Site index curves for young *Populus tremula* stands on Athos Peninsula (northern Greece)

Kyriaki KITIKIDOU*, Minas KAYMAKIS, Elias MILIOS

Department of Forestry and Management of the Environment and Natural Resources, Democritus University of Thrace, Pantazidou 193, 682 00 Orestiada - GREECE

Received: 22.06.2010

**Abstract:** Difference equations derived on the basis of the McDill-Amateis differential functions and from the integral form of the Bertalanffy, Richards, and Korf growth functions were used to model the dominant height growth of young European aspen (*Populus tremula* L.) stands on Athos Peninsula (northern Greece). Data from stem analysis were used for fitting. Both numerical and graphical analyses were used to compare alternative models. The cross-validation approach was used to analyze the predictive ability of the models. The algebraic difference form of the differential function proposed by McDill and Amateis resulted in the best compromise between biological and statistical aspects and produced the most adequate site index curves. Therefore, it is recommended for height growth prediction and site classification of European aspen stands on Athos Peninsula. This equation is base-age invariant, so any number of points ($A_1, H_1$) on a specific site curve can be used to make predictions for a given age $A_2$ and the predicted height $H_2$ will always be the same.

**Key words:** Algebraic difference equation, height growth, site index

**Introduction**

The classification of forest land in terms of its productivity is an important issue for forestry managers as well as for forestry enterprise administrators. An index that expresses this productivity is a required variable for the modeling of present and future growth and yield and can also be used for forest land stratification for the purposes of forest inventory and forest exploitation on a sustainable yield basis (García 1983). Conceptually, site quality is considered an inherent property of plots of land, whether or not trees are being grown at the time of interest (Beaulieu et al. 2011). Site index models that include a good description of the state of the stand at any point in time, in addition to change rates as a function of this state, are able to adequately predict stand development (Garcia 1994), evaluate forest instant changes due to thinning, and assess forest productivity (Álvarez González et al. 2002). The development of a flexible growth model could facilitate the decision making required to guarantee sustainable management. The modeling level is often dictated by available data, the desired prediction level, and the time horizon for projection (Burkhart 2003).

The productivity of a specific stand can vary greatly due to a host of factors including the

---

* E-mail: kkitikid@fmenr.duth.gr
underlying soil conditions, climatic variables, and management practices (Beaulieu et al. 2011). For timber production purposes, site quality is commonly expressed as a species-specific site index (Carmean 1975; Davis et al. 2001). Site index may be defined as the height, at a predetermined age, of dominant trees that have always been dominant and healthy (Goelz and Burk 1992). Empirical evidence from thinning experiments indicates that, for many commercially important species, height growth is not greatly affected by the manipulation of stand density (Krieger 1998). The average height of the stand may be affected by thinning, depending on the method used. Within the limits of stand density, however, height growth appears to be unaffected, particularly when the comparison is restricted to dominant trees (Clutter et al. 1983).

The objective of the present study was to develop site index curves for European aspen stands on Athos Peninsula (northern Greece) with an approach designed to account for the desired features of the models discussed above. Today, poplars are planted for different purposes in both the northern and southern hemispheres (FAO 2001). The largest area of planted poplars was reported in China (6 million ha), whereas 6 other countries (France, Hungary, Turkey, Italy, Spain, and the USA) reported more than 100,000 ha of poplar plantations (FAO 2001). In Greece, Populus tremula is not forming forests or very large stands. It is found sporadically as a component of damp mixed forests at altitudes between 600 and 1500 m (Christensen 1997). In a few cases, it has been observed at lower altitudes. Poplars are typically trees with a high production potential, consistently allocating gathered resources to growth-related processes (Mattson et al. 2001). They produce leaves continuously and are able to take advantage of a prolonged growing season or favorable growth conditions (Dickmann et al. 2001). This distinguishes poplars from other species with determinate growth or multiple flushing growth patterns (Dickmann et al. 2001). The approach of this study was applied to young trembling aspen trees growing on a variety of sites. Evaluation of the selected model was performed by using a validation dataset.

Materials and methods

Data

The study was carried out in the winter of 2007-2008 in a broad area of approximately 234 ha in the Simonos Petras monastery forest of Athos Peninsula in northern Greece (40°12'40.74''N, 24°15'25.53''E). The forest stands are located at elevations between 460 and 893 m. The main species is Castanea sativa, and it is managed under the coppice system. Moreover, there are small stands or groups of Populus tremula. These stands and groups are not managed. Most Populus tremula trees are probably root sprouts. They were initiated after the clear cutting of small areas as well as the cutting of individual trees. The substratum in almost all areas is mica schist and granite, and the soils vary from sandy clay to clay sandy (Siamidis 2006). The soil depth in the Populus tremula stands and groups ranges from deep to shallow (Siamidis 2006). On average, the annual precipitation at the Arnea station, the closest meteorological station, is 651 mm, and the mean yearly temperature is 12.4 °C.

By using simple random sampling, 59 dominant trees were selected. A tree can be characterized as dominant if it has a well-developed crown extending above the general level of the crown cover. In addition, dominant trees are fairly crowded on the sides, receive full light from above and partial light from the sides, and are bigger than the average trees in the stand (Smith et al. 1997). Cross-sectional disks were cut from each tree and removed at the 0.3 level, breast height (1.3 m) level, the 2.3 level, and at 2-m intervals up the bole. The last disk was collected at a bole diameter of 5 cm. These disks were taken to the laboratory to measure the number of rings. In each cross-sectional disk, the number of annual growth rings was counted using the LINTAB system of RINNTECH and the TSAP-Win program (Rinn 2003). For height calculation in particular, the improved version of Carmean’s formula was used (Carmean 1972; Newberry 1991).

Summary statistics, including number of observations, mean, standard deviation, and minimum and maximum values, were calculated for the total tree height variable grouped by age classes (Table 1).
According to Clutter et al. (1983), most techniques for site index curve construction can be viewed as special cases of 3 general methods: 1) the guide curve method, 2) the parameter prediction method, and 3) the difference equation method. Although the 3 methods are not mutually exclusive, the difference equation method has been the preferred form for developing site index curves (Álvarez González et al. 2005).

The difference equation method makes direct use of the fact that observations corresponding to a given dominant tree should belong to the same site curve (Parresol and Vissage 1998). A height-by-age equation can be differentiated to provide an equation for height growth rather than accumulated height. An equation in this form is referred to as an algebraic difference equation (Parresol and Vissage 1998). In this method, height \( H_2 \) at age \( A_2 \) is expressed as a function of \( A_2 \), height \( H_1 \) at age \( A_1 \), and \( A_1 \). The expression is obtained through substitution of one parameter in the growth model. The choice of this parameter determines the behavior of the model, which is capable of producing anamorphic or polymorphic (with a single asymptote) curve families (Bailey and Clutter 1974).

With the difference equation method, short observation periods of stem analysis from trees whose total age is under or over the reference age can be used, the curves pass through the site index at the reference age, and they are base-age invariant (Cieszewski and Bailey 1999). The invariant or unchanging property refers to predicted heights; any number of points \( (A_i, H_i) \) on a specific site curve can be used to make predictions for a given age \( A_2 \), and the predicted height \( H_1 \) will always be the same. This includes forward and backward predictions and the path invariance property, which ensures that the result of projecting first from \( A_i \) to \( A_j \) and then from \( A_j \) to \( A_k \) is the same as that of the one-step projection from \( A_i \) to \( A_k \) (Cieszewski 2003). Equations derived using this technique define both height-growth and site index models as special cases of the same equation (Cieszewski 2003).

### Function selection

Growth functions describe variations in the global size of an organism or a population with age; they can also describe the changes in a particular variable of a tree or a stand with age, in this case dominant height (Curtis et al. 1974). Site index curves are used to indicate the potential productivity of forest land (Curtis et al. 1974). The height growth curves define the average pattern of height development for the tallest trees in stands of a given site quality and are appropriately used for construction of yield tables (Curtis et al. 1974). The most important desirable attributes for site index models are polymorphism, a sigmoid growth pattern with an inflexion point, a horizontal asymptote at old age, logical behavior (height should be zero at age zero and equal to site index at the reference age), and base-age invariance (Parresol and Vissage 1998). The fulfillment of these attributes depends on both the construction method and the mathematical function used to develop the curves.

A total of 3 models were selected for fitting the relationship between height and age (Table 2), where \( \hat{h}_i \) is estimated height at age \( a_i \), \( h_1 \) is height at index age (m), \( h_2 \) is the estimated height at total age (m), \( a_i \) is the index age (years), \( a_2 \) is tree age (years), and \( b_i \) values are regression coefficients. We fitted the

### Table 1. Total tree height statistics (m) by age class (years).

<table>
<thead>
<tr>
<th>Age class</th>
<th>Number of observations</th>
<th>Mean</th>
<th>Standard deviation</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>15-20</td>
<td>1</td>
<td>10.00</td>
<td>-</td>
<td>10.00</td>
<td>10.00</td>
</tr>
<tr>
<td>20-25</td>
<td>22</td>
<td>16.16</td>
<td>4.29</td>
<td>10.40</td>
<td>24.50</td>
</tr>
<tr>
<td>30-35</td>
<td>3</td>
<td>26.37</td>
<td>1.29</td>
<td>25.20</td>
<td>27.75</td>
</tr>
<tr>
<td>35-40</td>
<td>1</td>
<td>25.00</td>
<td>-</td>
<td>25.00</td>
<td>25.00</td>
</tr>
</tbody>
</table>
Site index curves for young *Populus tremula* stands on Athos Peninsula (northern Greece)

The Chapman-Richards model (M1; Chapman 1961), the Korf model (M2; Lundqvist 1957), and the McDill-Amateis model (M3; McDill and Amateis 1992) have been widely used to develop height/age curves (Trincado et al. 2002; Calama et al. 2003; Diéguez Aranda et al. 2005). All 3 candidate models were fitted to the data from young aspen trees by applying the following bootstrapping technique. To fit a model to a given set of sample data, model parameters were estimated using nonlinear regression. The loss function was defined as the sum of squared prediction errors (residuals), and its global minimum was sought using a conjugate gradient method (Press et al. 1994).

### Model comparison and model selection

There is no single criterion for choosing the best regression model through a number of models (Draper and Smith 1998). In this study, 5 criteria were used; these are described in Table 3.

### Table 2. Algebraic difference models considered.

<table>
<thead>
<tr>
<th>Model</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chapman-Richards</td>
<td>$\hat{h}_2 = b_0 \left( \frac{h_i}{b_0} \right)^{\frac{\ln(1-e^{-b_1}y_i)}{\ln(1-e^{-b_1})}}$</td>
</tr>
<tr>
<td></td>
<td>$\hat{h}_i = b_0 \left( 1 - e^{-b_1}y_i \right)$</td>
</tr>
<tr>
<td>Korf</td>
<td>$\hat{h}_2 = b_0 e^{-b_1} \frac{\ln \left( \frac{h_i}{b_0} \right)}{\ln \frac{h_i}{b_0}}$</td>
</tr>
<tr>
<td></td>
<td>$\hat{h}_i = b_0 e^{-b_1} y_i$</td>
</tr>
<tr>
<td>McDill-Amateis</td>
<td>$\hat{h}_2 = \frac{b_0}{1 - \left( 1 - \frac{b_0}{h_i} \right)^{a_1} / a_1}$</td>
</tr>
<tr>
<td></td>
<td>$\hat{h}_i = \frac{b_0}{1 + \frac{b_0}{a_1}}$</td>
</tr>
</tbody>
</table>

### Table 3. Statistics for model fitting (training data).

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Absolute mean error</th>
<th>Standard error of estimate</th>
<th>Fit index</th>
<th>Relative mean squared error %</th>
<th>Regression coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Formula</td>
<td>$\sum_{i=1}^n</td>
<td>h_i - \hat{h}_i</td>
<td>/ n$, $\sqrt{\frac{1}{n-p} \sum_{i=1}^n (h_i - \hat{h}<em>i)^2}$, $1 - \frac{1}{\sqrt{\sum</em>{i=1}^n (h_i - \hat{h}<em>i)^2 / n}}$, $\frac{1}{\sqrt{n}} \sum</em>{i=1}^n \left( \frac{h_i - \hat{h}_i}{\hat{h}_i} \right)^2$, $100$</td>
<td>$\alpha$, $\beta$</td>
<td></td>
</tr>
<tr>
<td>Optimum</td>
<td>0.0024</td>
<td>14.3856</td>
<td>0.6545</td>
<td>15.7578</td>
<td>-0.1004, 1.0091</td>
</tr>
<tr>
<td>M1</td>
<td>0.0706</td>
<td>14.4470</td>
<td>0.6530</td>
<td>12.5418</td>
<td>0.1433, 0.9936</td>
</tr>
<tr>
<td>M3</td>
<td>0.0163</td>
<td>14.3760</td>
<td>0.6547</td>
<td>15.7397</td>
<td>-0.0870, 1.0062</td>
</tr>
</tbody>
</table>

$h_i$ = observed values of the dependent variable;  
$\hat{h}_i$ = expected values of the dependent variable;  
n = number of observations;  
p = number of regression coefficients;  
$\hat{h}$ = arithmetic mean of the observed dependent variable.
After various tests and comparisons, it is likely that a chosen model will seem to be the best but in fact be unstable. Another sample of the population may thus lead us to a completely different regression model (a completely different model or different regression coefficients for the same model). As a result, after assessing the regression coefficients, the stability of the model should be controlled. Therefore, before giving a regression model for general use, some controls that relate to the behavior of the environment should be tested where the model is to be used. The predictive ability of the selected regression model is of particular interest. Predictive ability or reliability is the percentage of variation of the dependent variable explained by the existence of independent variables in the model when the independent variables take new values, values that did not participate in training the model. These statistics are known as validation criteria (Draper and Smith 1998). For site index model validation purposes, data were split into 2 groups. The first group (80% of the trees) was used for fitting (training) data, and the second group (20% of the trees) was used for validation data. This validation method is known as cross-validation (McCarthy 1976). The same statistics used for fitting purposes were calculated.

Apart from these statistics, one of the most efficient ways of ascertaining the overall picture of model performance is by visual inspection. Graphical analyses, consisting of plots of observed against predicted values of the dependent variable and plots of studentized residuals against the predicted dominant height were carried out. These graphs are useful for detection of possible systematic discrepancies (Neter et al. 1996). Additionally, graphs showing the appearance of the fitted curves overlaid on the trajectories of stem analysis over time were examined.

Practical use of the models to estimate site quality from any given pair of heights and ages requires the selection of a base age to which the site index will be referenced. Conversely, the site index and its associated base age could be used to estimate dominant height at any desired age. Selection of the base age for the site index equations was made while considering that the base age should be chosen so that it is a reliable predictor of height at other ages. In order to address this consideration, different base ages and their corresponding observed heights were used to estimate heights at other ages for each tree. The results were compared with the values obtained from stem analyses, and the average deviation was then calculated.

**Results**

Statistics for the fitting and the cross-validation phases are shown in Tables 3 and 4, respectively. All model parameters were found to be significant at a 5% level. Parameter values were as follows.

### Table 4. Statistics for model validation.

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Absolute error</th>
<th>Standard error of estimate</th>
<th>Fit index</th>
<th>Relative mean squared error %</th>
<th>Regression coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Formula</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sum_{i=1}^{n} \frac{</td>
<td>h_i - \hat{h}_i</td>
<td>}{n}$</td>
<td>$\sqrt{\frac{\sum_{i=1}^{n} (h_i - \hat{h}_i)^2}{n-p}}$</td>
<td>$1 - \frac{\sum_{i=1}^{n} (h_i - \hat{h}<em>i)^2}{\sum</em>{i=1}^{n} (h_i - \bar{h})^2}$</td>
<td>$100 \frac{\sum_{i=1}^{n} \left(\frac{h_i - \hat{h}_i}{h_i}\right)^2}{n}$</td>
</tr>
<tr>
<td>Optimum</td>
<td>0.0051</td>
<td>10.6594</td>
<td>0.7442</td>
<td>15.0882</td>
<td>-0.0639 1.0057</td>
</tr>
<tr>
<td>M1</td>
<td>0.0896</td>
<td>10.7987</td>
<td>0.7409</td>
<td>11.3694</td>
<td>0.2021 0.9906</td>
</tr>
<tr>
<td>M2</td>
<td>0.0070</td>
<td>10.6467</td>
<td>0.7445</td>
<td>15.4524</td>
<td>-0.0810 1.0061</td>
</tr>
</tbody>
</table>

$h_i =$ observed values of the dependent variable;  
$\hat{h}_i =$ expected values of the dependent variable;  
$n =$ number of observations;  
$p =$ number of regression coefficients;  
$\bar{h} =$ arithmetic mean of the observed dependent variable.
Site index curves for young *Populus tremula* stands on Athos Peninsula (northern Greece)

Best values for each criterion are highlighted in Tables 3 and 4. These results suggest that, for this species, the McDill-Amateis model (M3) is the most suitable*.

As previously noted, visual or graphical inspection of the models is considered an essential point in selecting the most accurate representation. Therefore, plots showing the site curves for heights of 13, 19, and 25 m at 25 years overlaid on the trajectories of observed values over time (Figure 1) were examined. Height-age pairs from selected dominant trees were used in curve fitting. We observe that the behavior of all 3 models was similar.

In selecting the base age, it was found that a base age of 25-30 years was superior for predicting height at other ages with a minimum of reliability (Figure 2). Even though at older ages the error in predicting

---

**M1:**

\[
\hat{h}_2 = 45.610 \left( \frac{h_1}{45.610} \right)^{\ln(1-e^{-0.016a})/\ln(1-e^{-0.016b})}
\]

**M2:**

\[
\hat{h}_2 = 105.693e^{0.045a} - 5.045
\]

**M3:**

\[
\hat{h}_2 = \frac{96.081}{1-\left(1-\frac{96.081}{h_1}\right)^{0.836}}
\]

---

*Figure 1. Plots showing the site curves for heights of 13, 19, and 25 m at 25 years overlaid on the trajectories of observed values over time: a) M1 model, b) M2 model, c) M3 model.*

* As regards the regression coefficients a and b in Table 4, M1 and M3 models have the best values. However, there is no dilemma in choosing the McDill-Amateis model (M3), since the model selection is based on Table 3, Table 4 does not forbid our selection and the graphical tests confirm our choice.
height was lower, the scarcity of data would lead to an incorrect decision as the data were not representative enough. According to Goelz and Burk (1992), this selection procedure should be devised so that variance of the potential volume estimates for the forest of interest is minimized. Nevertheless, the lack of necessary information forced us to conclude that a reference age of 25 years is appropriate for European aspen on Athos Peninsula.

Discussion

In this study, site index model selection was viewed as a compromise between biological and statistical considerations. Model M3, the algebraic difference form of the differential function proposed by McDill and Amateis (1992), produced curves with an adequate graphical behavior as well as good values for goodness-of-fit statistics. Based on 5 statistical criteria and 2 graphical analyses, we propose its use for height growth prediction and site classification of European aspen stands on Athos Peninsula (northern Greece).

On Athos Peninsula, we observed a high growth rate and a potential for rapid volume growth of Populus tremula. Mean height for dominant trees at 40 years (reference age) was 23 and 26 m at the best sites in Norway and Sweden (Opdahl 1992, as referenced in Johansson 1996; Johansson 1996), while in our study this height was reached at 25-30 years. Average height for Populus tremula was up to 1.5 m at 4 years of age in another Swedish study (de Chantal and Granström 2007), while in our study, this height was reached at 1-3 years. The growth rate in our study is more similar to that of an American study regarding other poplar species. Pallardy et al. (2003) found that Populus nigra and Populus deltoides had an average height of 1.71 m during the first year, and in the second year, mean height almost tripled.

---

Figure 2. Relative mean squared error in height predictions related to choice of reference age: a) M1 model, b) M2 model, c) M3 model.
A graphical comparison with height growth models for European aspen in Sweden (Johansson 1996) showed a higher growth rate for young (up to 10 years old) trees in our study (Figure 3). This may be due to different environmental conditions (moisture, nutrients, heat and light, and length of growing season) or genetic background.

Acknowledgements

The authors are grateful to Simonos Petras Monastery (Athos Peninsula, northern Greece) for help and permission in data collection.

References


Cieszewski C (2003) Developing a well-behaved dynamic site equation using a modified Hossfeld IV function \( Y = \frac{a(x^n)}{c + x^{m-1}} \), a simplified mixed-model and scant subalpine fir data. For Sci 49: 539-554.


K. KITIKIDOU, M. KAYMAKIS, E. MILIOS


