class I, while the medium and the poorest site quality classes were II and III respectively. This implies that a stand with height development pattern similar to curve SI11 is growing on a better site than a stand with height development pattern similar to curve S18. A stand with a mean dominant height of about 4 m at age 15 years is expected to follow the height development pattern exhibited by class I, while a stand with a mean dominant height of about 3 m at the age of 15 years is expected to follow the height development pattern exhibited by class II. However, there could be a shift in the site index, caused by climatic variation, measurement error, damage to dominant height by diseases and fire, as well as removal of dominant trees through thinning (Teshome & Petty 2000).

Even though there are more productive sites in adjacent areas, the Hungarian oak stands of the present study are not productive. Tree height growth, even in the best site quality, is very low. Matis (2000), using dominant and codominant trees, presented site index curves for Hungarian oak derived from coppice stands in the University forest of Taxiarchis in Chalkidiki (northern Greece). Tree growth rates were significantly higher as compared with the results of the present study. In particular, he distinguished five site qualities, noting that the tree height is up to 11.72 m in the V site quality at the breast height age of 50 years, while in the best (I) site quality trees showed a height of at least 21.68 m. On the other hand, when he distinguished three site qualities, trees of the third (III) site quality at the same breast height age had a height up to 13.93 m, and trees of the best (I) site quality were at least 19.47 m tall. In Hungarian oak stands from approximately the same area, Smiris et al. (1992), using stem analyses of 15 trees that had the dimensions of the mean bole, noted that the height of the “mean bole” tree was 19.17 m at the age of 49 years. Smiris et al. (1998), in experimental plots in sprout origin Hungarian oak stands, where different intensity thinnings had been applied, notice that, at the age of 52 years, the mean height of trees was ranging between 14.18 (control plot) and 16.19 m (light thinning plot). In north-eastern Chalkidiki, Dafis (1966) distinguished the site index curves for Hungarian oak stands constructed using the model of eqn. 1. (SI): Site Index. (SI6), (SI8), (SI11): average height of dominant trees at the age of 50 years on the sites III, II, and I, was 6, 8 and 11 m, respectively.
six site qualities for coppice Hungarian oak stands. Average height of trees of the main stand in the V site quality, at the age of 40 years, was the same as the height of trees, at the age of 50 years, in the best site quality of the previous inventory.

Collecting data for a longer period of time from permanent sample plots may allow site index curves to be validated. Indeed, a good knowledge of growth and yield is necessary for sound forest management plans, especially silvicultural decision-making processes such as time of first and subsequent thinnings. The site index curves developed in this study may be useful in quantifying Hungarian oak growth and yield on various low productivity site conditions. Site index curves may also help as a guide in making thinning decisions for stands of species in analogous low productivity conditions. Among the various methods used to determine the time of first and subsequent thinnings, the dominant height approach is preferred because it reflects site effect more than other methods (e.g., age, basal area and volume methods - Teshome & Petty 2000).

As for Simpson’s paradox, it has been mentioned in several forestry researches (Thomas & Parresol 1989, Scheiner et al. 2000, Loh 2006, Piñeiro et al. 2006). This can happen due to a shift in the distribution of the stands from a previous survey to a more recent survey, and not from a fundamental decline in tree growth, in case of successive forest inventories. In such a case, we can assume that the shift is due to the cutting in saw timber stands that removed the large mature trees, resulting in a reclassification of the stands in the next inventory (Thomas & Parresol 1989).

In our study, Simpson’s paradox can be a key to understanding the apparent inconsistency between height growth rate per tree and the performance of stand level statistics. Measurements of per tree growth rates over time, when used as a whole for site index curve fitting, may give misleading estimates. Actually, foresters must always recognize the effect that each felled tree may have on the forest level growth. Interpretation of our site index curves development, with the awareness of Simpson’s paradox, indicates that results from one region should not be extrapolated to another one too far apart on the basis of equations or graphs showing height growth rate by age class. Analysis of growth data for large areas can be risky, considering that the tree-stand interaction over time is quite composite.

Conclusions

From the analysis of degraded Quercus prinus stands in central Greece, we obtained site index curves showing the height development pattern that the stands are expected to follow. These site index curves can be used to quantify the production potential of Hungarian oak in low productivity sites. However, before their application outside the study area, the model must be validated using data from permanent sample plots. Lastly, the comprehension of the Simpson’s paradox and a cautious analysis of data can increase our understanding of forest growth rates and stand processes.

Acknowledgements

We would like to thank the Bodossaki Foundation for the financial support of A. Stampoulidou. We also want to thank the Forest Service of Sperchitaia (Greece) for their kind cooperation.

References


