Original article

Site quality equations for *Pinus sylvestris* L. plantations in Galicia (northwestern Spain)

Ulises DIÉGUEZ-ARANDA*, Juan Gabriel ÁLVAREZ GONZÁLEZ, Marcos BARRIO ANTA,
Alberto ROJO ALBORECA

Departamento de Ingeniería Agroforestal, Universidad de Santiago de Compostela, Escuela Politécnica Superior,
Campus universitario, 27002 Lugo, Spain

Abstract – Difference equations derived on the basis of the Sloboda and McDill-Amateis differential functions, and from the integral form of the Bertalanffy-Richards, Korf and Hossfeld growth functions were used to model the dominant height growth of Scots pine (*Pinus sylvestris* L.) in Galicia (north-western Spain). Data from stem analysis and permanent sample plots were combined and used for fitting. Both numerical and graphical analyses were used to compare alternative models. The cross-validation approach was used to analyse 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, producing the most adequate site curves. It is therefore recommended for height growth prediction and site classification of Scots pine plantations in Galicia. 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.

height growth / site classification / algebraic difference equation / even-aged forest stand

Résumé – Équations prédictives de la fertilité des stations pour des plantations de *Pinus sylvestris* L. en Galice (nord-ouest de l'Espagne). À partir des équations différentielles de Sloboda et McDill-Amateis, ainsi que de la forme intégrale de Bertalanffy-Richards, Korf et Hossfeld, un modèle a été établi pour caractériser la croissance en hauteur dominante du pin sylvestre (*Pinus sylvestris* L.) en Galice (nord-est de l'Espagne). Des données, issues à la fois de l'analyse de tiges et de l'accroissement de placettes semi-permanentes, ont été combinées pour établir ce modèle. Des analyses numériques et graphiques ont été utilisées pour comparer les différents modèles existants. Les résidus de la validation croisée ont été utilisés pour évaluer le comportement des équations. L'équation en différences algébriques obtenue à partir de la fonction différentielle de McDill et Amateis donna le meilleur compromis entre les aspects biologique et statistique, fournissant les courbes de qualité les plus adéquates. C'est donc l'équation recommandée pour prédire la croissance en hauteur et pour réaliser les classifications de qualité des sites de plantation des pins sylvestres en Galice. Cette équation est invariante quant à l'âge de référence, de sorte que la hauteur à un âge donné peut être estimée à partir de la hauteur à un autre âge, sans compromettre la validité des prédictions.

croissance en hauteur / classement des sites / équation en différence algébrique / peuplement équienne

1. 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 which expresses this productivity is a required variable for the modelling of present and future growth and yield, and can also be used for forest land stratification for purposes of forest inventory, and for forest exploitation on a sustainable yield basis [24].

Conceptually, site quality is considered an inherent property of plots of land, whether or not trees are being grown at the time of interest. The productivity of specific stand can vary greatly due to a host of factors including the underlying soil conditions, climatic variables and management practices. For timber production purposes, and especially for even-aged forest stands, site quality is commonly expressed as a species-specific site index [14, 21]. Site index may be defined as height, at a predetermined age, of dominant or codominant trees that have always been dominant or codominant and healthy [29]. Empirical evidence from thinning experiments indicates that for many commercially important species, height growth is not greatly affected by the manipulation of stand density. However, the average height of the stand may be affected by thinning, depending on the method used, but within limits of stand density, height growth appears to be unaffected, particularly when the comparison is restricted to dominant and codominant trees [20].

---

* Corresponding author: udieguez@lugo.usc.es

Site quality associated with Scots pine (*Pinus sylvestris* L.) in Spain has been studied by several authors. The earliest site index curves were constructed using methods that could all be classified within the guide curve method [25–27, 41, 47, 51]. All of these studies together covered the three main areas of distribution of the species in Spain: the *Sistema Central* Mountains, the *Sistema Ibérico* Mountains and the Pyrenees. A further description of studies related to site quality of Scots pine up to 1996 is provided by Rojo and Montero [51]. More recently, Bravo and Montero [7] developed a system for site index estimation for this species in the High Ebro Basin (northern Spain), by considering soil attributes and using an extension of the Richards' model. Finally, Palahí et al. [44] developed a site quality system for Scots pine in the northeast of Spain using data from permanent plots and stem analysis. These authors tested eleven equations for modelling dominant height growth. Most of the equations were derived from functions frequently used in forest growth modelling, by means of the algebraic difference approach [5, 12, 20, 37] and the generalized algebraic difference approach [18]. Other recent studies related to growth and yield of Scots pine in Spain are from Palahí et al. [43], Palahí and Pukkala [42], Bravo and Montero [8], and Bravo and Díaz-Balteiro [6].

The objective of the present study was to develop a site index system for pure Scots pine plantations in Galicia (northwest of Spain).

2. MATERIALS AND METHODS

2.1. Data

The data used to develop the site index curves were obtained from two different sources. Initially, in the winter of 1996 and 1997 a network of 185 plots was established in pure Scots pine plantations. The plots were located throughout the area of distribution of this species in Galicia, and were subjectively selected to represent the existing range of ages, stand densities and sites. The plot size ranged from 625 m² to 1200 m², depending on stand density, in order to achieve a minimum of 60 trees per plot. We adopted this procedure because the plots were established for developing a whole stand model and an adequate number of trees is required to accurately estimate yield and growth. Two dominant trees were destructively sampled at 118 locations. These trees were selected as the first two dominant trees found outside the plots but in the same stands within ± 5% of the mean diameter at 1.3 m above ground level and mean height of the dominant trees (considered as the 100 largest-diameter trees per hectare). The trees were felled leaving stumps of average height 0.11 m; total bole length was measured to the nearest 0.01 m. The logs were cut at 2 to 2.5 m intervals for the first 4 to 5 m of bole length and at 1 m intervals thereafter. Number of rings was counted at each cross-sectioned point, and then converted to stump age, which can be considered equal to plantation age. As cross section lengths do not coincide with periodic height growth, it was necessary to adjust height/age data from stem analysis to account for this bias using Carmean's method [13], and the modification proposed by Newberry [40] for the topmost section of the tree. A test of six methods of estimating true heights from stem analysis data [22] showed that Carmean's algorithm provided the most accurate estimates. The data of these 236 stem analyses composed the first source of data.

A subset of 79 of the above-mentioned plots was re-measured in the winter of 2003. These plots were selected for developing a dynamic growth model for the species in the area of study, thus the initial number of plots installed in 1996–1997 was considerably reduced

Table I. Total tree height or plot dominant height statistics, given in meters (age class 15 = 10–19 years, etc.).

Age class	Number of obs.	Mean	Standard deviation	Minimum value	Maximum value
15	18	6.11	1.41	4.2	8.6
25	53	8.56	2.84	4.6	17.3
35	162	11.47	3.40	4.8	20.5
45	69	15.65	4.21	6.6	24.0
55	7	17.13	3.57	9.8	20.1
65	4	20.45	2.49	17.7	23.1
75	2	26.71	0.20	26.6	26.9

because the plots measured twice provided better information concerning the development of Scots pine stands. The dominant height of each of these plots was calculated as the mean height of the 100 thickest trees per hectare, both for the first and the second inventories. The data on age (excluding seedling age for plantation) and dominant height in these plots measured twice constituted the second source of data used in the study.

Summary statistics including number of observations, mean, standard deviation, minimum, and maximum values were calculated for the total tree height and plot dominant height variables grouped by age classes (Tab. I). All the data were converted to a two-year interval structure (i.e. heights for ages at 2, 4, 6, etc. years) using Carmean's algorithm (in the case of trees) or interpolating between the observed heights at the ages of measurement for the plots.

2.2. Methods for constructing a site index system

According to Clutter et al. [20], most techniques for site index curves construction can be viewed as special cases of three general methods: (1) the guide curve method, (2) the parameter prediction method, and (3) the difference equation method. Although the three methods are not mutually exclusive, the difference equation method has been the preferred form for developing site index curves [1, 2, 5, 11, 45, 48].

The difference equation method makes direct use of the fact that observations corresponding to a given plot or dominant tree should belong to the same site curve. 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 [45]. 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 behaviour of the model, which is capable of producing anamorphic or polymorphic (with single asymptote) curve families.

With the difference equation method (1) short observation periods of temporary plots or stem analysis from trees whose total age is under or over the reference age can be used, (2) the curves pass through site index at the reference age, and (3) they are base-age invariant [19, 20]. The invariant or unchanging property refers to predicted heights: 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. This includes forward and backward predictions, and the path invariance property that ensures the result of projecting first from A_1 to A_2, and then from A_2 to A_3, being the same as that of the one-step projection from A_1 to A_3. Equations derived using this technique define both height-growth and site index models as special cases of the same equation [16].

Table II. Algebraic difference models considered.

Model	Algebraic difference model	Base equation
M1	$$H_2 = b_0\left(\frac{H_1}{b_0}\right)^{e^{\left(\frac{b_1}{(b_2-1)A_2^{(b_2-1)}} - \frac{b_1}{(b_2-1)A_1^{(b_2-1)}}\right)}}$$	Sloboda [53] $$\frac{dH}{dA} = \ln\left(\frac{b_0}{H}\right)b_1\left(\frac{H}{A^{b_2}}\right)$$
M2	$$H_2 = \frac{b_0}{1 - \left(1 - \frac{b_0}{H_1}\right)\left(\frac{A_1}{A_2}\right)^{b_1}}$$	McDill and Amateis [37] $$\frac{dH}{dA} = \left(1 - \frac{H}{b_0}\right)b_1\frac{H}{A}$$
M3	$$H_2 = b_0\left(1 - \left(1 - \left(\frac{H_1}{b_0}\right)^{\frac{1}{b_2}}\right)^{\frac{A_2}{A_1}}\right)^{b_2}$$	Bertalanffy-Richards $$H = b_0(1 - e^{-b_1A})^{b_2}$$ solved by b_1
M4	$$H_2 = b_0\left(\frac{H_1}{b_0}\right)^{\frac{\ln\left(1 - e^{-b_1A_2}\right)}{\ln\left(1 - e^{-b_1A_1}\right)}}$$	Bertalanffy-Richards $$H = b_0(1 - e^{-b_1A})^{b_2}$$ solved by b_2
M5	$$H_2 = b_0\left(\frac{H_1}{b_0}\right)^{\left(\frac{A_1}{A_2}\right)^{b_2}}$$	Korf $$H = b_0e^{-b_1A^{-b_2}}$$ solved by b_1
M6	$$H_2 = b_0\exp\left(-b_1A_2^{\left(\ln\left(\frac{\ln(H_1/b_0)}{-b_1}\right)/\ln A_1\right)}\right)$$	Korf $$H = b_0e^{-b_1A^{-b_2}}$$ solved by b_2
M7	$$H_2 = \frac{H_1 + d + r}{2 + \frac{4b_3}{A_2^{b_2}(H_1 - d + r)}}$$ with $d = \frac{b_3}{Asi^{b_2}}$ and $r = \sqrt{(H_1 - d)^2 + 4b_3H_1A_1^{-b_2}}$	Hossfeld IV $$H = \frac{b_0}{1 + \frac{b_1}{A^{b_2}}}$$ solved by b_0 and assuming $b_1 = b_3/S$

H_1 and H_2 are dominant height (m) at ages A_1 and A_2 (years), respectively; Asi is an age between 5 and 50 years used to reduce the mean square error; ln is the natural logarithm; and b_0, b_1, b_2 and b_3 are parameters to be estimated.

2.3. 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.

There are many growth functions that can be employed in forestry, such as the 74 documented by Kiviste et al. [35]. The most important desirable attributes for site index equations are: (1) polymorphism, (2) sigmoid growth pattern with an inflexion point, (3) horizontal asymptote at old ages, (4) logical behaviour (height should be zero at age zero and equal to site index at the reference age), and (5) base-age invariance [2, 28, 45]. The fulfilment of these attributes depends on both the construction method and the mathematical function used to develop the curves, and it cannot always be achieved.

Multiple asymptotes (i.e. asymptote varying with site index, typically higher site indices having higher asymptotes) may also be a desirable attribute [15, 18], although some of the frequently used functions have a common asymptote. This does not appear to be of great importance, as the behaviour of the curves is generally suitable for the range of ages that would be used in practice, and the common asymptote is usually achieved at very old ages. Limits for usage must always be appended to the curves [28].

A total of seven models were selected for fitting the height/age relationship (Tab. II). The first six models were derived using the difference equation method. Models M1 and M2 were formulated on the basis of the differential equations proposed by Sloboda [53] and McDill and Amateis [37], respectively. Models M3 to M6 were formulated from the integral form of the Bertalanffy-Richards model [3, 4, 50]; and based on the integral form of Korf's model (cited in [36]).

Model M7 was proposed by Cieszewski and Bella [19] from the Hossfeld IV equation (cited in [46]) by relating a model parameter to site index. The approach used to derive model M7 can be referred to as initial-condition site index substitution in expanded dynamic equations [17].

These algebraic difference equations are base-age invariant, polymorphic, and model M7 has multiple asymptotes. All the models have been widely used to develop height/age curves [10, 19, 23, 24, 54].

2.4. Data structure and model fitting

The data structure used for fitting the seven models was arranged with all the possible combinations among height/age pairs for each tree and plot, including descending growth intervals. All possible growth intervals typically produce fitted models with a better predictive performance as compared to, for example, forward moving first differences [28, 32]. However, this data structure may lead to the rejection of the error assumptions because it automatically introduces a lack of independence among observations [28]. Although under noninde-pendence the parameter estimates are even unbiased, standard error estimators are biased [30].

The potential problem of lack of independence among observations and heteroscedasticity can be solved using generalised nonlinear least squares (GNLS) methods [28, 31, 38]. In this case, autocorrelation was modelled by expanding the error term in the following way [28, 29, 45]:

$$H_{ij} = f(H_j, A_i, A_j, \beta) + e_{ij} \text{ with } e_{ij} = \rho e_{i-1,j} + \gamma e_{i,j-1} + \varepsilon_{ij} \qquad (1)$$

where H_{ij} depicts prediction of height i by using H_j (height j), A_i (age i), and A_j (age $j \neq i$) as predictor variables; β is the vector of parameters to be estimated; e_{ij} is the corresponding error term; the ρ parameter accounts for the autocorrelation between the current residual and the residual from estimating H_{i-1} using H_j as a predictor; the γ parameter accounts for the autocorrelation between the current residual and the residual from estimating H_i using H_{j-1} as a predictor; and ε_{ij} are independent and identically distributed errors.

To avoid the possible problem of heteroscedasticity, the variance of errors was assumed to be a power function of the predicted dominant height [32, 33]. The weighting factors used were $w_i = \text{pred}.H_i^k$, where k is a constant. Since the predicted dominant heights are initially unknown, different values of k (e.g. $k = -0.4, -0.2, 0, 0.2, 0.4$) were tested until heteroscedasticity was corrected.

In using all possible differences, the number of observations is artificially inflated and the corresponding standard errors for the parameters are therefore too small. Thus, the standard errors were expanded by $\sqrt{n(\text{apd})/n(\text{fd})}$, where $n(\text{apd})$ is the number of observations using all possible differences and $n(\text{fd})$ is the number of observations if using only first differences [29].

Fitting was carried out by modelling the mean and the error structure simultaneously, using the SAS/ETS® MODEL procedure [52]. This method of proceeding simultaneously optimizes the regression of H on A and A on H, and avoids parameter bias due to independent estimation of (1) site index at base age given height at some other age, and (2) height at some desired age given height (site index) at base age [28, 29, 45].

For model M7, the parameter Asi was ranged from 5 to 50 years in order to reduce the mean square error [23, 54].

2.5. Model comparison and model selection

The comparison of the estimates of the eight models fitted for predicting dominant height over age was based on numerical and graphical analyses. Three statistical criteria obtained from the residuals were examined: root mean square error ($RMSE$), which analyses the accuracy of the estimates; the adjusted coefficient of determination (R^2_{adj}), which shows the proportion of the total variance that is explained by the model, adjusted for the number of model parameters and the

number of observations; and Akaike's information criterion differences ($AICd$), which is an index for selecting the best model on the basis of minimizing the Kullback-Liebler distance [9]. Their expressions may be summarized as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n - p}} \qquad (2)$$

$$R^2_{adj} = 1 - \frac{(n-1) \sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{(n-p) \sum_{i=1}^{n} (y_i - \bar{y})^2} \qquad (3)$$

$$AICd = n \log \hat{\sigma}^2 + 2k - \min (n \log \hat{\sigma}^2 + 2k) \qquad (4)$$

where y_i, $\hat{y}_i$ and $\bar{y}$ are the measured, predicted and average values of the dependent variable, respectively; n is the total number of observations used to fit the model; p is the number of model parameters; $k = p + 1$; and $\hat{\sigma}^2$ is the estimator of the error variance of the model obtained with the following equation:

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n}. \qquad (5)$$

The cross-validation of each model was based on the analysis of root mean square error of the estimates, the adjusted model efficiency (ME_{adj}, equivalent to the R^2_{adj} of the previous phase), and the Akaike's information criterion differences, obtaining the residual of each tree or plot by refitting the model without that tree or plot.

Apart from these three statistics, one of the most efficient ways of ascertaining the overall picture of model performance is by visual inspection, so 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 both for detection of possible systematic discrepancies and for selecting the weighting factor [39]. Additionally, graphs showing the appearance of the fitted curves overlaid on the trajectories of the stem analysis or plot data over time were examined.

Practical use of the models to estimate site quality from any given pair of height and age requires the selection of a base age to which site index will be referenced. Inversely, 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 according to the following considerations [28]: (1) the base age should be less than or equal to the youngest rotation age under typical management, (2) the base age should be close to the rotation age, and (3) the base age should be chosen so that it is a reliable predictor of height at other ages.

In order to address the third consideration, different base ages and their corresponding observed heights were used to estimate heights at other ages for each tree or plot. The results were compared with the values obtained from stem analyses and plot re-measurements, and the relative error in predictions ($RE\%$) was then calculated as follows:

$$RE\% = \frac{\sqrt{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2/(n-p)}}{\bar{y}} 100 \qquad (6)$$

where y_i, $\hat{y}_i$ and $\bar{y}$ are the measured, predicted and average values of the dependent variable, respectively; n is the total number of observations used to fit the model; and p is the number of model parameters.
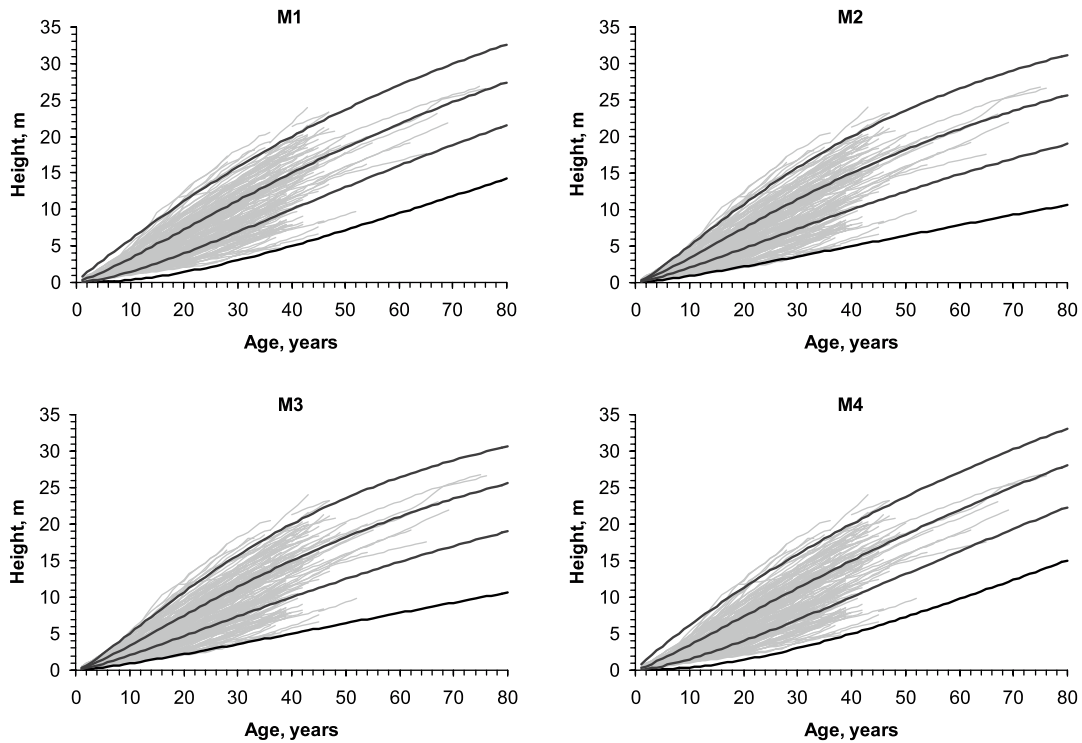
Figure 1. Plots showing the site curves for heights of 5, 10, 15 and 20 m at 40 years overlaid on the trajectories of observed values over time.

In addition to the graphs mentioned above, and because by definition site index is a fixed stand attribute that should be stable over time [32, 38], graphs showing the stability or the consistency of site index predictions over time were also constructed.

Finally, H_2 at A_2 was estimated considering previous H_1 at A_1 as predictors, and using different intervals of 2, 4, 6, etc. years, in order to find out for how long the estimations could be made from any given pair of height and age. Root mean square error over age was calculated for different age lags. For each of the different lags tested, a critical error ($E_{crit.}$) expressed as a percentage of the observed mean was also computed by re-arranging Freese's χ_n^2 statistic [49]:

$$E_{crit.} = \frac{\sqrt{\tau^2 \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 / \chi_{crit.}^2}}{\bar{y}} \qquad (7)$$

where n is the total number of observations in the data set, y_i is the observed value, $\hat{y}_i$ is its prediction from the fitted model, $\bar{y}$ is the average of the observed values, τ is a standard normal deviate at the specified probability level ($\tau = 1.96$ for $\alpha = 0.05$), and $\chi_{crit.}^2$ is obtained for $\alpha = 0.05$ and n degrees of freedom. If the specified allowable error, expressed as a percentage of the observed mean, is within the limit of the critical error, the χ_n^2 test will indicate that the model does not give satisfactory predictions; otherwise, it will indicate that the predictions are acceptable.

3. RESULTS AND DISCUSSION

The parameter estimates for each model and the statistics for both the fitting and the cross-validation phases are shown in Table III. All the parameters were found to be significant at a 5% level when the expansion factor proposed by [29] was applied.

In general, weightings factors of $w_i = 1/\text{pred}.H_i^{0.2}$ and $w_i = 1/\text{pred}.H_i^{0.4}$ showed the best results when plots of studentized residuals against the predicted heights were examined for detection of possible systematic trends of unequal error variance.

The values of the statistics used to compare the models indicate that all the models, except model M6 (Korf solved by b_2), produced a reasonable performance with low *RMSE* for fitting and cross-validation. These results suggest that, for this species, the indicated solution of Korf's model is not suitable for modelling the height/age relationship, as the b_2 parameter does not depend on site quality. Similar results were obtained by [1]. The best results for model M7 [19] were obtained when $Asi = 30$ was used. Models M1 (Sloboda [53]) and M4 (Bertalanffy-Richards solved by b_2) provided the best results for the goodness-of-fit statistics calculated, although models M2 (McDill and Amateis [37]) and M3 (Bertalanffy-Richards solved by b_1) represented the data almost as well, with both behaving similarly. Thereafter, the selection of the best model focused on these four models.

As previously commented, 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 5, 10, 15 and 20 m at 40 years overlaid on the trajectories of observed values over time (Fig. 1) were examined. They indicated that at young ages models M1 (Sloboda [53]) and M4 (Bertalanffy-Richards solved by b_2) overestimated

Table III. Parameter estimates and statistics for model comparison.

Model	Parameters		Approx. expanded std. error	Fitting			Cross-validation		
				RMSE	R^2_{adj}	AICd	RMSE	ME_{adj}	AICd
	b_0	67.90	2.53						
	b_1	0.1571	0.002						
	b_2	0.6179	0.009	0.484	0.9912	0	0.949	0.9662	0
	ρ	0.9276	0.013						
	γ	0.1358	0.011						
	b_0	51.39	0.97						
	b_1	1.277	0.004						
M2	ρ	0.9693	0.011	0.527	0.9896	11659	1.029	0.9603	11153
	γ	0.1032	0.010						
	b_0	37.90	0.61						
	b_2	1.294	0.005						
M3	ρ	0.9689	0.011	0.528	0.9895	12021	1.031	0.9601	11469
	γ	0.1035	0.010						
	b_0	64.22	1.91						
	b_1	0.008475	0.000						
M4	ρ	0.9410	0.012	0.501	0.9906	4798	0.984	0.9637	5012
	γ	0.1232	0.011						
	b_0	19438	4608						
	b_2	0.1323	0.004						
M5	ρ	0.9767	0.011	0.589	0.9870	27161	1.188	0.9470	30925
	γ	0.09856	0.010						
	b_0	158.1	11.6						
	b_1	7.034	0.057						
M6	ρ	1.041	0.008	1.125	0.9525	116067	2.242	0.8115	118258
	γ	0.04942	0.007						
	b_2	1.262	0.004						
	b_3	3116	55						
M7	Asi	30	–	0.542	0.9890	15757	1.068	0.9572	16281
	ρ	0.9707	0.011						
	γ	0.1030	0.010						

heights for the best sites and underestimated heights for the poorest sites. At older ages the curves generated from these models seem to increase quicker than the trajectories of trees and plots show.

In selecting the base age, it was found that a base age of 40 to 45 years was superior for predicting height at other ages with a minimum of reliability (Fig. 2). Even though at older ages the relative error in predicting height was lower, the scarcity of data
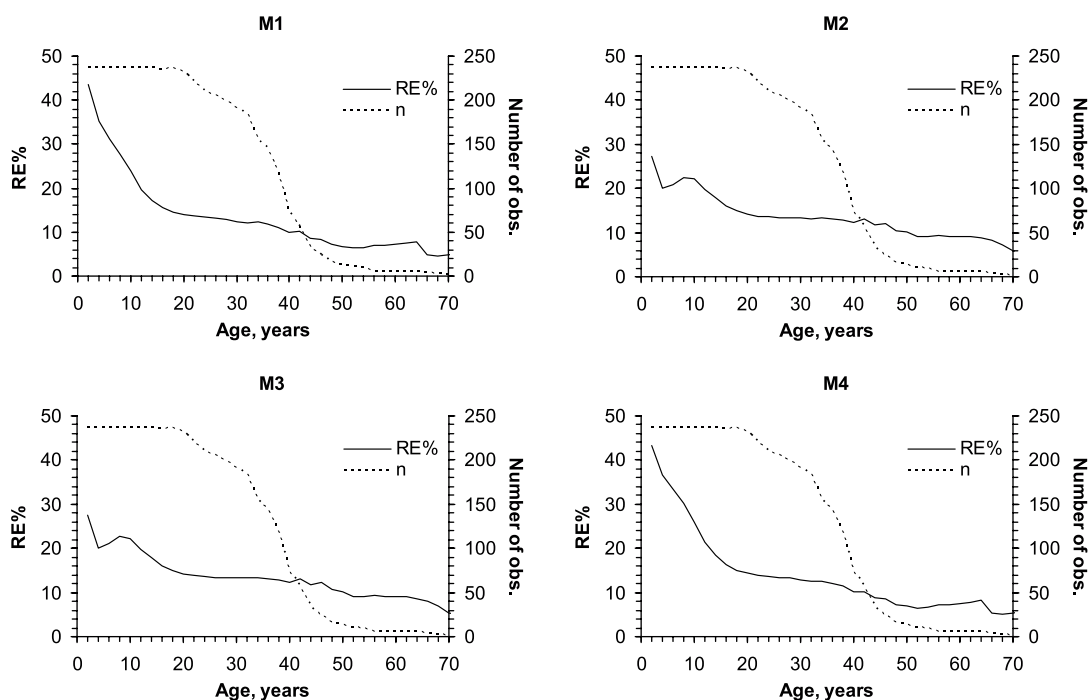
Figure 2. Relative error in height predictions related to choice of reference age. Ages older than 44 years are not representative enough due to the lack of data.

would lead to an incorrect decision as the data were not representative enough (less than 30 trees or plots). According to [28], this selection procedure should be devised so that variance of the volume estimates for the forest of interest are minimized, which requires that site index equations be integrated in a growth and yield system. Nevertheless, the lack of necessary information forced us to conclude that a reference age of 40 years is appropriate for Scots pine in Galicia.

The present investigation was based primarily on trees with a height at 40 years of between 6–22 m. Some of the trees were older than 55 years. Taking into account that most of the stands of the type covered in this study will be clear-felled at around 80–100 years, curves for heights up to 80 years were constructed. However, the use of the curves should always be approached with caution for ages above 60 years. Moreover, curves should be used for ages greater than 10 years, since for younger ages the erratic behaviour of the trees at initial ages may lead to erroneous classifications.

Finally, plots showing the stability of site index predictions over time (Fig. 3) show that models M2 (McDill and Amateis [37]) and M3 (Bertalanffy-Richards solved by b_1) provided the best results, judged by the consistency of predicted values over time. These two models performed adequately (taking into account numerical and graphical criteria) and both would be suitable for developing a site index system.

Summing up, model selection has been viewed as a compromise between biological and statistical considerations. Model M2 – the algebraic difference form of the differential function proposed by McDill and Amateis [37] – produced curves with an adequate graphical behaviour as well as good values of the goodness-of-fit statistics. Based on these consid-

erations, we propose its use for height growth prediction and site classification of Scots pine stands in Galicia:

$$H_2 = \frac{51.39}{1 - \left(1 - \frac{51.39}{H_1}\right)\left(\frac{A_1}{A_2}\right)^{1.277}} \qquad (8)$$

where H_1 and A_1 represent the predictor height (meters) and age (years), and H_2 is the predicted height at age A_2.

It should be noted that model M2 is parsimonious, as it includes only two parameters (excluding the correlation ones), and fits the data as well as, or better than, other models that were tried.

To use model M2 to estimate average stand height (H) for some desired age (A), given site index (S) and its associated base age (A_b), substitute S for H_1 and A_b for A_1 in equation (8):

$$H = \frac{51.39}{1 - \left(1 - \frac{51.39}{S}\right)\left(\frac{A_b}{A}\right)^{1.277}}. \qquad (9)$$

Similarly, to estimate site index at some chosen base age, given stand height and age, substitute S for H_2 and A_b for A_2 in equation (8):

$$S = \frac{51.39}{1 - \left(1 - \frac{51.39}{H}\right)\left(\frac{A}{A_b}\right)^{1.277}}. \qquad (10)$$

As regards how long the curves should be used for estimating height at any age given height at any other age, the plot of *RMSE*
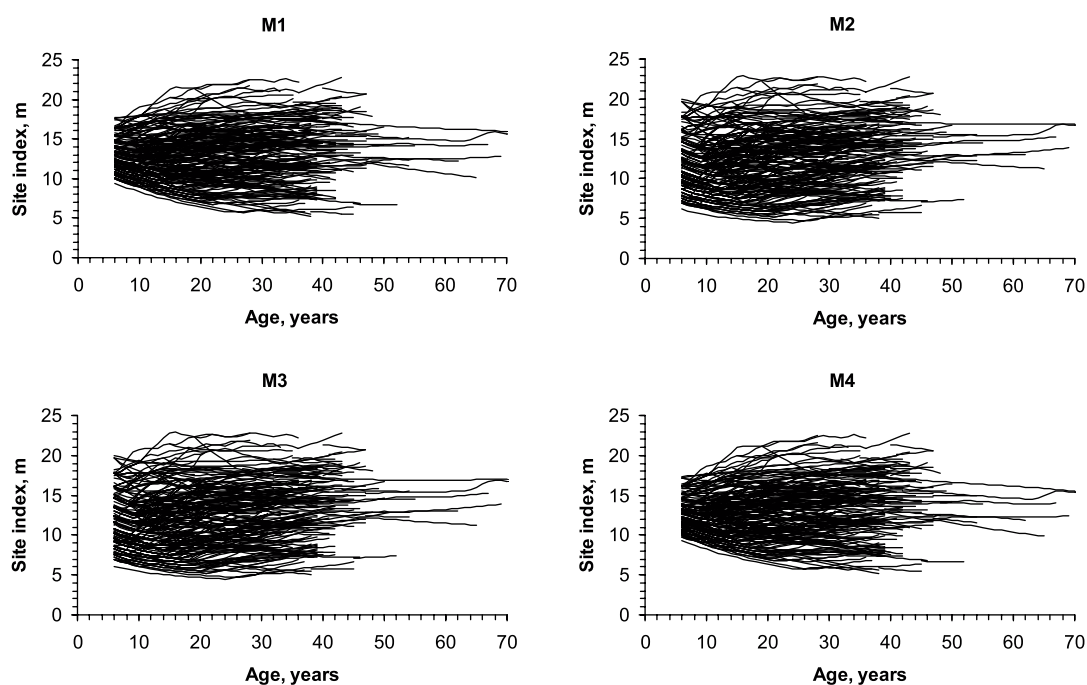
Figure 3. Site index predictions against total age using the stem analysis data and the data of the plots measured twice.
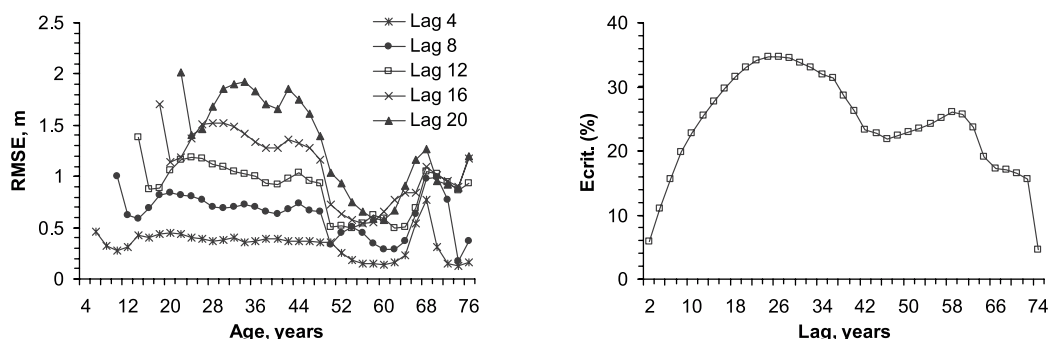

Figure 4. Plot of *RMSE* against age for different year lags (left) and plot of critical error against different year lags (right) for model M2.

against age for different year lags (Fig. 4 left) shows that as lag increases *RMSE* also increases for all ages, being more or less stable over time for the ages where enough data are available, and is, e.g. a maximum of 1 m for lag = 8 years and 1.8 m for lag = 20 years. From the plot showing the critical error against different year lags (Fig. 4 right), it can be observed that for lags of more than eight years the critical error exceeded 20%. In both cases, and for either ages or lags of more than 40–45 years, the calculated statistics are not reliable because of a lack of data. Considering the required accuracy in forestry growth modelling, where a mean prediction error of the observed mean at 95% confidence intervals within ± 10%–20% is generally realistic and reasonable as a limit for the actual choice of the acceptance and rejection levels [34], it can be stated that projections of dominant height should not be made directly for differences in

age of more than eight years; that is, for a given stand, projections of more than this length of time should be tested with real data as time passes, and these new data should be used to make new projections.

4. CONCLUSIONS

In practical terms, equation (8) is recommended for height growth prediction and site classification in the age interval 10–60 years in Scots pine plantations in Galicia, although the curves could also be used with caution for ages up to 70–80 years. It is also recommended that field studies of height growth in old pine plantations are continued.

Equation (8) can be solved for any base age, so estimates of Scots pine stand height and site index are easily obtained through direct evaluation of the function; there is no need for iterative numerical evaluation methods.

Visual inspection of the models is essential because numerical analyses may provide the best results for models that do not fit the data well enough.

Finally, it should be noted that projections of dominant height should not be made directly for differences in age of more than eight years, i.e. for a given stand, projections of more than this length of time should be tested with real data as time passes, and these new data should be used to make new projections.

REFERENCES

[1] Álvarez González J.G., Ruiz González A.D., Rodríguez Soalleiro R.J., Barrio Anta M., Development of ecoregion-based site index models for even-aged stands of *Pinus pinaster* Ait. in Galicia (northwestern Spain), Ann. For. Sci. 62 (2005) 117–129.

[2] Bailey R.L., Clutter J.L., Base-age invariant polymorphic site curves, For. Sci. 20 (1974) 155–159.

[3] Bertalanffy L.v., Problems of organic growth, Nature 163 (1949) 156–158.

[4] Bertalanffy L.v., Quantitative laws in metabolism and growth, Q. Rev. Biol. 32 (1957) 217–231.

[5] Borders T.H., Bailey R.L., Ware K.D., Slash pine site index from a polymorphic model by joining (splining) nonpolynomial segments with an algebraic difference method, For. Sci. 30 (1984) 411–423.

[6] Bravo F., Díaz-Balteiro L., Evaluation of new silvicultural alternatives for Scots pine stands in northern Spain, Ann. For. Sci. 61 (2004) 163–169.

[7] Bravo F., Montero G., Site index estimation in Scots pine (*Pinus sylvestris* L.) stands in the High Ebro Basin (northern Spain) using soil attributes, Forestry 74 (2001) 395–406.

[8] Bravo F., Montero G., High-grading effects on Scots pine volume and basal area in pure stands in northern Spain, Ann. For. Sci. 60 (2003) 11–18.

[9] Burnham K.P., Anderson D.R., Model selection and inference. A practical information-theoretic approach, Springer-Verlag, New York, 1998.

[10] Calama R., Cañadas N., Montero G., Inter-regional variability in site index models for even-aged stands of stone pine (*Pinus pinea* L.) in Spain, Ann. For. Sci. 60 (2003) 259–269.

[11] Cao Q.V., Estimating coefficients of base-age invariant site index equations, Can. J. For. Res. 23 (1993) 2343–2347.

[12] Cao Q.V., Baldwin V.C., Lohrey R.E., Site index curves for direct-seeded loblolly and longleaf pines in Lousinia, North J. Appl. For. 21 (1997) 134–138.

[13] Carmean W.H., Site index curves for upland oaks in the central states, For. Sci. 18 (1972) 109–120.

[14] Carmean W.H., Forest site quality evaluation in the United States, Adv. Agron. 27 (1975) 209–267.

[15] Cieszewski C.J., Comparing fixed-and variable-base-age site equations having single versus multiple asymptotes, For. Sci. 48 (2002) 7–23.

[16] Cieszewski C.J., Developing a well-behaved dynamic site equation using a modified Hossfeld IV function $Y^3 = (ax^m)/(c + x^{m-1})$, a simplified mixed-model and scant subalpine fir data, For. Sci. 49 (2003) 539–554.

[17] Cieszewski C.J., Bailey R.L., The method for deriving theory-based base-age invariant polymorphic site equations with variable asymptotes and other inventory projection models, University of Georgia, PMRC Technical Report 1999–4.

[18] Cieszewski C.J., Bailey R.L., Generalized algebraic difference approach: theory based derivation of dynamic equations with polymorphism and variable asymptotes, For. Sci. 46 (2000) 116–126.

[19] Cieszewski C.J., Bella I.E., Polymorphic height and site index curves for lodgepole pine in Alberta, Can. J. For. Res. 19 (1989) 1151–1160.

[20] Clutter J.L., Fortson J.C., Pienaar L.V., Brister G.H., Bailey R.L., Timber management: a quantitative approach, Krieger Publishing Company, New York, 1983, Reprint ed. 1992.

[21] Davis L.S., Johnson K.N., Bettinger P.S., Howard T.E., Forest management: to sustain ecological, economic, and social values, McGraw-Hill, New York, 2001.

[22] Dyer M.E., Bailey R.L., A test of six methods for estimating true heights from stem analysis data, For. Sci. 33 (1987) 3–13.

[23] Elfving B., Kiviste A., Construction of site index equations for *Pinus sylvestris* L. using permanent plot data in Sweden, For. Ecol. Manage. 98 (1997) 125–134.

[24] García O., A stochastic differential equation model for the height growth of forest stands, Biometrics 39 (1983) 1059–1072.

[25] García Abejón J.L., Tablas de producción de densidad variable para *Pinus sylvestris* en el Sistema Ibérico, Comunicaciones INIA Serie: Recursos Naturales, nº 10, 1981.

[26] García Abejón J.L., Gómez Loranca J.A., Tablas de producción de densidad variable para *Pinus sylvestris* en el sistema Central, Comunicaciones INIA Serie: Recursos Naturales, nº 29, 1984.

[27] García Abejón J.L., Tella G., Tablas de producción de densidad variable para *Pinus sylvestris* L. en el Sistema Pirenaico, Comunicaciones INIA Serie: Recursos Naturales, nº 43, 1986.

[28] Goelz J.C.G., Burk T.E., Development of a well-behaved site index equation: jack pine in north central Ontario, Can. J. For. Res. 22 (1992) 776–784.

[29] Goelz J.C.G., Burk T.E., Measurement error causes bias in site index equations, Can. J. For. Res. 26 (1996) 1586–1593.

[30] Gregoire T.G., Schabenberger O., Linear modelling of irregularly spaced, unbalanced, longitudinal data from permanent-plot measurements, Can. J. For. Res. 25 (1995) 137–156.

[31] Huang S., Development of a subregion-based compatible height-site index-age model for black spruce in Alberta, Alberta Land and Forest Service, For. Manag. Res. Note No. 5, Pub. No. T/352, Edmonton, Alberta, 1997.

[32] Huang S., Development of compatible height and site index models for young and mature stands within an ecosystem-based management framework, in: Amaro A., Tomé T. (Eds.), Empirical and process based models for forest tree and stand growth simulation, Edicões Salamandra, 1999, pp. 61–98.

[33] Huang S., Price D., Titus S.J., Development of ecoregion-based height-diameter models for white spruce in boreal forests, For. Ecol. Manage. 129 (2000) 125–141.

[34] Huang S., Yang Y., Wang Y., A critical look at procedures for validating growth and yield models, in: Amaro A., Reed D., Soares P. (Eds.),

Modelling forest systems, CAB International, Wallingford, Oxfordshire, UK, 2003, pp. 271–293.

[35] Kiviste A.K., Álvarez González J.G., Rojo A., Ruiz A.D., Funciones de crecimiento de aplicación en el ámbito forestal, Monografía INIA: Forestal nº 4, Ministerio de Ciencia y Tecnología, Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, Madrid, 2002.

[36] Lundqvist B., On height growth in cultivated stands of pine and spruce in Northern Sweden, Medd. Fran Statens Skogforsk. 47 (1957) 1–64.

[37] McDill M.E., Amateis R.L., Measuring forest site quality using the parameters of a dimensionally compatible height growth function, For. Sci. 38 (1992) 409–429.

[38] Monserud R.A., Height growth and site index curves for inland Douglas-fir based on stem analysis data and forest habitat type, For. Sci. 30 (1984) 943–965.

[39] Neter J., Kutner M.H., Nachtsheim C.J., Wasserman W., Applied linear statistical models, 4th ed., McGraw-Hill, New York, 1996.

[40] Newberry J.D., A note on Carmean's estimate of height from stem analysis data, For. Sci. 37 (1991) 368–369.

[41] Ortega A., Modelos de evolución de masas de *Pinus sylvestris* L., Tesis doctoral, Escuela Técnica Superior de Ingenieros de Montes, Universidad Politécnica de Madrid, 1989 (unpublished).

[42] Palahí M., Pukkala T., Optimising the management of Scots pine (*Pinus sylvestris* L.) stands in Spain based on individual-tree models, Ann. For. Sci. 60 (2003) 105–114.

[43] Palahí M., Pukkala T., Miina J., Montero G., Individual-tree growth and mortality models for Scots pine (*Pinus sylvestris* L.) in north-east Spain, Ann. For. Sci. 60 (2003) 1–10.

[44] Palahí M., Tomé M., Pukkala T., Trasobares A., Montero G., Site index model for *Pinus sylvestris* in north-east Spain, For. Ecol. Manage. 187 (2004) 35–47.

[45] Parresol B.R., Vissage J.S., White pine site index for the southern forest survey, USDA For. Serv. Res. Pap. SRS-10, 1998.

[46] Peschel W., Die mathematischen Methoden zur Herteitung der Wachstums-gesetze von Baum und Bestand und die Ergebnisse ihrer Anwendung, Tharandter Forstliches Jarbuch 89 (1938) 169–274.

[47] Pita P.A., La calidad de la estación en las masas de *Pinus sylvestris* de la Península Ibérica, Anales del Instituto Forestal de Investigaciones y Experiencias 9 (1964) 5–28.

[48] Ramírez-Maldonado H., Bailey R.L., Borders B.E., Some implications of the algebraic difference approach for developing growth models, in: Ek A.R., Shifley S., Burk T.E. (Eds.), Proceedings of IUFRO conference on forest growth modelling and prediction, USDA For. Serv. Gen. Tech. Rep. NC-120, 1988, pp. 731–738.

[49] Reynolds M.R. Jr., Estimating the error in model predictions, For. Sci. 30 (1984) 454–469.

[50] Richards F.J., A flexible growth function for empirical use, J. Exp. Bot. 10 (1959) 290–300.

[51] Rojo A., Montero G., El pino silvestre en la Sierra de Guadarrama, Ministerio de Agricultura, Pesca y Alimentación, Madrid, 1996.

[52] SAS Institute Inc., SAS/ETS User's Guide, Version 8, Cary, North Carolina, 2000.

[53] Sloboda B., Zur Darstellung von Wachstumprozessen mit Hilfe von Differentialgleichungen erster Ordung. Mitteillungen der Badenwürttem-bergischen Forstlichen Versuchs und Forschungsanstalt, 1971.

[54] Trincado G., Kiviste A.K., Gadow K.v., Preliminary site index models for native roble (*Nothofagus oblicua*) and raulí (*N. alpina*) in Chile, N.Z. J. For. Sci. 32 (2002) 322–333.