

Modelling dominant height growth of Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco) in Portugal

LUIS FONTES¹, MARGARIDA TOMÉ², MARTA BAPTISTA COELHO², HOWARD WRIGHT¹, JAIME SALES LUIS³ AND PETER SAVILL¹

¹ Oxford Forestry Institute, University of Oxford, South Parks Road, Oxford OX1 3RB, England

² Instituto Superior de Agronomia, Universidade Técnica de Lisboa, Tapada da Ajuda 1399, Lisboa Codex, Portugal

³ Secção Florestal, Universidade de Trás-os-Montes e Alto Douro, Quinta de Prados, Apartado 1013-5000-911, Vila Real Codex, Portugal

*Corresponding author. E-mail: luisfontes@isa.utl.pt

Summary

The first plantations of the North-West American conifer Douglas-fir were established in Portugal in 1904. Investigations into growth and yield patterns were started in 1969. Since then others have carried out further studies. This study includes data from the previous studies and covers the whole range of site conditions where Douglas-fir grows in Portugal. Dominant height data from 87 stem analysis trees from 12 sites were used to model dominant height growth of Douglas-fir using the McDill–Amateis, Chapman–Richards and Lundqvist–Korf growth functions. The Chapman–Richards and Lundqvist–Korf growth functions were used in their integral and difference forms. For the evaluation of the candidate growth models' performance, three steps were adopted: (1) all the candidate growth equations were fitted with the data available from stem analysis from 87 trees, eliminating growth equations with non-logical and non-biological consistency and poor statistical properties; (2) the remaining growth equations were cross-validated using two sub-samples of the stem analysis data and selection of the growth equations with the best statistical results; (3) validation, using the whole data set from stem analysis for fitting growth models and the stand data for calculation of the prediction errors, was carried out. Out of the nine models evaluated five were rejected in the first step and two in the second step. The two best models had similar results in the third step and were compared with previous Douglas-fir site index curves. The McDill–Amateis function performed best.

Introduction

To obtain relevant information for prudent decision-making in forest management and forest policy, it is necessary to understand the growth patterns of a species and to assess site quality

(Avery and Burkhart, 1994). Site quality can be defined as the timber production potential of a site for a particular tree species or forest type (Clutter, 1983). Thus, site quality has meaning only with respect to the one or more species that may be considered at a particular site. For

example, a given site might have an excellent site quality for maritime pine (*Pinus pinaster* Ait.) and be very poor for Douglas-fir. Additionally, for species such as Douglas-fir, provenance trials show that site index can vary with provenance.

The most widely used indirect method for assessing site quality in single species, even-aged stands is based on the dominant height–age relationship and is termed the site index (Clutter, 1983; Tomé, 1988; Avery and Burkhart, 1994; Savill *et al.*, 1997). The dominant height of a uniform stand, at a given age, is a good indicator of the potential productivity of that type of forest on that particular site (Cailliez and Alder, 1980). This is based on Eichhorn's hypothesis (Eichhorn, 1904) that total production from a fully stocked stand, which is the volume currently standing plus anything removed in previous thinnings, is a function of its height (Savill *et al.*, 1997).

Site index curves are families of curves representing dominant height and are developed from data collected from as wide a range of sites as possible. The total range of dominant heights at the base-age is divided into a number of classes, each having an equal but limited range of heights, usually 2 or 3 m (Philip, 1994). There are three methods for site index curve construction (Clutter, 1983): (1) the guide curve method; (2) the difference equation method; and (3) the parameter prediction method.

Before the development of computers, the most common way of constructing site index curves was to use the guide curve method, a graphical version of the current guide curve version (Spurr, 1952; Curtis, 1964; Cailliez and Alder, 1980; Tomé, 1988). The data for dominant height and age are plotted on a graph and an average curve, or guide curve, is drawn by hand. From this, other curves are drawn at proportional distances from the guide curve. This method originates an anamorphic family of curves. Using it for any two curves, the height of one at any age is a constant proportion of the height of the other at the same age. There are two other types of curve families: polymorphic disjoint and polymorphic non-disjoint curves. Neither has a constant proportionality relationship but in the disjoint curve family the curves do not cross within the age range of interest. In the polymorphic curve family it is assumed that the shape of the height growth curve is dependent on the site quality, whereas in

the anamorphic family it is assumed to be independent. The height of the site index curve relates to the site index and the shape to the pattern of growth of the species.

The difference equation method requires the development of a difference form of the height/age equation being fitted. Therefore it is necessary to re-write the height/age equation to express height at re-measurement (h_{t_2}) as a function of re-measurement age (t_2), initial measurement age (t_1) and height at initial re-measurement (h_{t_1}). It is a flexible method and can be applied to any height/age equation to produce anamorphic or polymorphic curve families (Clutter, 1983).

The parameter prediction method involves fitting linear or nonlinear functions to the data on a tree-by-tree or plot-by-plot basis. It requires preliminarily knowledge of the site index or site class of the plot before fitting the model. Two types of models have been used, constrained and unconstrained models. Using constrained models the dominant height (h_{t_0}) is forced to equal site index (S) when the age is equal to the index age (t_0). Using unconstrained models the curves must be conditioned after fitting to ensure that the dominant height corresponds to the site index at the index age (Cailliez and Alder, 1980).

According to Furnival *et al.* (1990), the three methods are identical for fitting a site index equation that can be transformed into a simple linear model. However, these results cannot be applied to a majority of current site index models, which are non-linear (Cao, 1993). In addition, although the parameter prediction method has the advantage of great flexibility in the type of model adopted, it requires an *a priori* estimation of site index (Cailliez and Alder, 1980). With stem analysis data, the site index for each tree is usually known or assessed by interpolation, provided that the tree is older than the base-age. The problem is that the derived site index curves are dependent on the base-age which in the future may be changed (the other two methods are base-age invariant). For these reasons the parameter prediction method was not considered for modelling Douglas-fir dominant height growth in Portugal, only the guide curve and difference equation methods seemed suitable.

Following model construction, it is necessary to test the model to determine its validity and

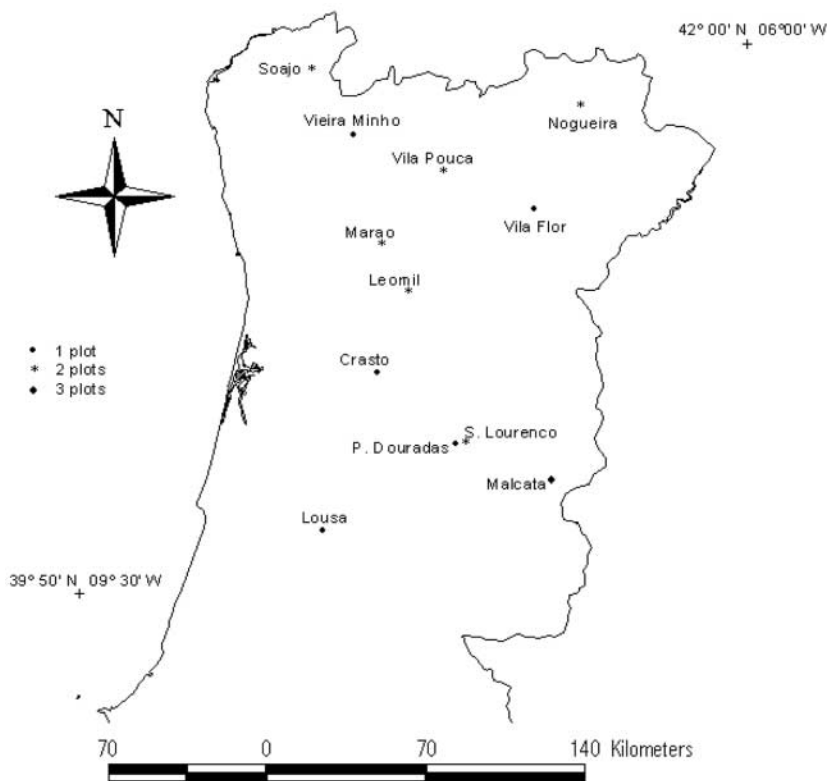


Figure 1. Study area: locations of the 20 plots that were established. Each contained 30–103 trees and was 589–2808 m² in area.

precision. This is best done using data collected independently from that used in the model construction, which was not used to fit any of the functions in the model (Cailliez and Alder, 1980; Myers, 1990; Soares *et al.*, 1995).

Studies have already been made of Douglas-fir dominant height growth in Portuguese stands (Diniz, 1969; Loureiro, 1986; Luis, 1989). However, Diniz (1969) based his work only on temporary sample plots and very few sites, while Loureiro (1986) based his on stem analysis for just one site (Serra de Vila Pouca), and Luis (1989) his on stem analysis over a range of sites using predominantly young stands. The aim of this work was to construct site index curves based on stem analysis covering the whole range of site conditions where Douglas-fir grows in Portugal.

Materials and methods

Study area and preliminary field measurements

Data for this study came from the re-measurement of existing plots complemented with data gathered in new plots. One plot established in 1988 by Costa (1992) at Lousã was re-measured. All except two of the plots established by Luis (1989) were re-measured. These plots range from 600 to 2800 m² and from 75 to 150 trees. One of the two excluded plots had been burnt, the other had had the dominant trees thinned. New plantations were chosen if they had uniform stocking and an area that permitted the establishment of a plot of 80–100 trees, excluding the edges. The location of the plots is shown in Figure 1 and a summary of stand data is given in Table 1.

Table 1: Stand variables of the Douglas-fir plots from which the stem analysis were selected

Site	Year of measurement	Plot no.	Age (years)	N (tree ha ⁻¹)	Hdom (m)	QM d.b.h. (cm)	G (m ² ha ⁻¹)
Aveleira	1986	1	23	1717	12.6	14.9	29.9
		2	23	1571	13.1	15.2	28.5
		3	23	1845	13.9	15.8	36.2
Bornes	1986	1	14	1400	8.5	12.7	17.7
		2	14	1231	8.5	12.0	13.9
		3	14	1344	8.6	12.6	16.8
Crasto	1999	1	27	929	18.9	20.9	31.7
Estrela	1985	1	47	848	30.9	32.7	71.2
		2	47	1431	27.5	23.9	64.2
		3	47	939	33.2	31.1	71.3
	1999	2	61	640	32.3	29.5	57.2
		3	61	574	37.0	35.5	53.5
		1	25	1353	15.0	17.2	31.4
Leomil	1986	2	25	1423	15.2	17.2	33.1
		3	25	1281	14.9	16.6	27.7
		2	38	1283	23.3	23.0	53.5
	1999	3	38	1185	22.6	23.2	33.4
		1	61	316	33.4	46.3	53.2
		1	21	1011	11.1	14.3	16.2
Lousa	1999	2	19	1034	6.9	8.2	5.5
Malcata	1999	3	20	1280	13.0	14.5	21.2
		1	21	1664	13.2	19.0	47.2
		2	21	1545	12.8	18.2	40.2
Marao	1986	3	21	1406	12.3	18.1	36.2
		1	34	1578	21.3	23.4	67.7
		3	34	1357	20.1	26.1	72.6
Nogueira	1986	1	15	875	9.7	16.9	19.6
		2	15	967	9.7	15.5	18.2
		3	15	893	9.7	16.1	18.2
	1999	1	28	776	20.6	30.1	55.0
		2	28	951	20.8	25.5	48.7
		1	48	324	31.4	45.9	53.6
Padrela	1986	2	48	403	26.4	36.9	43.1
		3	48	415	29.6	34.5	38.8
		1	62	274	38.4	52.8	60.1
	1999	2	62	287	32.8	45.2	46.0
		1	94	298	35.0	51.1	61.2
		1	35	938	25.3	25.1	46.4
P. Douradas	1985	2	35	1008	25.8	24.2	46.4
		3	35	695	23.3	26.6	38.6
		1	49	588	30.6	32.8	49.5
Soajo	1999	2	49	595	31.3	32.7	50.0
		1	28	927	15.3	18.9	26.1
V. Flor	1999	1	28	927	15.3	18.9	26.1
V. do Minho	1999	1	36	692	28.0	24.6	33.0

The 1985 and 1986 measurements are from Luis (1989).

Stem analysis

There are references to the use of stem analysis from at least as early as 1898 (Młodziansky,

1898). Stem analysis involves the analysis of a complete tree stem and entails counting and measuring the annual growth rings on a series of cross-sections taken at different heights, to

determine past rates of growth and changes in stem form. It can be used with species where it is possible to recognize annual growth rings.

In all but one plot, two dominant trees were selected and felled, using the method of Tennent and Burkart (1981), giving a total of 39 trees. This method demonstrated that the two trees with the d.b.h. closest to the mean diameter of the 100 largest stems per hectare and with measured heights within 5 per cent of the mean dominant height are closely related to the plot mean dominant height development. Felling was done ~0.10 m above the ground and the exact height of the first disk was recorded. Tree age was obtained by counting the rings in the base of the stem analysis trees. Sectioning was carried out through percentages of total height. This method was used because, in even-aged plantations, using percentages of height, the disks of trees are likely to have similar ages. Precision increases with the number of the disks sampled. Two levels of sampling were adopted since there were big and old trees and small and young trees. If the same number of disks had been used for all trees the intensity of sampling of the older trees would have been much smaller. Hence for larger trees 12 disks were cut: at the base (0.1 m), d.b.h. (1.3 m), 10 per cent, 15 per cent, 25 per cent, 35 per cent 45 per cent, 55 per cent, 65 per cent, 75 per cent and 85 per cent of total height, and at 7 cm top diameter. For smaller trees, 10 disks were cut; those at 10 per cent and 85 per cent were not cut. In general for smaller trees, the disk at 10 per cent of total height was very close to the d.b.h. disk and the disk at 85 per cent of total height was very close to 7 cm top diameter.

Correction of heights

When stem analysis is used to determine height/age pairs the determination of past height growth presents a problem. Because of the conical growth pattern exhibited by trees, the true height of a tree at the age corresponding to the ring count at a cross-cut will almost always be located some distance above the cross-cut (Dyer and Bailey, 1987).

For estimating the true height at a given age several proposed algorithms exist: graphic (Mlodziansky, 1898), Graves' (1906), Lenhart's (1972), Carmean's (1972), Newberry's (1991)

and Ratio (Brister and Schultz). Fabbio *et al.* (1994) proposed another method named ISSA. Dyer and Bailey (1987) in a comparison of the first six methods concluded that Carmean's method (1972) was the most accurate. Fabbio *et al.* (1994) compared ISSA, Lenhart's (1972), Carmean's (1972) and branch whorl methods. They concluded the ISSA method is the most precise if the interval between cross-cuts is ~0.5 m and Carmean method is best at an interval of 2 m. Since the average interval between cross-cuts was 2.2 m, Carmean's method was chosen for height corrections in this work.

Carmean's method can be expressed as follows:

$$h_{ij} = b_i + [(b_{i+1} - b_i)/(r_i - r_{i+1})]/2 + (j - 1) * [(b_{i+1} - b_i)/(r_i - r_{i+1})] \quad (1)$$

where h_{ij} = estimated total height for growth ring j based on section point i , b_i = height at the i th section point, r_i = number of growth rings at the i th section point, j = growth ring number (assuming the pith as the starting point), $j = 1, \dots, r_i$. In this study the modification proposed by Newberry (1991) by subtracting 0.5 years, in formula (1), for the portion of the tree bole above the highest section point, was used.

Candidate growth equations for dominant height modelling

It was aimed to have growth equations for modelling growth in dominant height with desirable attributes such as polymorphism, inflection point, asymptote, logical behaviour and base-age invariance (Tomé, 1988; Goelz and Burk, 1992; Elfving and Kiviste, 1997). The equations most often used for site index purposes have been the Chapman-Richards (Mitscherlich, 1919) and Schumacher (1939) growth equations. Thus, the candidate growth equations considered were: Chapman-Richards (Mitscherlich, 1919); Lundkvist-Korf (Korf, 1939; Lundqvist, 1957) which is a generalization of the Schumacher equation; and McDill-Amateis (McDill and Amateis, 1992).

Chapman-Richards growth equation

$$Y = A(1 - Ce^{-kt})^{\frac{1}{1-m}} \quad (2)$$

where A is the asymptote of dominant height, k is a growth related parameter, t is the age of the stand and m is a shape parameter. The parameter C ($C = e^{-kt_0}$ was considered 1, since $t_0 = 0$).

Lundkvist–Korf growth equation

$$Y = Ae^{-\frac{k}{t^n}} \quad (3)$$

where A is the asymptote of dominant height, k is a growth-related parameter, t is the age of the stand and n is a shape parameter.

McDill–Amateis growth equation

$$Y_2 = \frac{A}{1 - \left(1 - \frac{A}{Y_1}\right) \left(\frac{t_1}{t_2}\right)^a} \quad (4)$$

where A is the asymptote of dominant height, a is a growth rate/shape parameter, t is the age and the subscripts 1 and 2 refer to age 1 and age 2 years.

Difference equations To apply the difference equation method of Clutter *et al.* (1983), the growth functions must be expressed in the form of difference equations. A difference form expresses height at remeasurement (h_{i2}) as a function of remeasurement age (t_2), initial or previous measurement age (t_1), and height at initial or previous measurement (h_{i1}):

$$h_{i2} = f(t_2, h_{i1}, t_1).$$

In order to obtain difference forms of the equations one of the parameters may be left free leaving two parameters to be statistically estimated. The difference forms of the Chapman–Richards (Mitscherlich, 1919) and Lundkvist–Korf (Korf, 1939; Lundqvist, 1957) growth equations were taken from Amaro *et al.* (1998).

Selection of a growth equation for dominant height modelling

For the selection of a growth function to be used in this study three steps were followed: (1) fitting the candidate growth equations; (2) cross-validation for model selection; (3) validation using stand data.

(1) Fitting all the candidate growth equations

with the data available from stem analysis of 87 trees (39 felled for this study, plus 48 felled by Luis (1989) in 1985/6). The candidate models were fitted by nonlinear regression using the PROC NLIN (SAS, 1996) function based on the Gauss–Newton procedure. The starting values have to be relatively close to the estimates or the procedure may result in convergence to a local minimum of the SS_{Res} function (Myers, 1990). In this case the starting values found by Amaro (1997) for *Eucalyptus* for the same functions were used. The criteria for comparing models were as follows.

(a) Adjusted coefficient of determination R_a^2 (Draper and Smith, 1998):

$$R_a^2 = 1 - \frac{\frac{RSS_{n-p}}{n-p}}{\frac{CTSS}{n-1}} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (5)$$

$$= 1 - (1 - R^2) \left(\frac{n-1}{n-p} \right)$$

where RSS_{n-p} is the residual sum of squares, n the number of observations, p is the total number of parameters in a fitted model and CTSS the corrected total sum of squares. The R_a^2 value is an attempt to correct the tendency for over-fitting of R^2 by adjusting both the numerator and the denominator by their respective degrees of freedom. The R_a^2 can decline in value if the contribution to the explained deviation by the additional variable is less than the impact on the degrees of freedom. Thus R_a^2 will react to alternative equations for the same dependent variable in a manner similar to the standard error of the estimate; i.e. the equation with the smallest standard error of the estimate will most likely also have the highest R_a^2 . While the R^2 is a percentage, the R_a^2 is not and should be referred to as an index value (Draper and Smith, 1998).

(b) Significance of the parameters. The final estimate of the parameters, and an ~95 per cent confidence band, should exclude zero, indicating that there are only non-zero values for the parameters and then they are always significant.

(c) Normality of the studentized residuals. A graph of studentized residuals plotted against their expected values under normality – the so

called QQ-plots – will display a straight line with a slope of 1 in the ideal situation, when the errors are normal. The residuals are ranked in increasing order $e_{[1]} < e_{[2]} < \dots < e_{[n]}$ and are plotted against the cumulative probability of a normal distribution $P_i = (i - 1/2)/n, i = 1, 2, \dots, n$ (Myers, 1990).

(d) Homogeneity of residuals. Plotting the studentized residuals against the fitted values \hat{y}_i highlighting a deviation from the homogeneous variance assumption.

(e) Statistics based on prediction sum of squares (PRESS) residuals. Considers a set of data in which the i th observation is set aside from the sample and the remaining $n - 1$ observations are used to estimate the coefficients for the candidate model. The procedure continues until all the observations have been removed, one at a time, and thus the candidate model is fitted n times. Each candidate model will have n PRESS residuals $y_i - \hat{y}_{i,-1} = e_{i,-1} (i = 1, 2, \dots, n)$ (Myers, 1990). Since $\hat{y}_{i,-1}$ is independent from y_i PRESS residuals can be considered prediction errors. PRESS residuals are indicated for examining the predictive abilities of the candidate models when sample sizes are small (Green, 1983). The mean of the PRESS residuals (MPRESS) is a measure of bias of a model, the ideal value for the MPRESS is zero while the mean of the absolute PRESS residuals (MAPRESS) is used to evaluate the precision of a model (Soares and Tomé, 2001).

(f) Analysis of the logical and biological consistency of the estimates of the parameters. The estimates have to be logically consistent and biologically realistic (Vanclay, 1994; Soares *et al.*, 1995). The growth equations with non-logical and non-biological consistency and poor statistical properties were eliminated.

(2) Following the adjustment of the candidate functions and elimination of growth equations with non-logical and non-biological consistency and poor statistical properties, the remaining functions were submitted to cross-validation. The PRESS residuals are to a certain extent prediction residuals and therefore they were used in the first stage for model selection. The use of an independent data set for cross-validation was thought to be useful for a second stage.

The data set of the stem analysis trees was split into two sub-samples: a fitting sample and a validation sample. First, the fitting group was

used for fitting the model and estimating the coefficients, while the validation group was used to estimate the responses of the model, through the calculation of prediction errors. After this, the validation group was used for fitting while the fitting group was used for estimating the prediction errors. These errors are then prediction errors and can be used for comparative purposes. The average of prediction errors was used to assess the bias of the models. To evaluate the precision of the models, the mean of the absolute value of the prediction errors was used (Myers, 1990). To aid a better understanding of the prediction capacity of the models, the prediction errors were analysed in detail for each age class and site index through box plot graphs.

Modelling efficiency (ME) was another statistical parameter analysed for different models and can be expressed as follows (Soares *et al.*, 1995):

$$ME = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (6)$$

ME provides an index of model performance on a relative scale, where 1 indicates a ‘perfect’ fit, 0 reveals that the model is no better than a simple average, and negative values indicate a poor model (Vanclay and Skovsgaard, 1997).

(3) Validation using all the data set from stem analysis for fitting equations and the stand data for calculation of the prediction errors of the best models in step (2). Selection of the growth equation for dominant height modelling. The stand data were obtained from field measurements and from existing stand data from previous studies (Loureiro, 1986; Luis, 1989; Luis and Monteiro, 1998).

Results

The final data set included 87 trees corresponding to 39 trees felled during field-work in 1999 and 48 trees felled between 1985 and 1986 by Luis (1989). Figure 2 shows their growth after applying Carmean’s (1972) height correction.

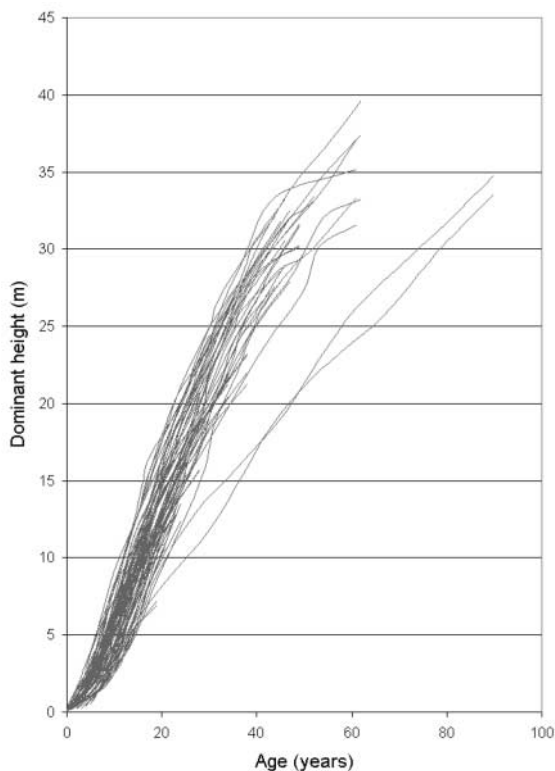


Figure 2. Height growth of 87 stem analysis trees after height correction using Carmean's method. The curves give the corrected heights in metres.

Fitting the candidate growth equations

The estimates of the parameters are summarized in Table 2. Lundqvist–Korf- n , Chapman–Richards- i and Lundqvist–Korf- k functions were rejected due to non-logical estimates of parameters. The Lundqvist–Korf- n and Chapman–Richards- i functions had estimates for the asymptotes (35.5 m and 36.5 m, respectively) smaller than the dominant height of trees from the largest stand (38 m, plot 1 at Vila Pouca). The estimate for the asymptote of the Lundqvist–Korf- k function was too large to be biologically realistic (350 m).

The non-linear least squares statistics were also used to evaluate the candidate functions (Table 3). The adjusted coefficients of correlation, R_a^2 , showed that all models have values above 93 per cent, with the exception of Lundqvist–Korf- n , with 84 per cent. All par-

ameters for all candidate functions were significantly different from zero (at 95 per cent confidence level).

The graphs of normality of the studentized residuals showed some deviations from the slope of one as, for example, where there is a skewed positive tendency which is often caused by model misspecification (Myers, 1990). However, small departures from normality do not greatly affect the model (Montgomery and Peck, 1982). In addition, no improvement was found by re-fitting the model using iteratively re-weighted least squares (Huber's M-estimator; Myers, 1990). Thus the estimates of the ordinary least squares were retained.

The graphs of studentized residuals against the fitted values were used to assess violations of the assumption of homogeneity of variance. Deviations were highlighted in graphs for

Table 2: Estimates of the parameters (notation of functions: i = integral form; k , A , n and m = when k , A , n and m are free parameters of difference equations)

Candidate function Estimates of the parameters

Lundqvist-Korf- i	$Y = 68.617e^{\frac{12.888}{t^{0.699}}}$
Lundqvist-Korf- k	$Y_2 = 349.760 \left(\frac{Y_1}{349.760} \right)^{\left(\frac{t_1}{t_2} \right)^{0.290}}$
Lundqvist-Korf- a	$Y_2 = Y_1 e^{7.862 \left(\frac{1}{t_1^{0.260}} - \frac{1}{t_2^{0.260}} \right)}$
Lundqvist-Korf- n	$Y_2 = \frac{35.503}{\sqrt[1nt_1]{\left(\frac{1}{0.001} \right)^{1n} \frac{t_2}{t_1} 1n \left(\frac{35.503}{Y_1} \right)^{1nt_2}}}$
Chapman-Richards- i	$Y = 36.503 (1 - e^{-0.048t})^{\frac{1}{1 - 0.498}}$
Chapman-Richards- k	$Y_2 = 43.526 \left\{ 1 - \left[1 - \left(\frac{Y_1}{43.526} \right)^{1 - 0.294} \right]^{\frac{t_2}{t_1}} \right\}^{\frac{1}{1 - 0.294}}$
Chapman-Richards- a	$Y_2 = Y_1 \left(\frac{1 - e^{-0.029t_2}}{1 - e^{-0.029t_1}} \right)^{\frac{1}{1 - 0.274}}$
Chapman-Richards- m	$Y_2 = 46.68^{\frac{1 - \frac{1n(1 - e^{-0.025t_2})}{1n(1 - e^{-0.025t_1})}}{1n(1 - e^{-0.025t_1})}} Y_1 \frac{1n(1 - e^{-0.025t_2})}{1n(1 - e^{-0.025t_1})}$
McDill-Amateis	$Y_2 = \frac{54.609}{1 - \left(1 - \frac{54.609}{Y_1} \right)^{\left(\frac{t_1}{t_2} \right)^{1.383}}}$

Table 3: Summary of the statistical analysis of the candidate functions using data from 87 stem analysis trees

Candidate function	R_a^2	Homogeneity of residuals*	MPRESS	MAPRESS
Lundqvist–Korf- <i>i</i>	0.933	3	0.126	1.757
Lundqvist–Korf- <i>k</i>	0.992	1	0.302	0.634
Lundqvist–Korf- <i>a</i>	0.993	1	0.237	0.569
Lundqvist–Korf- <i>n</i>	0.835	2	-0.269	2.939
Chapman–Richards- <i>i</i>	0.993	3	0.251	0.574
Chapman–Richards- <i>k</i>	0.949	1	0.059	1.490
Chapman–Richards- <i>a</i>	0.992	1	0.289	0.620
Chapman–Richards- <i>m</i>	0.993	2	0.244	0.574
McDill–Amateis	0.994	1	0.217	0.525

* 1 = almost homogeneous distribution of studentized residuals; 2 = near to homogeneous distribution of studentized residuals; 3 = non-homogeneous distribution of studentized residuals.

Table 4: Summary of the cross-validation statistics of the candidate functions

Candidate function	Sub-samples	Resp	Resp%	Aresp	Aresp%	ME
Lundqvist–Korf- <i>a</i>	1_2	0.35	7.67	0.64	11.28	0.99
	2_1	0.39	9.64	0.67	12.03	0.99
Chapman–Richards- <i>k</i>	1_2	0.26	7.23	0.58	11.08	0.99
	2_1	0.30	6.41	0.57	10.53	0.99
Chapman–Richards- <i>a</i>	1_2	0.32	8.38	0.62	11.78	0.99
	2_1	0.27	7.39	0.61	11.25	0.99
McDill–Amateis	1_2	0.27	7.30	0.58	11.12j	0.99
	2_1	0.24	6.51	0.56	10.60	0.99

Sub-samples 1_2 = using fitting sample for estimation of parameters and validation sample for calculation of errors; 2_1 = using validation sample for estimation of parameters and fitting sample for calculation of errors. Resp = mean prediction error (in m); Resp% = mean prediction error (in per cent); Aresp = mean absolute prediction error (in m); Aresp% = mean absolute prediction error (in per cent); ME = modelling efficiency.

Lundqvist–Korf-*i*, Lundqvist–Korf-*n*, Chapman–Richards-*i* and Chapman–Richards-*m* functions.

Fitting the candidate growth equations allowed the elimination of the Lundqvist–Korf-*n*, Chapman–Richards-*i* and Lundqvist–Korf-*k* functions due to non-biologically realistic estimates. In addition, Lundqvist–Korf-*i*, Lundqvist–Korf-*n*, Chapman–Richards-*i* and Chapman–Richards-*m* were eliminated due to failures in the assumption of homogeneity of variance.

Cross-validation for model selection

Following the adjustment of the candidate models the Lundqvist–Korf-*a*, McDill–Amateis, Chapman–Richards-*a* and Chapman–Richards-*k* models were retained for cross-validation. Data

for the set of 87 trees were split into two sub-samples of 45 trees and 42 trees, respectively. The results of cross-validation came up with two models with small prediction errors, Chapman–Richards-*k* and McDill–Amateis (Table 4).

The graphs of the prediction residuals against the fitted values showed reasonably homogeneous patterns of variance. However, the Lundqvist–Korf-*a* function showed some biased tendency for negative prediction errors for height estimations of ~0 m.

Validation using stand data

The two best models from the results of cross-validation were the McDill–Amateis and Chapman–Richards-*k* functions. These were

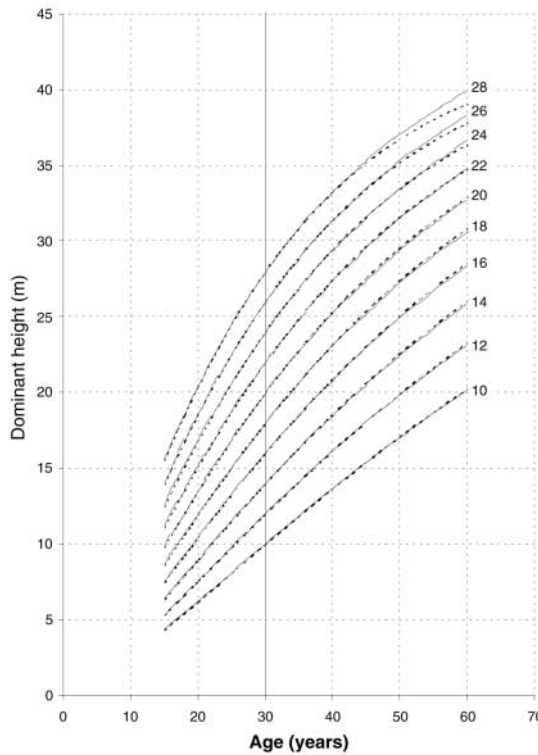


Figure 3. Site index curves for Douglas-fir in Portugal, base-age 30 years. Comparison of the two models with the best statistical properties: McDill–Amateis (continuous lines) and Chapman–Richards-*k* (broken lines) growth models.

Table 5: Summary of statistics when stand data was used for the calculation of the prediction errors

Candidate function	Resp	Resp%	Aresp	Aresp%	ME
Chapman–Richards- <i>k</i>	-0.22	-4.09	1.06	8.40	0.98
McDill–Amateis	-0.22	-4.04	1.04	8.32	0.98

Notations as in Table 4.

selected for validation using stand data. The results are similar for both models for characterization of prediction errors (Table 5) but the McDill–Amateis function had a slight superiority.

The two models were plotted in order to illustrate their differences (Figure 3). Both models are very similar except that on the best sites at older ages the McDill–Amateis function gives higher estimates for dominant height than Chapman–Richards-*k*.

Discussion

Validation using stand data in the third step led to very similar statistical characteristics of the McDill–Amateis and Chapman–Richards-*k* functions, differing mainly in the faster growth of higher productivity site classes at older ages in the McDill–Amateis model. They were compared with previous Douglas-fir site index curves, attempting to identify a general tendency of

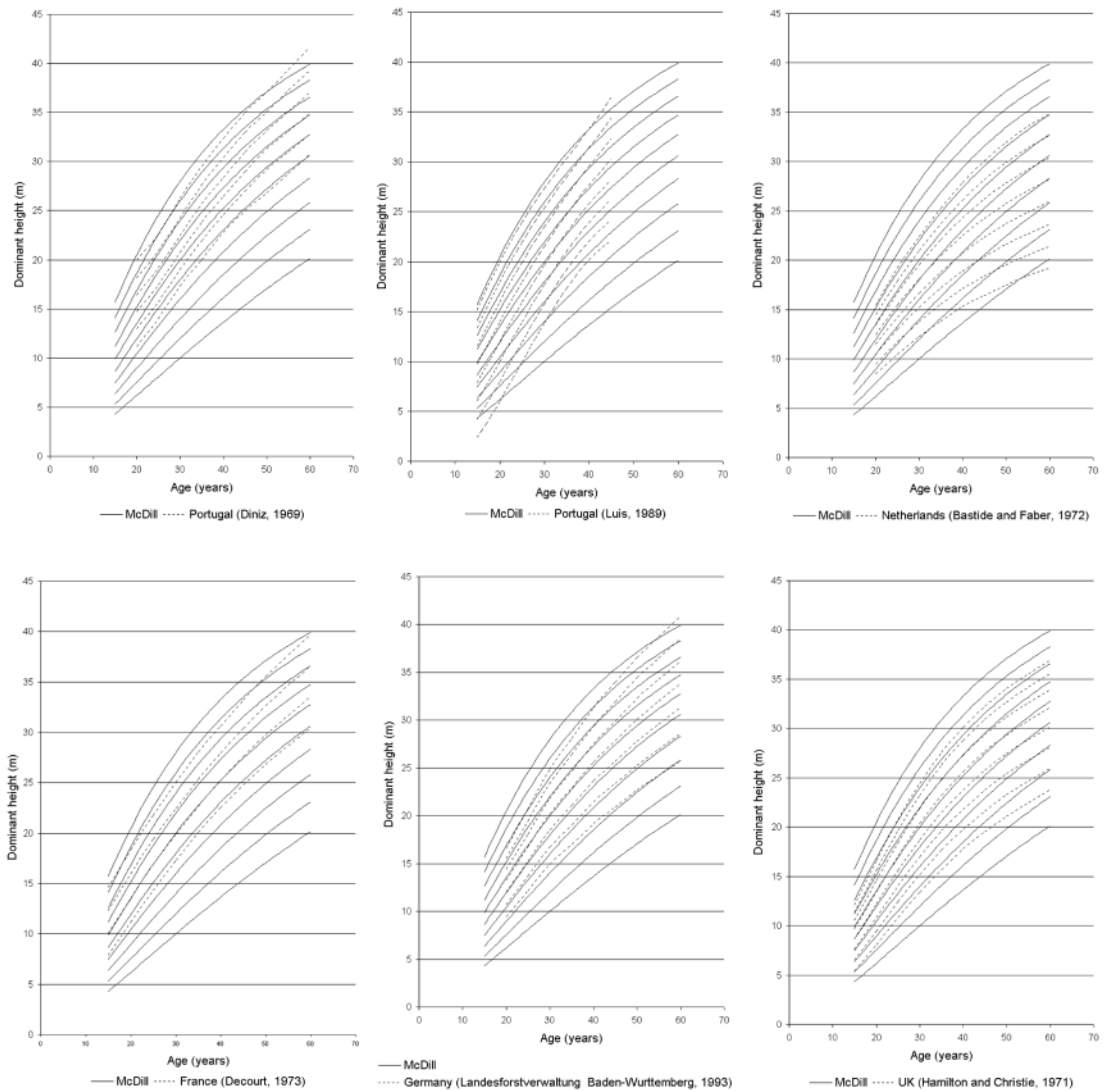


Figure 4. Comparison of site index curves derived from McDill–Amateis model and published site index curves for Douglas-fir in Portugal, the Netherlands, France, Germany and UK.

growth on the higher productivity site classes, so this difference could be used to indicate the better model.

Site index curves for Douglas-fir in Portugal (Diniz, 1969; Luis, 1989), the Netherlands (Bastide *et al.*, 1972), Italy (Maetzke and Nocentini, 1994), France (Decourt 1973), Canada (Mitchell and Cameron, 1985), USA (Curtis *et al.*, 1982), Germany (Landesforstverwaltung

Baden-Württemberg, 1993), Belgium (Rondeux and Thibaut, 1996) and Great Britain (Hamilton and Christie, 1971) were compared with the Chapman–Richards- k and McDill–Amateis models (Figure 4) derived here. These comparisons highlighted an overall tendency for lower estimations of dominant height for high productivity sites at older ages for both McDill–Amateis and Chapman–Richards- k models. This is even

clearer for the Chapman–Richards- k model, where only the curves for the Netherlands (Bastide *et al.*, 1972) do not show better performances for dominant height on high productivity sites at older ages. The tendency for lower estimates of dominant height on high productivity sites is related to a lower estimation of the asymptote. These low estimates are likely to be related to the absence of long-term data for high productivity sites. Estimates of the asymptote can vary considerably with different model formulations when long time-series of data are not available. There are inherent inaccuracies in attempting to project long-term growth using short-term data (McDill and Amateis, 1992). However, when information on long-term growth for high productivity sites becomes available it can be used to update the parameter estimates of the models without any fundamental model changes (McDill and Amateis, 1992). Since the McDill–Amateis model has higher estimates for high productivity sites at older ages, and all the other statistical parameters are similar to the Chapman–Richards- k model, it is proposed that it is used for modelling Douglas-fir dominant height growth in Portugal.

Figure 4 shows that the McDill–Amateis model has a faster growth rate in younger stands and slower growth in older ones when compared with previous Portuguese site index curves established with temporary sample plots (Diniz, 1969). A possible higher proportion of lower productivity temporary sample plots at younger ages and a higher proportion of temporary sample plots of high productivity at older ages might explain these differences. A comparison of the McDill–Amateis model with the Luis' (1989) site index curves showed a similar pattern for higher productivity site classes and a faster growth than the McDill–Amateis model for lower productivity site classes. The inclusion in this work of stands with lower productivity, such as Vila Flor, Penhas Douradas and Malcata, are likely to explain these differences.

The comparison of the McDill–Amateis model with site index curves from the US (Curtis *et al.*, 1982) and Canada (Mitchell and Cameron, 1985) show that these countries have faster growth on the higher productivity sites and lower growth on the lower productivity sites than in Portugal. Both these countries have extensive

areas of Douglas-fir corresponding to a wide range of sites where both higher and lower productivity sites, than the ones found in Portugal, might be expected. Comparing the McDill–Amateis model with Great Britain (Hamilton and Christie, 1971) and the Netherlands (Bastide *et al.*, 1972) site index curves there is a similar pattern for higher productivity site classes and slightly slower growth for Great Britain and the Netherlands on lower productivity site classes. French (Decourt, 1973) and German (Landesforstverwaltung Baden–Württemberg, 1993) site index curves have slightly higher growth compared with the McDill–Amateis model, which is more pronounced on higher productivity site classes. Italian (Maetzke and Nocentini, 1994) site index curves have considerably higher growth compared with the McDill–Amateis model.

Conclusions

The best growth function for modelling Douglas-fir dominant height in Portugal was the following McDill–Amateis function:

$$Y_2 = \frac{54.609}{1 - \left(1 - \frac{54.609}{Y_1}\right) \left(\frac{t_1}{t_2}\right)^{1.383}} \quad (7)$$

In its graphical form (Figure 3) it has practical relevance for the evaluation of site quality of Douglas-fir in Portuguese plantations. It produces estimates with a precision of ~0.6 m (mean of absolute prediction errors of cross-validation). The mean bias of its estimates is <0.3 m (mean prediction error from cross-validation).

The site index curves were constructed for an age range of 15–60 years, which includes the general rotations adopted for Douglas-fir (between 50 and 60 years). They are useful for quick qualitative assessments of Douglas-fir plantations. Extrapolations outside the range of ages of these site index curves are not recommended. A tendency for lower estimates of dominant height would be expected at high productivity site classes, which is likely to be related to the absence of long-term data for high productivity sites. If it is considered important to extend the range of ages for these site index curves,

information on long-term growth on high productivity sites will be required to update the parameter estimates of the model.

Acknowledgments

Many thanks to Eng. Maria Teresa Alves da Silva sub-director from Direcção Geral das Florestas for her encouragement and immense support, which made possible the completion of a large part of the field work. There are a great number of people from the following institutions whom we would like to thank: Direcção Geral das Florestas, CENACEF, COTF, and Serviços Florestais das Direcções Regionais da Agricultura of Entre-Douro e Minho, Trás-os-Montes e Alto Douro, Beira Interior and Beira Litoral. This research project was funded by Fundação para a Ciência e a Tecnologia grant PRAXIS XXI/BD/18258/98 to whom we are grateful.

References

- Amaro, A.P.N. 1997 *Modelação o crescimento de povoamentos de Eucalyptus globulus Labill de 1ª rotação em Portugal*. Instituto Superior Técnico, Universidade Técnica de Lisboa, Lisboa, 241 pp.
- Amaro, A., Reed, D., Tome, M. and Themido, I. 1998 Modeling dominant height growth: eucalyptus plantations in Portugal. *For. Sci.* **44**, 37–46.
- Avery, T.E. and Burkhart, H.E. 1994 *Forest Measurements*. McGraw-Hill, New York.
- Bastide, J.G.A.I., Faber, P.J. and La Bastide, J.G.A. 1972 Revised yield tables for six tree species in the Netherlands. *Uitvoerig Verslag, Stichting Bosbouwproefstation 'De Dorschkamp'* **11**, 1–64.
- Brister, G.H. and Schultz, R.C. *Height estimation in stem analysis*. The University of Georgia, School of Forest Resources (unpublished, quoted in Dyer and Bailey, 1987).
- Cailliez, F. and Alder, D. 1980 *Forest Volume Estimation and Yield Prediction*. Vol. 1. *Volume Estimation*. Vol. 2. *Yield Prediction*. Food and Agriculture Organization of the United Nations, Rome, 98 and 194 pp.
- Cao, Q.V. 1993 Estimating coefficients of base-age-invariant site index equations. *Can. J. For. Res.* **23**, 2343–2347.
- Carmean, W.H. 1972 Site index curves for upland oaks in the central states [of the USA]. *For. Sci.* **18**, 109–120.
- Clutter, J.L. 1983 *Timber Management: a Quantitative Approach*. Wiley, New York, Chichester.
- Costa, M.E.R.F.P. 1992 *Metodologias de Análise da Estrutura de Povoamentos Mistos*. Instituto Superior de Agronomia, Universidade Técnica de Lisboa, Lisboa.
- Curtis, R.O. 1964 A stem-analysis approach to site-index curves. *For. Sci.* **10**, 241–256.
- Curtis, R.O., Clendenen, G.W., Reukema, D.L. and DeMars, D.J. 1982 Yield tables for managed stands of coast Douglas-fir. *General Technical Report, Pacific Northwest Forest and Range Experiment Station, USDA Forest Service*, 182 pp.
- Decourt, N. 1973 Yield tables for Norway Spruce and Douglas Fir in the western Massif Central. *Revue Forestiere Francaise* **25**, 99–104.
- Diniz, D.M.A.M. 1969 *Estudo do crescimento da Pseudotsuga menziesii (Mirb.) Franco no Norte de Portugal*. Instituto Superior de Agronomia, Universidade Técnica de Lisboa, Lisboa, 70 pp.
- Draper, N.R. and Smith, H. 1998 *Applied Regression Analysis*. John Wiley, New York, Chichester.
- Dyer, M.E., and Bailey, R.L. 1987 A test of six methods for estimating true heights from stem analysis data. *For. Sci.* **33**, 3–13.
- Eichhorn, F. 1904 Beziehungen zwischen Bestandshöhe und Bestandsmasse. *Allgemeine Forst- und Jagdzeitung*, 45–49 (quoted in Savill *et al.*, 1997).
- Elfving, B. and Kiviste A. 1997 Construction of site index equations for *Pinus sylvestris* L. using permanent plot data in Sweden. *For. Ecol. Manage.* **98**, 125–134.
- Fabbio, G., Frattegiani, M. and Manetti, M.C. 1994 Height estimation in stem analysis using second differences. *For. Sci.* **40**, 329–340.
- Furnival, G.M., Gregoire, T.G. and Valentine, H.T. 1990 An analysis of three methods for fitting site-index curves. *For. Sci.* **36**, 464–469.
- Goelz, J.C.G. and Burk, T.E. 1992 Development of a well-behaved site index equation: jack pine in north central Ontario. *Can. J. For. Res.* **22**, 776–784.
- Graves, H.S. 1906 *Forest Mensuration*. John Wiley and Sons, New York (quoted in Dyer and Bailey, 1987).
- Green, E.J. 1983 Evaluating the predictive abilities of regressions with PRESS. *For. Sci.* **29**, 712–714.
- Hamilton, G.J. and Christie, J.M. 1971 Forest management tables (metric). *Forestry Commission Booklet* No. 34. HMSO, London.
- Korf, V. 1939 Prispěvek k matematické definici vzrostového zákona hmot lesních porostů. *Lesnická pracr*, 339–379.
- Landesforstverwaltung Baden-Württemberg 1993. *Hilfstabellen für die Forteinrichtung. Zusammenge stellt für den Gebrauch in der Landesforstverwaltung*. Herausgegeben vom Ministerium für Ländlichen Raum, Ernährung, Landwirtschaft und Forsten Baden-Württemberg, Stuttgart.
- Lenhart, J.D. 1972 An alternative procedure for improving height/age data from stem analysis. *For. Sci.* **18**, 332.

- Loureiro, A.C.M.M. 1986 *Estudos de crescimento e de produção de algumas espécies florestais com interesse para a região montana e subatlântica de Portugal*. Universidade de Trás-os-Montes e Alto Douro, Vila Real, Portugal.
- Luis, J.F.S. 1989 *Crescimento e competição em povoamentos de Pseudotsuga menziesii (Mirb.) Franco em Portugal*. Universidade de Trás-os-Montes e Alto Douro, Vila Real, 278 pp.
- Luis, J.F.S. and Monteiro, M.d.L. 1998 Dynamics of a broadleaved (*Castanea sativa*) conifer (*Pseudotsuga menziesii*) mixed stand in northern Portugal. *For. Ecol. Manage.* **107**, 183–190.
- Lundqvist, B. 1957 On the height growth in cultivated stands of pine and spruce in Northern Sweden. *Medd. fran Statens Skogforsk* **47**, 64 pp.
- McDill, M.E. and Amateis, R.L. 1992 Measuring forest site quality using the parameters of a dimensionally compatible height growth function. *For. Sci.* **38**, 409–429.
- Maetzke, F. and Nocentini, S. 1994 Dominant-height growth and site index curves for Douglas fir. *Italia Forest. Mont.* **49**, 582–594.
- Mitchell, K.J. and Cameron, J.R. 1985 Managed stand yield tables for coastal Douglas-fir: initial density and precommercial thinning. *Land Management Report*, Ministry of Forests, British Columbia, 69 pp.
- Mitscherlich, E.A. 1919 Problems of plant growth. *Landwirtschaftliche Jahrbucher* **53**, 167–182 (quoted in Vanclay, 1994).
- Mlodziansky, A.K. 1898 Measuring the forest crop. *USDA Division of Forestry Bulletin* No. 20 (quoted in Dyer and Bailey, 1987), 71 pp.
- Montgomery, D.C. and Peck, E.A. 1982 *Introduction to Linear Regression Analysis*. Wiley, New York.
- Myers, R.H. 1990 *Classical and Modern Regression with Applications*. Duxbury Press, Belmont, CA, USA.
- Newberry, J.D. 1991 A note on Carmean's estimate of height from stem analysis data. *For. Sci.* **37**, 368–369.
- Philip, M.S. 1994 *Measuring Trees and Forests*. CAB International, Wallingford.
- Rondeux, J. and Thibaut, A. 1996 *Yield Tables for Douglas fir (Pseudotsuga menziesii)*. Les Presses Agronomiques de Gembloux (ASBL), Gembloux.
- SAS 1996 *SAS Proprietary Software Release 6.12*. SAS Institute Inc., Cary, NC, USA.
- Savill, P., Evans, J., Auclair, D. and Falck, J. 1997 *Plantation Silviculture in Europe*. Oxford University Press, Oxford.
- Schumacher, F.X. 1939 A new growth curve and its application to timber-yield studies. *J. For.* **37**, 819–820.
- Soares, P. and Tomé, M. 2001. A tree crown ratio prediction equation for eucalypt plantations. *Ann. For. Sci.* **58**, 193–202.
- Soares, P., Tome, M., Skovsgaard, J.P. and Vanclay, J. 1995 Evaluating a growth model for forest management using continuous forest inventory data. Growth and Yield Estimation from Successive Forest Inventories. Selected papers from the IUFRO Conference, held in Copenhagen, 14–17 June 1993. *For. Ecol. Manage.* **1995**, **71**, 251–226.
- Spurr, S.H. 1952 *Forest Inventory*. Ronald Press, New York.
- Tennent, R.B. and Burkhart, H. E. 1981 *Selection of stem analysis trees for site index curve construction*. New Zealand Forest Service, Forest Research Institute, Rotorua (unpublished).
- Tomé, M.M.B.B.T. 1988 *Modelação do crescimento da árvore individual em povoamentos de Eucalyptus globulus Labill (1ª rotação)*. Instituto Superior de Agronomia, Universidade Técnica de Lisboa, Lisboa, 231 pp.
- Vanclay, J.K. 1994 *Modelling Forest Growth and Yield: Applications to Mixed Tropical Forests*. CAB International, Wallingford.
- Vanclay, J.K. and Skovsgaard, J.P. 1997 Evaluating forest growth models. *Ecol. Modelling* **98**, 1–12.

Received 5 May 2003