



**HAL**  
open science

## A stochastic height-diameter model for maritime pine ecoregions in Galicia (northwestern Spain)

Fernando Castedo Dorado, Marcos Barrio Anta, Bernard R. Parresol, Juan Gabriel Álvarez González

► **To cite this version:**

Fernando Castedo Dorado, Marcos Barrio Anta, Bernard R. Parresol, Juan Gabriel Álvarez González. A stochastic height-diameter model for maritime pine ecoregions in Galicia (northwestern Spain). *Annals of Forest Science*, 2005, 62 (5), pp.455-465. hal-00883904

**HAL Id: hal-00883904**

**<https://hal.science/hal-00883904>**

Submitted on 11 May 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## A stochastic height-diameter model for maritime pine ecoregions in Galicia (northwestern Spain)

Fernando CASTEDO DORADO<sup>a\*</sup>, Marcos BARRIO ANTA<sup>b</sup>, Bernard R. PARRESOL<sup>c</sup>,  
Juan Gabriel ÁLVAREZ GONZÁLEZ<sup>b</sup>

<sup>a</sup> Departamento de Ingeniería Agraria, Universidad de León, Escuela Superior y Técnica de Ingeniería Agraria, Campus de Ponferrada, 24400 Ponferrada (León), Spain

<sup>b</sup> Departamento de Ingeniería Agroforestal, Universidad de Santiago de Compostela, Escuela Politécnica Superior, Campus universitario, 27002 Lugo, Spain

<sup>c</sup> USDA Forest Service, Southern Research Station, PO Box 2680, Asheville, North Carolina 28802, USA

(Received 15 November 2004; accepted 25 March 2005)

**Abstract** – A stochastic height-diameter model was developed for maritime pine (*Pinus pinaster* Ait.) in Galicia (northwestern Spain). Four well-known growth functions were initially considered in this work, however, only Schnute's function performed adequately. A set of 20 695 pairs of height-diameter measures, collected in thinned and unthinned pure and even-aged stands, were used to fit the model. These stands were located throughout Galicia Autonomous Region covering a wide range of forest stands and site conditions. Since unequal error variance occurs, the generalized non-linear least squares method was used to take into account the error structure. Different weighting factors were employed to remove the heterogeneous variance of the errors. Because the local model (including only tree dimensions as explanatory variables) did not provide adequate results, stand variables were tested and incorporated into the height-diameter model. Ecoregion differences in the height-diameter relationship were analysed using the non-linear extra sum of squares method and the Lakkis-Jones test. Both tests showed that model parameters were significantly different between the two ecoregions normally defined for this species: coast and interior. The effect of thinning was examined; however, no benefits were obtained by introducing an additional thinning response variable in the prediction model. Finally, since trees with the same diameter usually do not have the same height, even within the same stand, a stochastic component was added to the deterministic height function. This approach mimics the natural variability of heights and therefore provides more realistic height predictions.

**generalized height-diameter model / stochastic component / ecoregions / thinning effect / *Pinus pinaster***

**Résumé** – Un modèle stochastique de hauteur-diamètre pour le pin maritime en Galice (nord-ouest de l'Espagne). Un modèle stochastique hauteur-diamètre de pin maritime a été développé en Galice (nord-ouest de l'Espagne). Quatre fonctions bien connues de croissance ont été initialement prises en compte dans cette étude, et cependant seule la fonction de Schnute s'est comportée correctement. Un échantillon de 20 695 couples de mesures hauteur-diamètre recueilli sur des peuplements clairsemés et des peuplements denses y compris sur d'anciens peuplements, a été utilisé pour établir ce modèle. Ces peuplements étaient situés dans la région autonome de Galice, et constituent un large éventail de peuplements forestiers et de conditions de terrain. Comme une divergence de variabilité est apparue, la méthode généralisée des moindres carrés non-linéaires a été utilisée pour prendre en compte la structure des erreurs. Différentes pondérations ont été utilisées pour enlever l'hétérogénéité des erreurs. Puisque le modèle local (comprenant seulement trois variables explicatives) n'a pas donné de résultats adéquats, des variables de peuplement ont été incluses dans le modèle. Les différences relatives à l'écorégion dans la relation hauteur-diamètre ont été analysées en utilisant la méthode des moindres carrés non-linéaires et le test de Lakkis-Jones. Les deux tests ont montré que les paramètres des modèles étaient significativement différents d'une région à l'autre, régions normalement naturelles pour ces essences : la côte et l'intérieur des terres. L'effet des éclaircies sélectives a été étudié ; cependant aucune amélioration ne fut obtenue en introduisant une caractéristique d'éclaircie dans la prévision de la hauteur des modèles. Finalement, comme des arbres de même hauteur ne présentent pas nécessairement un diamètre similaire, y compris ceux d'un peuplement donné, un élément aléatoire a été ajouté à la fonction de prédiction de la hauteur. Cette approche permet de prendre en compte la variabilité naturelle des hauteurs et par conséquent permet d'obtenir des prédictions plus fiables.

**modèle généralisé hauteur diamètre / composant stochastique / écorégions / effet de l'éclaircie / *Pinus pinaster***

### 1. INTRODUCTION

Maritime pine is the most important coniferous species of northwestern Spain, where 620 000 ha of pure or mixed stands

are present, derived both from plantations or natural regeneration after clear-cutting or fire. Its wide distribution and the variety of sites occupied have made *Pinus pinaster* a species of high relevance in Galician forestry with more than 50 million cubic

\* Corresponding author: diafcd@unileon.es

meters of standing timber [53]. The economic relevance of the species is also very high, with an annual cut volume of 2 380 000 m<sup>3</sup> in the period 1992–2001 [54].

Nowadays, maritime pine populations from Galicia show high levels of genetic diversity due to the use of seed lots from different origins. This lack of genetic homogeneity joined to an important genotype-by-environmental interaction favours the existence of adaptations to local ecological conditions [1, 2]. For these reasons differences in the height growth pattern among ecoregions were found [3].

To address this aspect it is necessary to adopt the principles of ecologically based forest management. Since the height-diameter relationship depends heavily on the local environmental conditions and varies within a geographic region [38], to account for the effects of climatic and ecological factors, the development of the height-diameter model should be based on the ecoregion classification system developed by Vega et al. [51] for this species in Galicia. This system differentiates interior and coastal ecoregions based on both environmental conditions and seed origin.

Individual tree height and diameter are essential forest inventory measures for estimating timber volume, site index, and other important variables in forest growth and yield, succession, and carbon budget models [37]. Devices which use ultrasound or laser pulses to measure distances have reduced the time needed to measure tree heights ( $h$ ), but measuring heights still requires more time than measuring the diameter at breast height ( $d$ ). For this reason, often only a subset of trees with measured diameters is also measured for height. Accurate height-diameter equations must be used to predict heights for the remaining trees, reducing data acquisition cost. Also, accurate height-diameter functions are basic for the estimation of stand development over time in growth and yield models (e.g., [9, 12]). Since this relationship is highly dependent on the stand conditions, the local height-diameter curves do not adapt well to all the possible situations that can be found within a forest, so a different diameter-height regression may be required for each stand [11, 31, 44, 61]. Normally, this relationship can be improved by taking into account stand variables that introduce into the model the dynamics of each stand [11, 27, 31, 45].

Due to the high variability in *Pinus pinaster* stands in Galicia, the objective of this study was to develop a generalized height-diameter model for both ecoregions (coast and interior) and to include a stochastic component to mimic the observed natural variability of heights. This model will be used as a component of a dynamic stand-level growth model for even-aged stands of maritime pine in this region.

## 2. MATERIALS AND METHODS

### 2.1. Data

We used three different inventories collected in even-aged stands throughout Galicia to develop the height-diameter model. The first data set is based on 249 samples in temporary and permanent rectangular plots. These data were derived from a thinning experiment established from 1965 to 1972 by the *Instituto Forestal de Investigaciones y Experiencias* (IFIE). The second group, a total of 188 plots, was sampled by the Lourizán Research Center with the objective of quantifying

the site quality and the effect of fertilization in stands of this species [4]. The third data set was obtained from 33 plots corresponding to a thinning trial experiment and temporary plots established by the *Escuela Politécnica Superior* in Lugo (University of Santiago de Compostela) in 2003. A total of 20 695 pairs of height-diameter measurements in both ecoregions were used in this study. These data sets cover a wide range of stand conditions in both ecoregions (coast and interior). In the first two data sources, most of the plots were measured from one to four times. Summary statistics, including the mean, minimum, maximum, and standard deviation of the main tree and stand variables are given in Tables I and II, respectively.

### 2.2. Methodology

#### 2.2.1. Candidate models

A large number of both local and generalized height-diameter equations have been reported in the forestry literature (e.g., [11, 14, 19, 21, 31, 36, 47, 50]). From a biological point of view, a curve of height growth does exhibit a sigmoidal or S-shaped pattern [10]. Thus, the selection of a functional form for the height-diameter relationship should not be restricted to the ease-of-fit to the data, but also should consider characteristics of the chosen models such as monotonic increment, functional inflection point and asymptotic value [30]. Sigmoid or S-shaped functions are preferred because they have these three properties; however, convex-shaped curves do not have inflection points. The number of parameters (flexibility), possible biological interpretation of the parameters (e.g., upper asymptote, maximum or minimum growth rate), and satisfactory predictions for height-diameter relationships are also important features [37].

Taking into account all these considerations, we considered four non-linear growth functions that have been frequently used (see [36, 37, 59]) for examination in this study: Bertalanffy-Richards [41]; Weibull [55], Korf [57] and Schnute [43]. The Bertalanffy-Richards and the Schnute models are probably the most flexible and versatile functions available for modelling height-diameter relationships [30]. The Bertalanffy-Richards function has been extensively used in describing height-diameter relationships (e.g., [14, 19, 36, 59]). The Schnute model has the advantage that it is easy to fit and quick to achieve convergence for any database [6, 29, 30]. This is particularly true in preliminary analysis for our database because non-convergence of parameter estimates for the first three equations was obtained. Thus, only the Schnute model (1) was considered for further study:

$$h = \left[ H_1^{\beta_0} + \left( H_2^{\beta_0} - H_1^{\beta_0} \right) \frac{1 - e^{-\beta_1(d-d_1)}}{1 - e^{-\beta_1(d_2-d_1)}} \right]^{(1/\beta_0)} + \varepsilon \quad (1)$$

where:

$h$  = total height of the tree;

$d$  = tree diameter at breast height;

$d_1$  = diameter at breast height of a small tree (lower range of data);

$d_2$  = diameter at breast height of a big tree (upper range of data);

$H_1$  = parameter representing mean tree height at  $d_1$ ;

$H_2$  = parameter representing mean tree height at  $d_2$ ;

$\beta_0$  = incremental acceleration in growth rate;

$\beta_1$  = constant acceleration in growth rate;

$\varepsilon$  = residual error.

**Table I.** Characteristics of the tree samples used for model fitting.

Sample of trees for model fitting ( $n = 20\ 695$ )				
Variable	Mean	Minimum	Maximum	Std. Dev.
$d$ (cm)	10.11	2.20	49.90	7.15
$h$ (m)	10.08	2.10	27.80	3.50

$d$  = diameter at breast height and  $h$  = total tree height.

**Table II.** Characteristics of the sample plots.

Sample of plots for model fitting ( $n = 493$ )				
Variable	Mean	Minimum	Maximum	Std. Dev.
$A$ (years)	19.24	8.00	50.00	7.35
$N$ (trees ha <sup>-1</sup> )	1666.61	275.00	4642.00	800.58
$\bar{d}$ (cm)	15.95	4.82	35.95	6.09
$d_g$ (cm)	16.40	5.06	36.53	6.21
$D_0$ (cm)	22.75	8.10	44.47	7.36
$G$ (m <sup>2</sup> ha <sup>-1</sup> )	29.95	3.17	72.48	12.11
$H_0$ (m)	11.56	3.71	24.60	4.48
$\bar{H}$ (m)	10.38	3.60	23.50	3.63
$S$ (m)	12.33	6.10	19.10	2.67

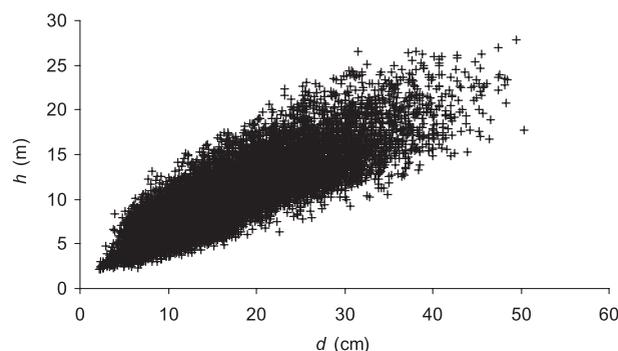
$A$  = stand age,  $N$  = number of trees per hectare,  $\bar{d}$  = mean diameter,  $d_g$  = quadratic mean diameter,  $D_0$ ,  $H_0$  = dominant diameter and dominant height respectively (using Assmann's criterion for both),  $G$  = basal area,  $\bar{H}$  = average height, and  $S$  = site index, defined as the dominant height (expressed in meters) that a stand reaches at 20 years, and determined from the site index curves available for both ecoregions [3].

In the context of height-diameter modelling, it is a common practice to force the curve to pass through the point (0, 1.3) to prevent negative estimates for small trees. Although, in reality, when diameter at breast height is zero height can take any value between 0 and 1.3, in this case the factors that control the height are independent and not part of the height-diameter relationship. Taking into account these considerations, we let  $d_1 = 0$  and  $H_1 = 1.3$ . This results in the modified Schnute model:

$$h = \left[ 1.3\beta_0 + \left( H_2\beta_0 - 1.3\beta_0 \right) \frac{1 - e^{-\beta_1 d}}{1 - e^{-\beta_1 d_2}} \right]^{(1/\beta_0)} + \varepsilon. \quad (2)$$

The relationship between diameter and height is highly influenced by stand variables; thus, some of these variables should be included into the model. There are two main approaches to incorporate the stand variables into a model [20, 49]. The first one is the parameter prediction approach [10], also known as the two-stage approach [15]. In this case, the height-diameter relationship is fitted individually for each sample plot; then in a second stage, parameters are explained using stand variables such as number of trees per hectare, basal area, dominant height, etc., as explanatory variables. The second approach is to add the stand variables directly into the model.

Because the first approach has biological relevance, leading to easier model interpretation [49], it was selected. Thus, the parameters  $\beta_0$ ,  $\beta_1$ , and  $H_2$  of model (2) were related with stand variables through correlation analysis and matrix-graphical analysis, and then were replaced by functions of the stand variables to develop a generalized height-diameter model.

**Figure 1.** Scatter plot of total tree height against diameter at breast height for the entire data set used.

### 2.2.2. Model fitting and selection

A fundamental least squares assumption is that the errors ( $\varepsilon$ ) in regression models are independent and identically distributed with mean zero and constant variance. However, in our case, in the scatter plot of total tree height against diameter at breast height for the entire data set (Fig. 1) an increasing height variance can be observed as values of the independent variable increase (heteroscedastic variance). Thus, weighted analysis is necessary to correct for heteroscedasticity. Without this correction, variance on the larger trees would be underestimated and variance on the smaller trees would be overestimated. Furthermore, minimum variance estimates and reliable prediction intervals can not be obtained [34].

In multiple regression, the error variance can be functionally related to two or more predictor variables [33]. Thus, to identify the correct function to model the unequal variance of the errors, the error variance was related to the explanatory variables. Following the proposal of Huang et al. [19], several assumptions about the nature of the heterocedasticity were analyzed. Finally, the weighting factors ( $\psi_j$ ) examined were:  $\psi_j = d^{K_1/K_2}$ ,  $\psi_j = d^{K_1/K_2} h_j^{K_1/K_2}$ ,  $\psi_j = D_0^{K_1/K_2}$ ,

$\psi_j = H_0^{K_1/K_2}$ , and  $\psi_j = \text{pred.} h_j^{K_1/K_2}$ , where  $K_1$  and  $K_2$  can take values of 0.1, 0.3, 0.5 and so on. For the weight factor  $\psi_j = \text{pred.} h_j^{K_1/K_2}$ , because the predicted heights are initially unknown, weighting is an iterative process. Parameters were estimated using generalized non-linear least squares (GNLS), also known as weighted non-linear regression, using the NLIN procedure of the SAS/STAT® statistical package [42].

Autocorrelation in the remeasured data set was ignored because there is only a small gain using complex techniques [47, 52] to account for this problem. Also, the impact of variance underestimation is likely masked by fitting each individual tree as an independent observation [16, 50].

The goodness-of-fit of the model was evaluated using two statistics: the root of the weighted mean square error ( $RMSE_\psi$ ) and the coefficient of determination ( $R^2$ ). The expressions for these statistics are the following:

$$RMSE_\psi = \sqrt{\frac{\sum_{j=1}^n \psi_j (h_j - \hat{h}_j)^2}{n-k}} \quad \text{and} \quad R^2 = 1 - \frac{\sum_{j=1}^n \psi_j (h_j - \hat{h}_j)^2}{\sum_{j=1}^n \psi_j (h_j - \bar{h})^2} \quad (3)$$

where  $h_j$ ,  $\hat{h}_j$ , and  $\bar{h}$  are the observed, the predicted and the average values of tree heights, respectively,  $n$  is the total number of observations,  $k$  is the number of parameters in the model and  $\psi_j$  is the weighting factor. Another important step in the evaluation of the fitted models was to perform a graphical analysis of the residuals, searching for dependencies or patterns that indicate systematic discrepancies [13, 22, 31, 44].

A portion of the data was not reserved for model validation. According to [18, 32, 59], the final estimation of the model parameters should come from the entire data set because the estimations obtained with this approach will be more precise than those obtained from the model fitted from the split data set. Other alternatives of validation, such as cross-validation, do not provide any additional information compared with the respective statistics obtained directly from the model fitted with the entire data set [26, 56].

### 2.2.3. Effect of thinning in height-diameter relationship

Thinning is perhaps the forester’s most basic silvicultural tool for moulding an even-aged stand, because it controls spacing, stand vigour, density, tree size distribution, and other stand and tree characteristics. Usually it is accepted that thinning has a very positive influence on residual tree diameter increment and fewer (if any) effects on tree height growth, so it is reasonable to expect that the relationship between total height and diameter at breast height will be different in thinned and unthinned stands [49, 61].

Two main questions were formulated in this work concerning the effect of thinning: (i) is there any difference on height-diameter relationship between thinned and unthinned stands? and if so (ii), is it possible to include an additional thinning response variable in the general model to explain these differences? If differences are found between these types of plots, modification of the generalized function might be done to model the response due to thinning. It is assumed that any formulation to take into account the effect of thinning must include

the thinning intensity, the stand age at the time of thinning and the time since thinning [17, 25, 39, 61].

In this work, to consider the thinning effect, the term proposed by Short and Burkhart [46] was selected and incorporated as a multiplier in model (2):

$$T_1 = \left(\frac{G_a}{G_b}\right)^{\alpha \left(\frac{A_i}{A}\right)} \quad (4)$$

where  $T_1$  is the effect of the thinning;  $G_a$  and  $G_b$  are the stand basal area after and before thinning, respectively;  $A_i$  is the stand age at thinning;  $A$  is the stand age at prediction and  $\alpha$  is the parameter to be estimated. This equation takes into account the thinning intensity and the time from thinning. Depending on the value of the parameter  $\alpha$ ,  $T_1$  has the following effects within model (2): If  $\alpha$  is zero, the  $h-d$  development of thinned plots is the same as that of unthinned ones; if  $\alpha$  is positive, the tree height of the thinned plots is less than that of the unthinned plots; and finally, if  $\alpha$  is negative, the tree height of the thinned plots is higher than the unthinned ones.

### 2.2.4. Comparison of height-diameter models between ecoregions

To compare the differences of the height-diameter function analysed between ecoregions we used two tests, both based on the likelihood-ratio test, for detecting simultaneous homogeneity among parameters: the non-linear extra sum of squares method [5, 23] and the  $\chi^2$  test proposed by Lakkis and Jones [24]. These tests are frequently applied to analyse differences among different geographic regions [3, 7, 19, 38, 40, 60].

Both methods require the fitting of reduced and full models. For the height-diameter model, the reduced model corresponds to the same set of parameters for the two ecoregions. The full model corresponds to different sets of parameters for each ecoregion and it is obtained by expanding each parameter including an associated parameter and a dummy variable to differentiate the two ecoregions:

$$\beta_i + \gamma_i I \quad i = 0, 1 \quad (5)$$

where  $\beta_i$  is a parameter of the reduced model;  $\gamma_i$  is the associated parameter of the full model and  $I$  is a dummy variable whose value is equal to 0 for interior ecoregion and 1 for coastal ecoregion.

The appropriate test statistics use the following expressions:

$$\text{Non-linear extra sum of squares} \quad F = \frac{SSE_R - SSE_F}{df_R - df_F} \div \frac{SSE_F}{df_F} \quad (6)$$

$$\text{Lakkis-Jones test} \quad L = (SSE_F / SSE_R)^{n/2} \quad (7)$$

where  $SSE_R$  is the error sum of squares of the reduced model;  $SSE_F$  is the error sum of squares of the full model;  $df_R$  and  $df_F$  are the degrees of freedom of the reduced and full model, respectively. The statistic  $-2 \ln(L)$  follows a  $\chi^2$  distribution with  $v = df_R - df_F$  degrees of freedom. The non-linear extra sum of squares follows an  $F$ -distribution.

### 2.2.5. Stochastic height prediction

All processes can be considered as the sum of two components: one deterministic and the other stochastic. Knowledge of the deterministic component is obtained with the model’s functional relations. The stochastic component represents influences beyond our present predictive capability, or deliberately omitted from the model [48]. Thus, if assessing the variability of the outcomes is one of the objectives of the

modelling procedures, most of the prediction equations would require a random component.

Stochastic prediction is appropriate when considering infinite sets of possible outcomes. In this sense, it is well-known that two trees with the same diameter within the same plot do not necessarily have the same height value. Thus, stochastic predictions are necessary to mimic the natural height variability observed in reality [35]. Basically, the variance components from a regression model linked with random numbers are used to create a stochastic prediction.

The Schnute function is a non-linear equation. A general non-linear model can be written as:

$$y_j = f(\mathbf{x}_j, \boldsymbol{\beta}) + \varepsilon_j \quad (8)$$

where  $y_j$  is the dependent variable (total height of the tree in our case),  $\mathbf{x}_j$  is a  $(P \times 1)$  nonstochastic vector of (tree and stand dimension) variables,  $\boldsymbol{\beta}$  is a  $(k \times 1)$  parameter vector,  $\varepsilon_j$  is a random error, and  $j$  represents the  $j$ th observation ( $j = 1, 2, \dots, n$ ). As it is shown, and unlike linear specifications, the number of parameters ( $k$ ), and the number of the independent variables ( $P$ ) do not necessarily coincide in non-linear models.

In non-linear estimation, the design matrix  $\mathbf{X}$  is replaced by the partial derivatives matrix  $\mathbf{Z}(\boldsymbol{\beta})$  defined as the transpose of the matrix  $\frac{\partial \mathbf{f}(\mathbf{X}, \boldsymbol{\beta})}{\partial \boldsymbol{\beta}}$ . That is:

$$\mathbf{Z}(\boldsymbol{\beta}) = \begin{bmatrix} \frac{\partial f(\mathbf{x}_1, \boldsymbol{\beta})}{\partial \beta_1} & \dots & \frac{\partial f(\mathbf{x}_1, \boldsymbol{\beta})}{\partial \beta_k} \\ \vdots & \ddots & \vdots \\ \frac{\partial f(\mathbf{x}_n, \boldsymbol{\beta})}{\partial \beta_1} & \dots & \frac{\partial f(\mathbf{x}_n, \boldsymbol{\beta})}{\partial \beta_k} \end{bmatrix}. \quad (9)$$

The generalized non-linear least squares estimate (GNLS) of the vector  $\boldsymbol{\beta}$  is that value of  $\boldsymbol{\beta}$  that minimizes the sum of squared errors:

$$SSE(\boldsymbol{\beta}) = \boldsymbol{\varepsilon}' \boldsymbol{\Psi}(\boldsymbol{\theta})^{-1} \boldsymbol{\varepsilon} = [\mathbf{y} - \mathbf{f}(\mathbf{X}, \boldsymbol{\beta})]' \boldsymbol{\Psi}(\boldsymbol{\theta})^{-1} [\mathbf{y} - \mathbf{f}(\mathbf{X}, \boldsymbol{\beta})] \quad (10)$$

where  $\boldsymbol{\Psi}(\boldsymbol{\theta})$  is a diagonal matrix of weights dependent on a fixed number  $q$  of parameters denoted by the  $(q \times 1)$  vector  $\boldsymbol{\theta}$ . The dimension of  $\boldsymbol{\theta}$ , and the precise way in which  $\boldsymbol{\Psi}$  depends on  $\boldsymbol{\theta}$ , relies on what assumptions are made about the error process. Under appropriate conditions, the GNLS estimate  $\mathbf{b}$  will be approximately normally distributed with mean  $\boldsymbol{\beta}$  and variance-covariance matrix that is consistently estimated by:

$$\mathbf{s}^2(\mathbf{b}) = \hat{\sigma}^2 [\mathbf{Z}(\mathbf{b})' \boldsymbol{\Psi}(\hat{\boldsymbol{\theta}})^{-1} \mathbf{Z}(\mathbf{b})]^{-1} \quad (11)$$

where the scalar  $\hat{\sigma}^2$  is the regression mean squared error, that is, (10) divided by the degrees of freedom:

$$\hat{\sigma}^2 = \frac{SSE(\hat{\boldsymbol{\beta}})}{n-k} = \frac{[\mathbf{y} - \mathbf{f}(\mathbf{X}, \hat{\boldsymbol{\beta}})]' \boldsymbol{\Psi}(\hat{\boldsymbol{\theta}})^{-1} [\mathbf{y} - \mathbf{f}(\mathbf{X}, \hat{\boldsymbol{\beta}})]}{n-k}. \quad (12)$$

This information can be used to form hypothesis tests and interval estimates on  $\mathbf{b}$  in an analogous manner to linear least squares.

In the special case that we want to know the prediction interval on an individual (new) outcome drawn from the distribution of  $y_j$ , the variance is:

$$\text{var}(\hat{y}_{j(NEW)}) = \hat{\sigma}^2 \psi_j(\hat{\boldsymbol{\theta}}) + \mathbf{z}(\mathbf{b})'_j \mathbf{s}^2(\mathbf{b}) \mathbf{z}(\mathbf{b})_j \quad (13)$$

where  $\psi_j(\hat{\boldsymbol{\theta}})$  is the  $j$ th diagonal element of the estimated weight matrix  $\boldsymbol{\Psi}(\hat{\boldsymbol{\theta}})$ , that is, the value of the weight function at observation  $j$ , and

the partial derivatives vector  $\mathbf{z}(\mathbf{b})'_j$  is the  $j$ th row of  $\mathbf{Z}(\mathbf{b})$  (see Eq. (9)).

The expression used to stochastically assign the heights to each tree of the sample is [35]:

$$\hat{h}_{sto} = \hat{h} + F_{U_1}^{-1} S_{\hat{h}} + F_{U_2}^{-1} S_{y.x} \quad (14)$$

where  $\hat{h}_{sto}$  is the stochastic height estimation,  $\hat{h}$  is the deterministic height obtained from (2),  $F_{U_i}^{-1}$  is the inverse of the standard normal distribution function, the  $U_i$ 's are independent uniform random variates on the interval  $[0, 1]$ ,  $S_{y.x}$  and  $S_{\hat{h}}$  are the standard error of estimate and the prediction, respectively. Considering (13), these standard errors are computed as:

$$S_{\hat{h}} = \sqrt{\mathbf{z}(\mathbf{b})'_j \mathbf{s}^2(\mathbf{b}) \mathbf{z}(\mathbf{b})_j} \quad (15)$$

$$S_{y.x} = \sqrt{\hat{\sigma}^2 \psi_j(\hat{\boldsymbol{\theta}})}. \quad (16)$$

To obtain the expected value of the random process, the whole sequence of computations is repeated with different random values. To generate these random values, the normal function was used through the NORMAL (SEED) function in SAS/STAT® package [42]. The NORMAL function is a scalar function that returns a pseudo-random number having a normal distribution with a mean of 0 and a standard deviation of 1. It needs an initial starting point, called a seed, that either the user or the computer clock supplies, and must be a non-negative integer: if a positive seed is used, it is possible to replicate the stream of random numbers, while if zero is used as the seed, the computer clock initializes the stream, and the stream of random numbers is not replicable.

The accuracy of the stochastic approach was evaluated by size classes for the explicative stand variables included in the model. By using this approach we attempted to decrease the current variability due to the different local height-diameter relationships for each plot, allowing study of the observed variability under more homogeneous conditions.

### 3. RESULTS AND DISCUSSION

#### 3.1. Height-diameter relationship

Initially, the local model (2) was fitted to the whole data set. All the parameters were significant, but it provided poor results ( $R^2 = 0.714$ ). Thus, it was necessary to relate tree height to diameter at breast height and a variety of stand variables to improve the fit. The equation (2) was first fitted to each one of the 493 plots of the database. Then, the estimated parameters  $H_2$ ,  $\beta_0$  and  $\beta_1$  were related to stand variables by means of graphical and correlation analysis. The coefficients  $\beta_0$  and  $\beta_1$  did not have a relationship with any stand variables, however  $H_2$  had a high correlation with dominant height  $H_0$  ( $R^2 = 0.980$ ). Thus, this parameter was replaced with this stand variable. Finally, the variable  $d_2$  (maximum diameter of the plot) was also replaced with the dominant diameter  $D_0$ , because  $d_2$  is strongly influenced by outlying observations [14]. With these substitutions, the equation (2) is constrained to estimate the dominant height of the stand when the diameter at breast height of the subject tree equals the dominant diameter of the stand.

Thus, the resulting model was:

$$h = \left[ 1.3^{\beta_0} + (H_0^{\beta_0} - 1.3^{\beta_0}) \frac{1 - e^{-\beta_0 d}}{1 - e^{-\beta_0 D_0}} \right]^{(1/\beta_0)} + \varepsilon. \quad (17)$$

**Table III.** Results of the Lakkis and Jones test ( $L$ -value) and of the non-linear extra sum of squares test ( $F$ -value) of the ecoregional differences for equation (18).

Equation	Reduced model			Full model			$n$	$L$ -value	$F$ -value
	$SSE_R$	$df_R$	$MSE_R$	$SSE_F$	$df_F$	$MSE_F$			
(18)	1130.4	20693	0.0546	1096.7	20691	0.053	20695	626.352*	16.851*

$SSE_R$  = error sum of squares of the reduced model,  $SSE_F$  = error sum of squares of the full model,  $df_R$  and  $df_F$  = degrees of freedom of the reduced and full model, respectively. Significant  $F$  and  $L$  values are marked with asterisk (\*).

As some works have pointed out (e.g., [8, 31, 44]), the inclusion of stand height (average or dominant height) as an independent variable in generalized height-diameter equations appears to be necessary in order to achieve acceptable predictions.

On the other hand, stand density, measured as the number of trees per hectare or as basal area, did not seem to have significant influence in the performance of the  $h$ - $d$  model. This result was not expected a priori, since stand density is the most obvious factor affecting a height-diameter relationship [58, 61]: in dense stands trees with the same diameter are taller than those in less dense stands. However, the inclusion of dominant diameter as an explanatory variable in equation (17) seems to be taking into account the competition level within the stand, since a close relationship between this variable and the number of trees per hectare was found.

Generalized non-linear least squares using the Marquardt algorithm of the NLIN procedure of SAS/STAT<sup>®</sup> package [42] was carried out to fit the final model to the entire data set. To correct the unequal variance of the residuals, the weighting function  $\psi_j = 1/d_j$  was chosen after several comparisons as the best assumption. That is, the variance of the error was directly proportional to a fixed function of the diameter at breast height:  $E(\varepsilon_j^2) = \sigma^2 d^{K_1/K_2}$  where  $K_1/K_2 = 0.3$ . This weighting factor stabilized the variance and provided a homogeneous residual plot.

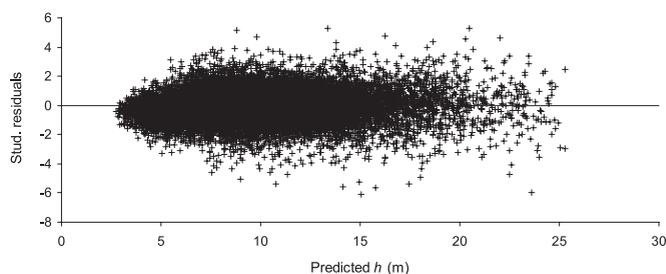
### 3.2. Comparison between ecoregions

Table III shows the values of the Lakkis-Jones test and the non-linear extra sum of squares method used to compare the differences in the height-diameter function between ecoregions. These results reveal that there are significant differences for the height-diameter relationship between ecoregions. Thus, the full model, which considers different sets of parameter for each ecoregion, was considered for further analysis. This model is written as follow:

$$h = \left[ 1.3\beta_0 + \gamma_0 I + \left( H_0 \beta_0 + \gamma_0 I - 1.3\beta_0 + \gamma_0 I \right) \frac{1 - e^{-(\beta_1 + \gamma_1 I)d}}{1 - e^{-(\beta_1 + \gamma_1 I)D_0}} \right]^{(1/(\beta_0 + \gamma_0 I))} + \varepsilon \quad (18)$$

where the dummy variable  $I$  is defined as 1 for the coastal region and as 0 for interior region.

On average, the addition of the stand variables to the height-diameter function (18) reduced the root mean square error by 53% compared to the results obtained by fitting equation (2)



**Figure 2.** Plot of studentized residuals against predicted values of height for equation (18).

(the  $RMSE_{\psi}$  of the full model is 0.8698). The percentage of variability explained by the model is 93.77%, which can be considered an adequate fit taking into account the great variability in the  $h$ - $d$  relationship in the database. A smaller value of the coefficient of determination  $R^2$  (0.919) was obtained by Schröder and Álvarez González [44] for the coastal region, working with a portion of the data used in this study and using the two stand variables dominant height and quadratic mean diameter.

The asymptotic 95% confidence intervals obtained for the parameter estimates from fitting model (18) with GNLS show reasonable values with all the parameters highly significant. This model also shows an approximately homogeneous variance over the full range of the predicted values and no systematic pattern in the variation of the residuals (Fig. 2).

In order to know the consequences of incorrectly mixing the height-diameter relationships in both ecoregions, and according to Huang et al. [19], the model fitted from the data in one ecoregion was used to make predictions for data from the other ecoregion. The mean error ( $\bar{\varepsilon}$ ) was calculated as follows:

$$\bar{\varepsilon} = \left[ \sum_{j=1}^n (h_j - \hat{h}_j) / n \right] \quad (19)$$

where  $n$  is the number of the predictions in the ecoregions.

The percent prediction bias was determined as bias (%) =  $\bar{\varepsilon} / \bar{h}$ , where  $\bar{h}$  is the average of the observed tree height.

Results showed a bias(%) of  $-7.41$  for the interior region and  $-1.28$  for the coastal region. The first value is 173 times larger than the value obtained using the appropriate equation and the second one is 4.8 times larger than the appropriate value. A  $t$ -test was also carried out to prove whether the mean of the new

prediction errors were zero or different. Results showed that means in both ecoregions were significantly different from zero and therefore overestimation (negative bias) occurs when a model developed for one of the ecoregions is applied to the other one. Also, the two prediction errors are not significantly different from zero (at  $\alpha=0.05$ ) when the appropriate equations for both ecoregions are used. These results, also obtained by Huang et al. [19], show that an important bias can be obtained using both equations interchangeably. Thus, a model with a different set of parameters for each ecoregion is needed.

### 3.3. Effect of thinning in height-diameter relationship

In the study of the effect of thinning in  $h$ - $d$  relationship, a new dummy variable to test if there were differences between thinned and unthinned plots was considered. The equation (18) was modified as follows:

$$h = \left[ 1.3^{\beta_0 + \gamma_0 I + \phi_0 Z} + \left( H_0^{\beta_0 + \gamma_0 I + \phi_0 Z} - 1.3^{\beta_0 + \gamma_0 I + \phi_0 Z} \right) \times \frac{1 - e^{-(\beta_1 + \gamma_1 I + \phi_1 Z)d}}{1 - e^{-(\beta_1 + \gamma_1 I + \phi_1 Z)D_0}} \right]^{(1/(\beta_0 + \gamma_1 I + \phi_0 Z))} + \varepsilon \quad (20)$$

where  $\phi_j$  is the parameter associated with the new dummy variable  $Z$ , which is defined as 1 when a thinning has been carried out and as 0 otherwise.

The results show that parameter  $\phi_1$  had a very large asymptotic confidence interval, even including zero, therefore it was removed of the model. The rest of the parameters were found to be highly significant. Thus, it was concluded that the thinning operations have a positive effect in the height-diameter relationship, increasing the height for a given diameter; however, the increase was very small. Model (20) shows a  $RMSE_{\psi} = 0.8654$  and a  $R^2 = 0.939$ . Comparing this model with (18), small differences between both can be observed (only a reduction of 0.5% in root weighted mean square error).

In an attempt to take into account the effect of thinning, equation (18) was refitted including the thinning effect term (4) proposed by Short and Burkhart [46]. All the parameters were found to be significant and the  $RMSE_{\psi}$  and  $R^2$  obtained were 0.8674 and 0.938 respectively (a reduction of only 0.3% in root weighted mean squared error over (18)). These results suggest that the stand explanatory variables in the general model (dominant height and dominant diameter) already account for thinning effects on the tree height-diameter relationship. Zhang et al. [61] and Leduc and Goelz [28] achieved similar results working with loblolly pine and longleaf pine stands, respectively. In fact, the evolution of dominant diameter in control (unthinned) and thinned plots was checked, observing a slight increment in this variable in thinned plots. These results are again similar to those obtained by Zhang et al. [61] in loblolly pine plantations.

These two stand variables,  $D_0$  and  $H_0$ , do not change immediately with a thinning from below, the type of thinning applied in many of the stands of this species in Galicia. Thus, the height-diameter relationship does not change immediately after this silvicultural treatment, which is consistent with the empirical reality.

Based on the above considerations, equation (21) was finally selected to express the deterministic height-diameter rela-

tionship for thinned and unthinned *Pinus pinaster* stands in Galicia.

$$\hat{h} = \left[ 1.3^{1.8943+1.4692I} + \left( H_0^{1.8943+1.4692I} - 1.3^{1.8943+1.4692I} \right) \times \frac{1 - e^{-(0.0461-0.0473I)d}}{1 - e^{-(0.0461-0.0473I)D_0}} \right]^{(1/(1.8943+1.4692I))} \quad (21)$$

### 3.4. Stochastic predictor of the tree heights

Deterministic values of height can be obtained using the previous model, while the prediction of the stochastic values involves the following steps:

1. Obtain the deterministic value of the height from equation (21).
2. Calculate the standard error of the prediction by means of equation (15), where the variance-covariance matrix of the parameter estimates is:

$$s^2(\mathbf{b}) = \begin{bmatrix} 1.290 \times 10^{-3} & -1.290 \times 10^{-3} & -6.024 \times 10^{-5} & 6.024 \times 10^{-5} \\ -1.290 \times 10^{-3} & 7.555 \times 10^{-3} & 6.024 \times 10^{-5} & -2.494 \times 10^{-4} \\ -6.024 \times 10^{-5} & 6.024 \times 10^{-5} & 3.109 \times 10^{-6} & -3.109 \times 10^{-6} \\ 6.024 \times 10^{-5} & -2.494 \times 10^{-4} & -3.109 \times 10^{-6} & 9.502 \times 10^{-6} \end{bmatrix}$$

The partial derivatives vector  $\mathbf{z}(\mathbf{b})_j$  and its transpose  $\mathbf{z}(\mathbf{b})_j'$  can be obtained substituting the estimated parameters in the partial derivatives of equation (18) shown in the Appendix.

3. Multiply the weighted root mean square error ( $RMSE_{\psi}$ ) obtained in the regression procedure (0.86983) times the square root of the weighting factor (in this case the diameter of the tree) to calculate the standard error of the estimate  $S_{y,x}$  (see Eq. (16)).

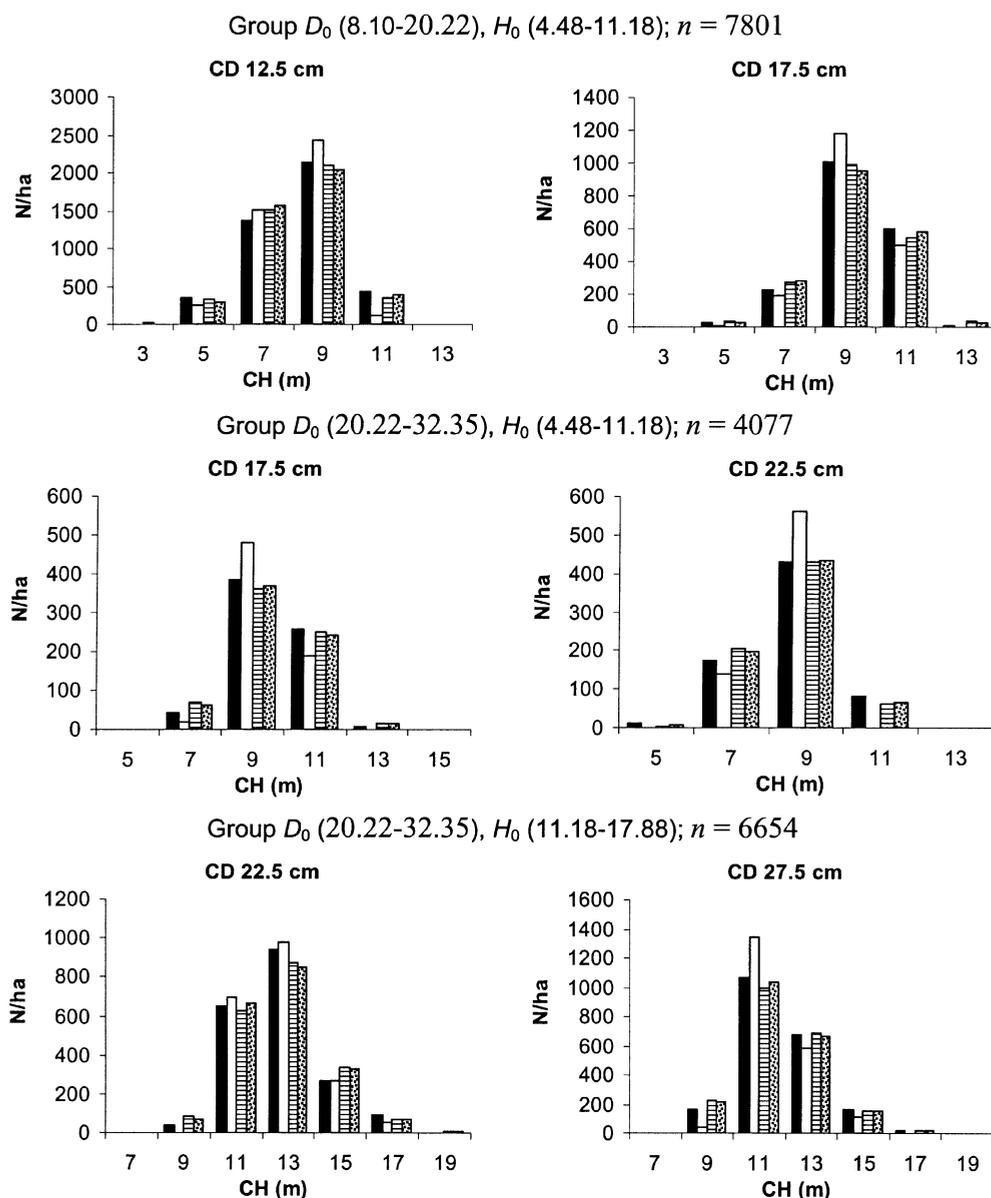
4. Calculate the inverse of a standard distribution function for two random numbers on the interval [0,1].

5. Estimate the stochastic height from equation (14) by substituting the values obtained in the previous steps.

In the present study, to obtain the expected values of the random process, a pseudo-random number was generated for each tree using the zero value as seed, since the stream of random numbers must change after each execution of the function. In addition, 100 repeated simulations were carried out for the entire data set, and the mean value from the 100 simulations was obtained for each tree.

Taking into account that dominant height and dominant diameter are the explicative stand variables of the selected model, the observed data set and the corresponding deterministic and stochastic estimates were split into nine groups, considering the three different combinations of the same width for dominant height and dominant diameter.

The height distribution for diameter classes with the largest number of observations within the three best represented groups is shown in Figure 3. The observed height distributions (which generally follow a normal distribution within each diameter class) were compared with those provided by the deterministic and two stochastic alternatives (one representing the mean value obtained from 100 simulations for each tree and the other selected at random). The deterministic estimate provided



**Figure 3.** Height distributions for diameter classes with the greatest number of observations within the three groups best represented in the data set. The black bars represent the observed distribution; the white bars represent the distribution obtained from the deterministic model; the striped and dotted bars represent the distributions generated by the stochastic approach using one pseudo-random number and the mean of 100 simulations, respectively. CH = height classes, CD = diameter classes.

a height distribution with low variability located around the observed mean value, whereas the two stochastic estimates provided greater variability, consistent with observed distribution (Fig. 3). The stochastic component therefore adequately mimics the variability in the observed heights within the same diameter classes, providing more realistic predictions at stand level.

#### 4. CONCLUSIONS

Four growth functions were considered to develop a height-diameter relationship; however, satisfactory results were only

obtained with the Schnute equation. Because poor results were found using a local function, a two-stage analysis was carried out that related the stand variables dominant height and dominant diameter to parameters of the Schnute model. The resulting generalization of the Schnute function achieved substantially better results. The new model implies a low sampling effort, since it only requires the measurement of diameters and one sample of heights for its practical application.

The generalized non-linear least squares method was used to account for the unequal errors variance, obtaining the inverse of the diameter at breast height as the best weighting function to remove the heterocedasticity problem.

The Lakkis-Jones and the non-linear extra sum of squares tests indicated that the height-diameter relationship is different between ecoregions. This is expected because the ecoregions have very different bio-geoclimatic conditions. Therefore, an ecoregion-based height-diameter model was developed.

Thinning had a minimal influence on the generalized height-diameter relationship. The results suggest that the stand explanatory variables in the general model (mainly dominant diameter) already account for the thinning effect on the tree  $h$ - $d$  relationship. Therefore, an additional thinning response variable was not included in the model. Model (21) was finally selected

to estimate deterministic values of height in thinned and unthinned *Pinus pinaster* stands in Galicia.

Finally, a stochastic component was added to the deterministic model. Stochastic height predictions were tested with real observations concluding that the model predictions performed acceptably well. The suggested approach allows for mimicking the natural variability in heights and therefore provides more realistic height predictions than the deterministic model. This feature is considered very important, since the height-diameter model developed in this study will be used to fill the missing heights for trees that have no height measurements.

**Appendix.** Partial derivatives of equation (18).

$$\begin{aligned} \frac{\partial h}{\partial \beta_0} &= \left[ 1.3^{(\beta_0 + \gamma_0 I)} + \left( H_0^{(\beta_0 + \gamma_0 I)} - 1.3^{(\beta_0 + \gamma_0 I)} \right) \frac{1 - e^{-(\beta_1 + \gamma_1 I)d}}{1 - e^{-(\beta_1 + \gamma_1 I)D_0}} \right]^{\frac{1}{(\beta_0 + \gamma_0 I)}} \times \\ &\quad \left[ \frac{1.3^{(\beta_0 + \gamma_0 I)} \ln 1.3 + \left( H_0^{(\beta_0 + \gamma_0 I)} \ln H_0 - 1.3^{(\beta_0 + \gamma_0 I)} \ln 1.3 \right) \frac{1 - e^{-(\beta_1 + \gamma_1 I)d}}{1 - e^{-(\beta_1 + \gamma_1 I)D_0}}}{(\beta_0 + \gamma_0 I) \left[ 1.3^{(\beta_0 + \gamma_0 I)} + \left( H_0^{(\beta_0 + \gamma_0 I)} - 1.3^{(\beta_0 + \gamma_0 I)} \right) \frac{1 - e^{-(\beta_1 + \gamma_1 I)d}}{1 - e^{-(\beta_1 + \gamma_1 I)D_0}} \right]} - \right. \\ &\quad \left. \frac{\ln \left[ 1.3^{(\beta_0 + \gamma_0 I)} + \left( H_0^{(\beta_0 + \gamma_0 I)} - 1.3^{(\beta_0 + \gamma_0 I)} \right) \frac{1 - e^{-(\beta_1 + \gamma_1 I)d}}{1 - e^{-(\beta_1 + \gamma_1 I)D_0}} \right]}{(\beta_0 + \gamma_0 I)^2} \right] \\ \frac{\partial h}{\partial \gamma_0} &= \left[ 1.3^{(\beta_0 + \gamma_0 I)} + \left( H_0^{(\beta_0 + \gamma_0 I)} - 1.3^{(\beta_0 + \gamma_0 I)} \right) \frac{1 - e^{-(\beta_1 + \gamma_1 I)d}}{1 - e^{-(\beta_1 + \gamma_1 I)D_0}} \right]^{\frac{1}{(\beta_0 + \gamma_0 I)}} \times \\ &\quad \left[ \frac{1.3^{(\beta_0 + \gamma_0 I)} I \ln 1.3 + \left( H_0^{(\beta_0 + \gamma_0 I)} I \ln H_0 - 1.3^{(\beta_0 + \gamma_0 I)} I \ln 1.3 \right) \frac{1 - e^{-(\beta_1 + \gamma_1 I)d}}{1 - e^{-(\beta_1 + \gamma_1 I)D_0}}}{(\beta_0 + \gamma_0 I) \left[ 1.3^{(\beta_0 + \gamma_0 I)} + \left( H_0^{(\beta_0 + \gamma_0 I)} - 1.3^{(\beta_0 + \gamma_0 I)} \right) \frac{1 - e^{-(\beta_1 + \gamma_1 I)d}}{1 - e^{-(\beta_1 + \gamma_1 I)D_0}} \right]} - \right. \\ &\quad \left. \frac{I \ln \left[ 1.3^{(\beta_0 + \gamma_0 I)} + \left( H_0^{(\beta_0 + \gamma_0 I)} - 1.3^{(\beta_0 + \gamma_0 I)} \right) \frac{1 - e^{-(\beta_1 + \gamma_1 I)d}}{1 - e^{-(\beta_1 + \gamma_1 I)D_0}} \right]}{(\beta_0 + \gamma_0 I)^2} \right] \\ \frac{\partial h}{\partial \beta_1} &= \frac{1}{(\beta_0 + \gamma_0 I)} \left[ 1.3^{(\beta_0 + \gamma_0 I)} + \left( h_2^{(\beta_0 + \gamma_0 I)} - 1.3^{(\beta_0 + \gamma_0 I)} \right) \frac{1 - e^{-(\beta_1 + \gamma_1 I)d}}{1 - e^{-(\beta_1 + \gamma_1 I)D_0}} \right]^{\left( \frac{1}{(\beta_0 + \gamma_0 I)} - 1 \right)} \left( h_2^{(\beta_0 + \gamma_0 I)} - 1.3^{(\beta_0 + \gamma_0 I)} \right) \times \\ &\quad \left[ \frac{d e^{-(\beta_1 + \gamma_1 I)d} \left( 1 - e^{-(\beta_1 + \gamma_1 I)D_0} \right) - D_0 e^{-(\beta_1 + \gamma_1 I)D_0} \left( 1 - e^{-(\beta_1 + \gamma_1 I)d} \right)}{\left( 1 - e^{-(\beta_1 + \gamma_1 I)D_0} \right)^2} \right] \\ \frac{\partial h}{\partial \gamma_1} &= \frac{1}{(\beta_0 + \gamma_0 I)} \left[ 1.3^{(\beta_0 + \gamma_0 I)} + \left( h_2^{(\beta_0 + \gamma_0 I)} - 1.3^{(\beta_0 + \gamma_0 I)} \right) \frac{1 - e^{-(\beta_1 + \gamma_1 I)d}}{1 - e^{-(\beta_1 + \gamma_1 I)D_0}} \right]^{\left( \frac{1}{(\beta_0 + \gamma_0 I)} - 1 \right)} \left( h_2^{(\beta_0 + \gamma_0 I)} - 1.3^{(\beta_0 + \gamma_0 I)} \right) \times \\ &\quad \left[ \frac{d I e^{-(\beta_1 + \gamma_1 I)d} \left( 1 - e^{-(\beta_1 + \gamma_1 I)D_0} \right) - D_0 I e^{-(\beta_1 + \gamma_1 I)D_0} \left( 1 - e^{-(\beta_1 + \gamma_1 I)d} \right)}{\left( 1 - e^{-(\beta_1 + \gamma_1 I)D_0} \right)^2} \right] \end{aligned}$$

**Acknowledgements:** This research study was conducted when the first two authors were Visiting Scholar at the Southern Research Station in Asheville, invited by Dr. Bernard Parresol. We are very grateful for the kind hospitality of the Station's staff. Helpful review comments were provided by Dr. Ulises Diéguez-Aranda and two anonymous referees. We also thank to Dr. Roque Rodríguez-Soalleiro for providing part of the necessary data for this study. This study was financed by the Comisión Interministerial de Ciencia y Tecnología (CICYT), project No AGL2001-3871-C02-01.

## REFERENCES

- [1] Alía R., Gil L., Pardos J.A., Performance of 43 *Pinus pinaster* provenances on 5 locations in central Spain, *Silvae Genet.* 44 (1995) 75–81.
- [2] Alía R., Moro J., Denis J.B., Performance of *Pinus pinaster* Ait. provenances in Spain: interpretation of the genotype-environment interaction, *Can. J. For. Res.* 27 (1997) 1548–1559.
- [3] Álvarez González J.G., Ruiz A.D., Rodríguez R.J., Barrio M., Development of ecoregion-based site index models for even-aged stands of *Pinus pinaster* Ait. in Galicia (north-western Spain), *Ann. For. Sci.* 62 (2005) 115–127.
- [4] Bará S., Toval G., Calidad de estación del *Pinus pinaster* Ait. en Galicia, Comunicaciones INIA, nº 24, Madrid, 1983.
- [5] Bates D.M., Watts D.G., Nonlinear regression analysis and its applications, John Wiley & Sons, New York, 1988.
- [6] Bredenkamp B.V., Gregoire T.G., A Forestry application of Schnute's generalized growth function, *For. Sci.* 34 (1988) 790–797.
- [7] 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.
- [8] Cañadas N., García C., Montero G., Relación altura-diámetro para *Pinus pinea* L. en el Sistema central, in: Rojo et al. (Eds.), Actas del Congreso de Ordenación y Gestión Sostenible de Montes, Santiago de Compostela, Tomo I, 1999, pp. 139–153.
- [9] Castedo F., Modelo dinámico de crecimiento para las masas de *Pinus radiata* D. Don en Galicia. Simulación de alternativas selvícolas con inclusión del riesgo de incendio, Ph.D. thesis, Escuela Politécnica Superior, University of Santiago de Compostela, 2004.
- [10] Clutter J.L., Fortson J.C., Pienaar L.V., Brister H.G., Bailey R.L., Timber management: a quantitative approach, John Wiley & Sons, New York, 1983.
- [11] Curtis R.O., Height-diameter and height-diameter-age equations for second-growth Douglas-fir, *For. Sci.* 13 (1967) 365–375.
- [12] Diéguez-Aranda U., Modelo dinámico de crecimiento para masas de *Pinus sylvestris* L. procedentes de repoblación en Galicia, Ph.D. thesis, Escuela Politécnica Superior, University of Santiago de Compostela, 2004.
- [13] Draper N.R., Smith H., Applied regression analysis, John Wiley & Sons, New York, 1981.
- [14] Fang Z., Bailey R.L., Height-diameter models for tropical forests on Hainan Island in Southern China, *For. Ecol. Manage.* 110 (1998) 315–327.
- [15] Ferguson I.S., Leech J.W., Generalized least squares estimation of yield functions, *For. Sci.* 24 (1978) 27–42.
- [16] Hamilton D.A., Edwards B.M., Modelling the probability of individual tree mortality, USDA Forest Service, Research Paper INT-185, Intermountain Forest and Range Experimental station, Ogden, UT, 1976.
- [17] Hasenauer H., Burkhardt H.E., Amateis R.L., Basal area development in thinned and unthinned loblolly pine plantations, *Can. J. For. Res.* 27 (1998) 265–271.
- [18] Hirsch R.P., Validation samples, *Biometrics* 47 (1991) 1193–1194.
- [19] 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.
- [20] Huang S., Titus S.J., An age-independent individual tree height prediction model for boreal spruce-aspen stands in Alberta, *Can. J. For. Res.* 24 (1994) 1295–1301.
- [21] Huang S., Titus S.J., Wiens D.P., Comparison of nonlinear height-diameter functions for major Alberta tree species, *Can. J. For. Res.* 22 (1992) 1297–1304.
- [22] 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.
- [23] Judge G.G., Carter R., Griffiths W.E., Lutkepohl H., Lee T.C., Introduction to the theory and practice of econometrics, John Wiley & Sons, New York, 1988.
- [24] Khattree R., Naik D.N., Applied multivariate statistics with SAS software, 2nd ed., SAS Institute Inc., Cary, North Carolina, 1999.
- [25] Knoebel B.R., Burkhardt H.E., Beck B.E., A growth and yield model for thinned stands of yellow-poplar, *For. Sci. Monogr.* 27, 1986.
- [26] Kozak A., Kozak R., Does cross-validation provide additional information in the evaluation of the regression models? *Can. J. For. Res.* 33 (2003) 976–987.
- [27] Larsen D.R., Hann D.W., Height-diameter equations for seventeen tree species in southwest Oregon, *Oreg. State. Univ. For. Res. Lab.* 46, 1987.
- [28] Leduc D.J., Goelz J.C., Personal communication, 2004.
- [29] Lei Y., Modelling forest growth and yield of *Eucalyptus globulus* Labill. in central-interior Portugal, Ph.D. thesis, Universidade de Trás-os-Montes e Alto Douro, 1998 (unpublished).
- [30] Lei Y., Parresol B.R., Remarks on height-diameter modelling, Research Note SRS-10, USDA Forest Service, Southern Research Station, Asheville, North Carolina, 2001.
- [31] López Sánchez C.A., Gorgoso J.J., Castedo F., Rojo A., Rodríguez R., Álvarez González J.G., Sánchez Rodríguez F., A height-diameter model for *Pinus radiata* D. Don in Galicia (Northwest Spain), *Ann. For. Sci.* 60 (2003) 237–245.
- [32] Myers R.H., Classical and modern regression with applications, 2nd ed., Duxbury Press, Belmont, California, 1990.
- [33] Parresol B.R., Modeling multiplicative error variance: An example predicting tree diameter from stump dimensions in baldcypress, *For. Sci.* 39 (1993) 670–679.
- [34] Parresol B.R., Additivity of nonlinear biomass equations, *Can. J. For. Res.* 31 (2001) 865–878.
- [35] Parresol B.R., Lloyd F.T., The stochastic tree modelling approach used to derive tree lists for the GIS/CISC identified stands at the Savannah River Site, Internal Report, USDA Forest Service, Southern Research Station, Asheville, North Carolina, 2004.
- [36] Peng C., Nonlinear height-diameter models for nine boreal forest tree species in Ontario, Ministry of Natur. Resour., Ontario For. Res. Inst., OFRI-Rep. 155, 1999.
- [37] Peng C., Developing and validating nonlinear height-diameter models for major tree species of Ontario's boreal forest, *North. J. Appl. For.* 18 (2001) 87–94.
- [38] Peng C., Zhang L., Huang S., Zhou X., Parton J., Woods M., Developing ecoregion-based height-diameter models for jack pine and black spruce in Ontario, Ministry of Natur. Resour., Ontario For. Res. Inst., OFRI-Rep. 159, 2001.
- [39] Pienaar L.V., Shiver B.D., Basal area prediction and projection equations for pine plantations, *For. Sci.* 32 (1986) 626–636.
- [40] Pillsbury N.H., McDonald P.M., Simon V., Reliability of tanoak volume equations when applied to different areas, *West. J. Appl. For.* 10 (1995) 72–78.
- [41] Richards F.J., A flexible growth function for empirical use, *J. Exp. Bot.* 10 (1959) 290–300.

- [42] SAS Institute Inc., SAS/STAT User's Guide, Release 8.2, Cary, North Carolina, 2000.
- [43] Schnute J., A versatile growth model with statistically stable parameters, *Can. J. Fish. Aquatic. Sci.* 38 (1981) 1128–1140.
- [44] Schröder J., Álvarez González J.G., Comparing the performance of generalized diameter-height equations for maritime pine in Northwestern Spain, *Forstw. Cbl.* 120 (2001) 18–23.
- [45] Sharma M., Zhang S.Y., Height-diameter models using stand characteristics for *Pinus banksiana* and *Picea mariana*, *Scand. J. For. Res.* 19 (2004) 442–451.
- [46] Short E.A., Burhart H.E., Predicting crown-height increment for thinned and unthinned loblolly pine plantations, *For. Sci.* 38 (1992) 594–610.
- [47] Soares P., Tomé M., Height-diameter equation for first rotation eucalypt plantations in Portugal, *For. Ecol. Manage.* 166 (2002) 99–109.
- [48] Stage A.R., How forest models are connected to reality: evaluation criteria for their use in decision support, *Can. J. For. Res.* 33 (2003) 410–421.
- [49] Staudhammer C., LeMay V., Height prediction equations using diameter and stand density measures, *For. Chron.* 76 (2000) 303–309.
- [50] Temesgen H., Gadow K.v., Generalized height-diameter models—an application for major tree species in complex stands of interior British Columbia, *Eur. J. For. Res.* 123 (2004) 45–51.
- [51] Vega P., Vega G., González M., Rodríguez A., Mejora del *Pinus pinaster* Ait. en Galicia, in: Silva Pando J. (Ed.), I Congreso Forestal Español, Vol. 2, 1993, pp. 129–134.
- [52] West P.W., Application of regression analysis to inventory data with measurements on successive occasions, *For. Ecol. Manage.* 71 (1995) 227–234.
- [53] Xunta de Galicia, O monte galego en cifras, Consellería de Medio Ambiente, Santiago de Compostela, 2001.
- [54] Xunta de Galicia, Anuario de estadística agraria 2000, Santiago de Compostela, 2002.
- [55] Yang R.C., Kozak A., Smith J.H., The potential of Weibull-type functions as a flexible growth curve, *Can. J. For. Res.* 8 (1978) 424–431.
- [56] Yang Y., Monserud R., Huang S., An evaluation of diagnostic tests and their roles in validating forest biometric models, *Can. J. For. Res.* 34 (2004) 619–629.
- [57] Zeide B., Accuracy of equations describing diameter growth, *Can. J. For. Res.* 19 (1989) 1283–1286.
- [58] Zeide B., Vanderschaaf C., The effect of density on the height-diameter relationship. In: Outcalt Kenneth W. (Ed.), Proc. 11th biennial southern silvicultural research conference, Gen. Tech. Rep. Southern Research Station—XX, Asheville, North Carolina, 2001, pp. 453–456.
- [59] Zhang L., Cross-validation of non-linear growth functions for modelling tree height-diameter relationships, *Ann. Bot.* 79 (1997) 251–257.
- [60] Zhang L., Peng L., Huang S., Zhou X., Development and evaluation of ecoregion-based tree height-diameter models for jack pine in Ontario, *For. Chron.* 78 (2002) 530–538.
- [61] Zhang S., Amateis R., Burkhart H.E., The influence of thinning on tree height and diameter relationships in loblolly pine plantations, *South. J. Appl. For.* 21 (1997) 199–205.