A generalized height–diameter model with random coefficients for uneven-aged stands in El Salto, Durango (Mexico)

BENEDICTO VARGAS-LARRETA[1*], FERNANDO CASTEDO-DORADO[2], JUAN GABRIEL ÁLVAREZ-GONZÁLEZ[3], MARCOS BARRIO-ANTA[4] AND FRANCISCO CRUZ-COBOS[1]

[1] Instituto Tecnológico de El Salto, Mesa del Tecnológico, 34942 El Salto, Durango, México
[2] Departamento de Ingeniería y Ciencias Agrarias, Universidad de León, Escuela Superior y Técnica de Ingeniería Agraria, Campus de Ponferrada, Avenida de Astorga, 24400 Ponferrada, Spain
[3] Departamento de Ingeniería Agroforestal, Universidad de Santiago de Compostela, Escuela Politécnica Superior, Campus universitario, 27002 Lugo, Spain
[4] Departamento de Biología de Organismos y Sistemas, Universidad de Oviedo, C/Gonzalo Gutiérrez de Quirós, 33600 Mieres, Spain
*Corresponding author. E-mail: bvargas@itf.edu.mx

Summary

A generalized height–diameter (*h-d*) model was developed in order to predict the total height of individual trees in uneven-aged stands in the region of El Salto, Durango (Mexico). Seven generalized *h-d* equations were evaluated and the equation proposed by Sharma and Parton, which includes the diameter at breast height of the tree and the quadratic mean diameter and the dominant height of the stand as independent variables, was selected as the best model. In order to address the among-plot variability, a non-linear mixed-effects modelling approach was used to fit the selected model for all the species or groups of species. The mixed model included a random parameter that affected the model asymptote. Calibration of the obtained *h-d* model for a particular species or group of species in a plot of interest was carried out with only a single randomly selected tree from the species or group of species within the plot. The stochastic component added to the mixed-effects model enabled the observed natural variability in heights within diameter classes for the same stand to be mimicked, thereby providing more realistic predictions. The equation developed represents a new tool for evaluation and management of uneven-aged stands in the region.

Introduction

Diameter at breast height (d.b.h.) over bark and total height are the variables most often measured in forest inventories. Devices that use ultrasound or laser pulses to measure distances have reduced the time needed to measure tree heights (*h*). However, measuring heights still requires more time than measuring d.b.h. (*d*) and therefore often the height of only a subset of trees of

Forestry, Vol. 82, No. 4, 2009. doi:10.1093/forestry/cpp016
Advance Access publication date 23 May 2009

known diameter is measured. Accurate height–diameter (*h-d*) equations must be used to predict heights of the remaining trees, thereby reducing data acquisition costs (Arabatzis and Burkhart, 1992; Huang *et al.*, 2000).

The *h-d* relationship is mainly used to characterize the vertical structure of forest stands (von Gadow *et al.*, 2001), to predict the height of individual trees in numerous forest growth simulators (Burkhart and Strub, 1974; Wykoff *et al.*, 1982; Larsen and Hann, 1987; Huang *et al.*, 1992; Nagel *et al.*, 2002) and to determine the dominant height for evaluating the productivity of a site (Lappi, 1997; Lei and Parresol, 2001). Knowledge of this relationship is also important in other contexts, including the estimation of total or merchantable volume and biomass of individual trees (Huang *et al.*, 1992; Peng *et al.*, 2001), simulation of the dynamics of forest stands (Canham *et al.*, 1994) and the analysis of the theoretical basis of tree growth (King, 1990).

Many studies have presented models for describing the *h-d* relationship. However, most of these models have been applied to pure, even-aged stands or plantations (e.g. Schröder and Álvarez-González, 2001; Soares and Tomé, 2002; López-Sánchez *et al.*, 2003; Calama and Montero, 2004; Gonda *et al.*, 2004; Diéguez-Aranda *et al.*, 2005; Castedo-Dorado *et al.*, 2006). Despite the homogeneous characteristics of this type of forest, a single *h-d* function is not usually adequate for all possible situations that can be found in pure and even-aged stands (e.g. Diéguez-Aranda *et al.*, 2005).

This is even more evident in mixed and uneven-aged stands in which different species, ages, sizes, crown types and tolerance levels coexist. Moreover, in this type of forest, the *h-d* relationships are often more difficult to describe because of the selective cuttings to which they are subjected and the changes in the natural dynamics of the stand (Oliver and Larson, 1990; Temesgen and von Gadow, 2004).

In summary, heterogeneous site conditions, the mixture of species and the different state of the stands are the reasons why a single *h-d* equation cannot describe all situations that can be found within a mixed, uneven-aged forest. A standard procedure in stand inventories is to establish a local *h-d* curve for each stand separately (van Laar and Akça, 2007). However, this procedure is costly and time consuming since at least 20–25 heights should be measured in each stand to obtain sufficiently accurate estimates (van Laar and Akça, 2007).

One practical alternative is to develop a generalized *h-d* relationship that describes the relationship between d.b.h. and total tree height and includes stand variables as predictors (e.g. dominant height, quadratic mean diameter, dominant diameter, number of trees per hectare, stand basal area, etc.).

The first generalized *h-d* function was developed by Wiedemann (1936). Since then many different formulations have been used to describe this relationship (e.g. Soares and Tomé, 2002; López-Sánchez *et al.*, 2003). However, little effort has been made to model the *h-d* relationship in uneven-aged stands with generalized *h-d* functions (e.g. Temesgen and von Gadow, 2004; Sharma and Parton, 2007).

With ~580 000 ha under management, the region of El Salto, Durango, is one of the most important forestry areas in Mexico. Timber production in the region is ~920 000 m^3 per year, which represents 44 per cent of the total production in the state of Durango (Vargas-Larreta, 2006).

Understanding the structure and development of these stands and the corresponding basic allometric relationships has increased during recent years. For example, Corral-Rivas *et al.* (2004, 2007a, b) have developed a site quality function, a merchantable volume system and equations for estimating tree volume from stump dimensions for the five main pine species in the region. One of the final allometric equations remaining to be developed is an *h-d* function.

Measurement of the height of four to six trees per plot and the assignment of an approximate value to the remaining trees according to the criteria and experience of the forest technician are very common practice in forest inventories carried out for management purposes in the region. This is due to a lack of *h-d* models that can be applied in estimating individual heights with a high degree of precision.

The main objective of the present study was therefore to develop equations that adequately describe the *h-d* relationship for the major species of mixed, uneven-aged stands in El Salto, Durango (Mexico). These equations represent a new management tool for forest managers and can be implemented in a whole-stand or in an individual-tree-level model for stands in the region.

Material and methods

Study area and data description

This research was carried out in the region of El Salto, Durango (Mexico), located in the Sierra Madre Occidental between a latitude of 23° 30′ and 24° 15′ N and a longitude of 105° 15′ and 105° 45′ W, 100 km to the south-west of the city of Durango (Figure 1). The predominant types of forest are pine–oak stands, which cover more than two-thirds of the region. The most important commercial species because of the physical and mechanical characteristics of their timber, their distribution and the usable volume are, in order of importance, *Pinus cooperi* Blanco, *Pinus durangensis* Martínez, *Pinus leiophylla* Sch. *et* Cham., *Pinus engelmannii* Carr, *Pinus cooperi* var. ornelasi, *Pinus teocote* Schl. *et* Cham., *Pinus herrerae* Martínez and some species of the genus *Quercus*.

Data including the d.b.h. over bark and total height of individual trees of five pine species, three evergreen oak and four other deciduous species were collected in the region. The data correspond to 55 permanent plots established by the *Unidad de Conservacion y Desarrollo Forestal de El Salto* in 1997 for developing growth and yield models for forest management programmes in the region. This network of plots was subjectively selected to represent the existing range of stand densities and site conditions. The size of the plot was 1000 m². In each plot, all the trees were labelled with a numbered plate. Callipers with millimetric graduation were used to make two measurements – at right angles to each other – of d.b.h. (1.3 m above ground level), and the arithmetic mean of the two

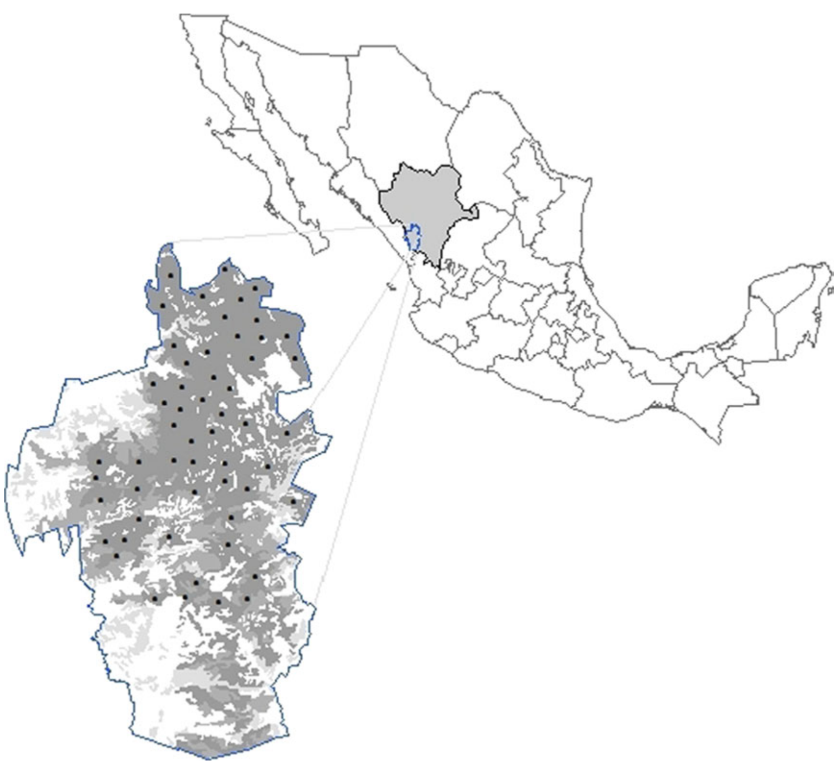

Figure 1. Location of the study area. The forest region of El Salto, Pueblo Nuevo, is in Durango State, in northern Mexico.

measurements was calculated. In addition, the total height of all the trees in the plot was measured with a digital hypsometer to the nearest 0.1 m.

In 2003, the 55 initially established plots were remeasured in the same way. For each inventory and plot, the following stand variables were calculated, considering all trees in each plot (regardless of the number of existing species): stand basal area, number of trees per hectare, quadratic mean diameter, dominant diameter (defined as the mean diameter of the 100 largest trees per hectare), dominant height (adopting the same criteria) and mean height. In addition, the basal area of largest trees (BAL) competition index was calculated for each tree as the basal area of all the trees per hectare that are larger than the tree under consideration (Wykoff *et al.*, 1982).

In total, 8270 pairs of *h*-*d* measurement were used. For statistical analysis, data from three species of the genus *Quercus* and from another five deciduous species were considered in two different groups because of the similarity among them and because only a few observations were available for some species. Summary statistics including the mean, minimum, maximum and standard deviation for tree and stand variables used for model fitting are shown in Table 1.

Models analysed

A large number of both linear and non-linear equations have been described in the forestry literature for modelling the *h*-*d* relationship with stand variables (see, e.g. Huang *et al.*, 1992; Soares and Tomé, 2002; López-Sánchez *et al.*, 2003; Diéguez-Aranda *et al.*, 2005; Trincado and Leal, 2006). Of the initially considered models, those that include the age of the stand as an independent variable were excluded because this variable has no practical use in uneven-aged stands (Huang and Titus, 1994). Seven generalized *h*-*d* equations were finally selected in the present study (Table 2). All of them have adequate mathematical properties and have performed satisfactorily in previous studies. These equations include the d.b.h. (d) and the stand variables dominant height (H_0) and quadratic mean diameter (D_g) as independent variables. One of them also includes a measure of a relative competitive position of a tree by means of a competition index (BAL).

Model selection and statistical analysis

In an initial exploratory phase, the pairs of *h*-*d* data were represented graphically for each species

Table 1: Summary statistics of the fitting data set

Species	Variable	No. of observations	Mean	SD	Maximum	Minimum
Pinus cooperi	d	1580	22.5	9.6	56.2	5.0
	h		14.9	5.4	27.3	3.0
Pinus durangensis	d	3226	17.9	8.6	58.5	5.0
	h		13.3	4.9	28	2.0
Pinus leiophylla	d	499	19.2	8.2	40	5.0
	h		12.5	5.0	24.5	2.0
Pinus teocote	d	595	20.5	10.1	51.4	5.0
	h		13.2	5.0	25.3	3.0
Pinus ayacahuite	d	352	14.8	7.1	47.5	5.0
	h		9.6	3.8	23.5	2.0
Quercus sp.	d	1330	17.7	12.2	90.0	5.0
	h		9.2	3.9	24.0	1.9
Other broadleaved species	d	688	15.2	8.7	61.4	5.0
	h		6.6	2.8	18	2.0
All species combined	d	8270	18.7	9.8	90.0	5.0
	h		12.2	5.3	28.0	1.9
	D_g		20.8	3.8	42.6	12.3
	H_0		18.6	3.0	25.2	10.8

d = d.b.h. over bark (cm); h = total tree height (m); D_g = quadratic mean diameter (cm); H_0 = dominant height (m); SD = standard deviation.

Table 2: Generalized *h-d* equations evaluated

Authors (year)	Mathematical form	Model
Hui and von Gadow (1993)	$h = 1.3 + b_0 \cdot H_0^{b_1} \cdot d^{b_2 \cdot H_0^{b_3}}$	M1
Soares and Tomé (2002)	$h = H_0 \cdot [1 + (b_0 + b_1 \cdot H_0 + b_2 \cdot D_g) \cdot e^{b_3 \cdot H_0}] \cdot \left(1 - e^{\frac{b_4 \cdot d}{H_0}}\right)$	M2
Mirkovic (1958)	$h = 1.3 + (b_0 + b_1 \cdot H_0 + b_2 \cdot D_g) \cdot e^{-\frac{b_3}{d}}$	M3
Schröder and Álvarez-González (2001)	$h = 1.3 + (b_0 + b_1 \cdot H_0 - b_2 \cdot D_g) \cdot e^{\frac{-b_3}{\sqrt{d}}}$	M4
Sharma and Parton (2007)	$h = 1.3 + b_0 \cdot H_0^{b_1} \left(1 - e^{-b_2 \cdot D_g^{-b_3} \cdot d}\right)^{b_4}$	M5
Temesgen and von Gadow (2004)	$h = 1.3 + (b_0 + b_1 \cdot \mathrm{BAL}) \cdot D_g^{b_2} \cdot d^{b_3}$	M6
Temesgen and von Gadow (2004) modified	$h = 1.3 + (b_0 + b_1 \cdot \mathrm{BAL}) \cdot D_g^{b_2} \cdot d^{b_3} \cdot H_0^{b_4}$	M7

d = d.b.h. over bark, h = total tree height, D_g = quadratic mean diameter, H_0 = dominant height, b_i = regression coefficients to be determined by model fitting.

and for the whole dataset (Figure 2). As was expected, in all cases, non-linear tendencies that described sigmoid- or concave-shaped curves were observed.

In a first step, estimation of the parameters was carried out with the NLIN procedure of SAS/STAT® statistical package (SAS Institute Inc., 2004). In order to ensure that the convergence obtained was global and did not correspond to a local optimum, the initial values of parameter estimates obtained in other similar studies were used.

The equations were evaluated to establish whether they fulfilled the fundamental least squares assumption that the errors (ϵ) are independent and identically distributed with mean zero and constant variance. Although more detailed statistical techniques are available, evaluation of these assumptions can easily be carried out by visual analysis of studentized residuals from regression fits (Huang *et al.*, 2000). Studentized residuals are used rather than ordinary residuals because the latter are intrinsically not independent and do not have common variance (e.g. Neter *et al.*, 1996).

Comparison of the estimation of the models was based on the graphical and numerical analysis of the residuals and three goodness-of-fit statistics: the root mean square error (RMSE), which analyses the precision of the estimations; the adjusted coefficient of determination (R^2_{adj}), which reflects the total variability that is explained by the model considering the total number of parameters to be estimated, and the Akaike's information criterion in differences (AICd), which is an index that is used to select the best model and is based on minimizing the distance of Kullback–Lieber (Burnham and Anderson, 1998). The expressions of these statistics are as follows:

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

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

$$\mathrm{AICd} = n \log \hat{\sigma}^2 + 2k - \min(n \log \hat{\sigma}^2 + 2k), \tag{3}$$

where y_i, $\hat{y}_i$ and $\bar{y}$ are the observed, predicted and average values of height, respectively; n is the total number of observations used to fit the model; p is the number of parameters to be estimated; $k = p + 1$ and $\hat{\sigma}^2$ is the maximum likelihood (ML) estimator of the variance of the error of the model obtained as

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}. \tag{4}$$

Equation (3) uses the ML estimator instead of the least square estimator of the variance of errors

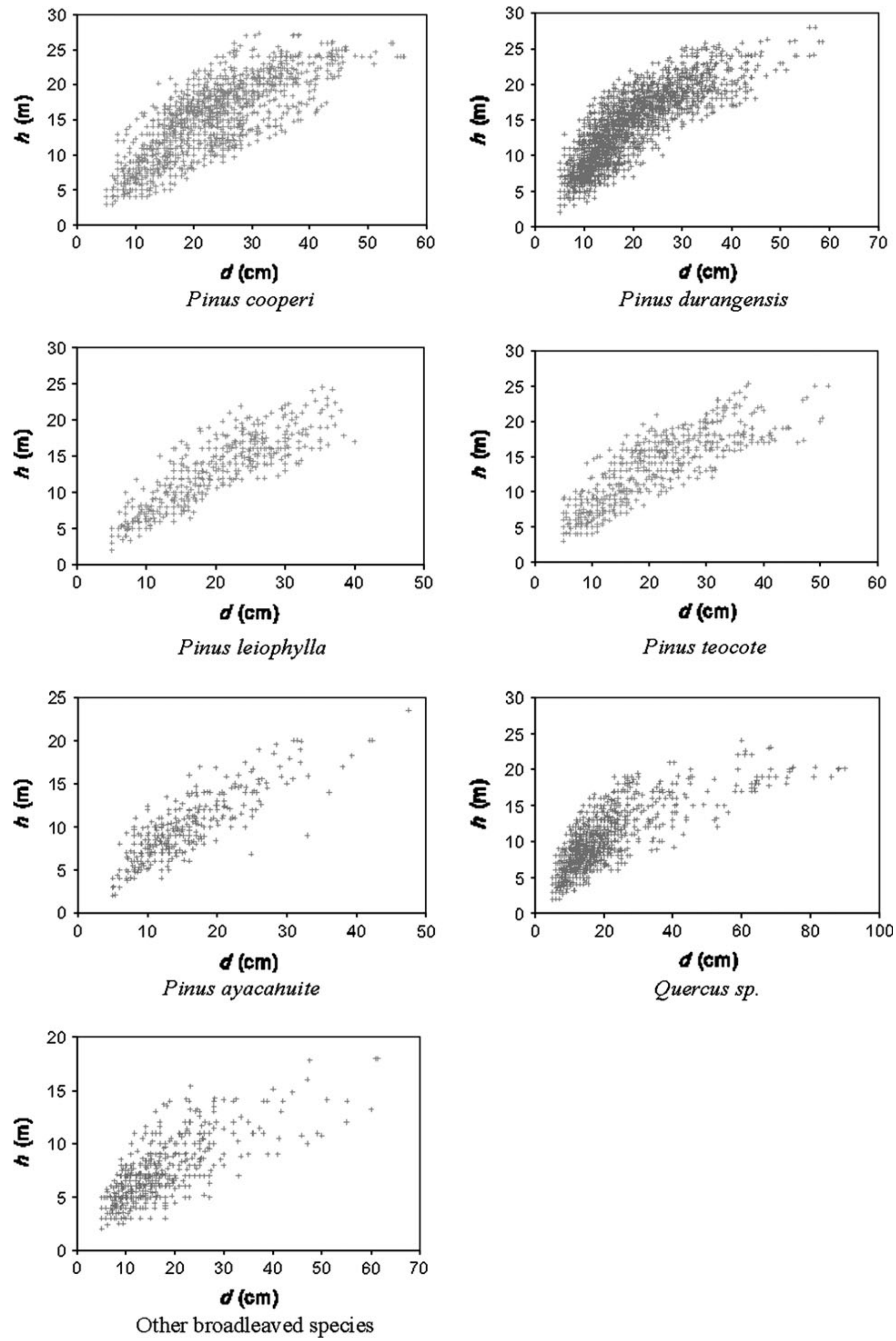

Figure 2. Scatter plots of total tree height against d.b.h. for each species.

because the original Akaike's information criterion is based on a relation between the relative Kullback–Lieber distance and the maximized log-likelihood (De Leeuw, 1992).

All *h-d* pairs of data were used to fit the models (i.e. a portion of the data was not reserved for model validation) as there was no really independent dataset available (which is a basic prerequisite for accomplishing a 'real validation'). The use of other alternatives, such as splitting the dataset in two parts, using one for fitting and the other for validation (cross-validation), does not provide any information in addition to the respective statistics obtained directly from the model fitted with the entire dataset (Kozak and Kozak, 2003; Yang *et al.*, 2004). Moreover, according to Hirsch (1991) and Zhang (1997), the final estimation of the model parameters should correspond to the entire dataset because the estimations obtained with this approach will be more precise than those obtained from the model fitted from the split dataset.

Non-linear mixed-effects modelling

Since the database contains multiple sample units (defined as each different species or group of species within each sample plot) from a large study area, it is reasonable to expect that important between-unit variability exists. For that reason, a mixed model approach, including random parameters to model that variability, has been used to fit the best model in the previous step. The NLMIXED procedure of SAS/STAT® statistical package (SAS Institute Inc., 2004) was used for this purpose.

A general expression for a non-linear mixed-effects model can be defined as (Lindstrom and Bates, 1990; Vonesh and Chinchilli, 1997; Pinheiro and Bates, 1998)

$$\boldsymbol{y}_i = f(\boldsymbol{\Phi}_i, \boldsymbol{x}_i) + \boldsymbol{e}_i, \tag{5}$$

where $\boldsymbol{y}_i$ is the $n_i \times 1$ vector of the n_i observations (trees) of the response variable taken from the ith unit, $\boldsymbol{x}_i$ is the $n_i \times 1$ predictor vector for the n_i observations of the predictor variable x taken from the ith unit, $\boldsymbol{\Phi}_i$ is a parameter vector, $r \times 1$ (where r is the number of parameters in the model), specific for each sampling unit, f is a non-linear function and $\boldsymbol{e}_i$ is an $n_i \times 1$ vector for the residual terms.

The main feature of mixed-effects models is that they allow parameter vectors to vary from unit to unit; therefore, the parameter vector $\boldsymbol{\Phi}_\mathrm{i}$ can be broken down into a fixed part, common to the population, and random components, specific to each unit:

$$\boldsymbol{\Phi}_i = \boldsymbol{A}_i\boldsymbol{\lambda} + \boldsymbol{B}_i\boldsymbol{b}_i, \tag{6}$$

where $\boldsymbol{\lambda}$ is the $p \times 1$ vector of fixed population parameters (where p is the number of fixed parameters in the model), $\boldsymbol{b}_i$ is the $q \times 1$ vector of random effects associated with the ith unit (where q is the number of random parameters in the model) and $\boldsymbol{A}_i$ and $\boldsymbol{B}_i$ are design matrices of size $r \times p$ and $r \times q$, for the fixed and random effects specific to each unit, respectively. A further explanation of the use of non-linear mixed models for modelling *h-d* relationships can be found in Calama and Montero (2004).

One important advantage of mixed models when predicting the value for a response variable is that the specific value of the random parameters vector ($\boldsymbol{b}_i$) for a given unit can be predicted if a supplementary sample of observations taken from that sampling unit is available. This approach makes a standard *h-d* model developed for a broad region more suitable for a specific unit, which is quite useful in a forest inventory or when it is desirable to assign heights to prospective stand tables arising from management planning. Therefore, the introduction of random parameters into the original model implies consideration of two different situations in a predictive role:

1. Fixed-effects response pattern: prediction of tree height in stands where only tree diameter and stand variables included in the model are measured. In this case, the predicted height is estimated using the model with only the fixed parameters and represents the mean behaviour of the pattern of the height variation for both the given diameter and the associated stand characteristics.
2. Calibrated response pattern: prediction of tree height in stands in which a subsample of tree heights, apart from diameter and stand variables, is available. If a subsample of k tree heights has been measured, such data can be used to predict the random-effects vector $\boldsymbol{b}_i$ of

the unit level. Prediction of $\boldsymbol{b}_i$ is carried out using the following expression (Vonesh and Chinchilli, 1997):

$$\hat{\boldsymbol{b}}_i \approx \hat{\boldsymbol{D}}\hat{\boldsymbol{Z}}_i^{T}(\hat{\boldsymbol{R}}_i + \hat{\boldsymbol{Z}}_i\hat{\boldsymbol{D}}\hat{\boldsymbol{Z}}_i^{T})^{-1}\hat{\boldsymbol{e}}_i, \tag{7}$$

where $\hat{\boldsymbol{D}}$ is the $q \times q$ variance–covariance matrix (where q is the number of random effects included in the model) for the among-unit variability, common to all units; $\hat{\boldsymbol{R}}_i$ is the $k \times k$ variance–covariance matrix for within-unit variability; $\hat{\boldsymbol{e}}_i$ is the residual vector $k \times 1$, whose components $\hat{e}_{ij}$ are given by the difference between the observed height for jth tree included in the subsample and the predicted height using the model including only fixed effects, and $\hat{\boldsymbol{Z}}_i$ is the $k \times q$ matrix evaluated at $\hat{\lambda}$:

$$\begin{bmatrix} \dfrac{\partial f(d_{i1},\Phi_i)}{\partial\lambda_1} & \cdots & \cdots & \dfrac{\partial f(d_{i1},\Phi_i)}{\partial\lambda_q} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \dfrac{\partial f(d_{ik},\Phi_i)}{\partial\lambda_1} & \cdots & \cdots & \dfrac{\partial f(d_{ik},\Phi_i)}{\partial\lambda_q} \end{bmatrix}, \tag{8}$$

where d_{ij} is the diameter of the jth tree in unit i.

Random-effects parameters predicted in such a way are added to the fixed parameters to obtain localized parameters. Heights of the trees measured only for diameters are then predicted in terms of diameter, stand variables and localized parameters. Details on the prediction of random-effects parameters in *h-d* context can be found in Calama and Montero (2004), Castedo-Dorado *et al.* (2006) and Trincado *et al.* (2007).

The number of trees measured for both height and diameter for a particular tree species in a given plot was not constant, and for some plot and species combinations, it was as low as 1. This is why only one tree was randomly selected from each unit (species or group of species within a sample plot) for prediction of random parameters. A SAS macro was implemented to predict the random parameter for each unit, with equation (7).

Stochastic height estimation

Two trees with the same diameter (or in the same diameter class) and within the same stand are not necessarily of the same height. Therefore, a deterministic model does not appear appropriate for mimicking the real natural variability in height (Castedo-Dorado *et al.*, 2005). To deal with this, an unstructured random component can be added to the mixed-effect model predictions. This approach assumes that the stochastic effects are entirely random and unstructured and adds a normally distributed random component, with variance equal to the residual mean square of a previously fitted model, to the model estimations (Dennis *et al.*, 1985; Fox *et al.*, 2001). Many other studies have found that incorporation of this random component is important for preserving variability in predictions (Stage, 1973, 2003; Daniels and Burkhart, 1975; Stage and Wykoff, 1993; Castedo-Dorado *et al.*, 2005).

Basically, this stochastic approach uses the standard error of prediction of a new observation in a similar way as when obtaining the prediction interval for an individual (new) in a regression model. Instead of using the t value corresponding to a fixed limit for all the trees, e.g. −1.96 and 1.96, respectively, for a probability of 2.5 and 97.5 per cent and infinite degrees of freedom, it is substituted by a randomly generated value from the inverse of the normal distribution function for each individual. Thus, to estimate the stochastic height $h_{ij(\mathrm{stoch})}$ of the ith tree in the jth plot, the following expression was used:

$$h_{ij(\mathrm{stoch})} = h_{ij} + F_U^{-1} S_{h_{ij(\mathrm{new})}} \tag{9}$$

where h_{ij} are the deterministic prediction values obtained in the previous step, $S_{h_{ij(\mathrm{new})}}$ is the standard error of the prediction for a new observation (see, e.g. Neter *et al.*, 1996) and F_U^{-1} is the inverse of the standard normal distribution function for U, a uniform random variable in the interval [0, 1]. More details about the stochastic height prediction are provided by Castedo-Dorado *et al.* (2005), Diéguez-Aranda *et al.* (2005) and Barrio-Anta *et al.* (2006).

The performance of the stochastic approach proposed was evaluated for diameter classes of 5 cm width and for each species (or group of species). A comparison between the observed height frequencies and the variability of the mixed-effects and stochastic approaches was carried out by analysis of box plots. Specifically, box plot graphs were used for visual comparison of the means, medians and 25/75 percentiles of the height distributions.

Results and discussion

Model selection

The parameter estimates and the goodness-of-fit statistics of the different generalized *h-d* models tested for each species (or group of species) are shown in Table 3. The model that performed best (lowest values of AICd and RMSE and highest values of R^2_{adj}) was, for most species, model M5 proposed by Sharma and Parton (2007). Specifically, for four of the five pine species (*P. cooperi*, *P. durangensis*, *P. leiophylla* and *P. teocote*) and for *Quercus* species, equation (M5) proposed by Sharma and Parton (2007) is the model that presents the greatest accuracy and precision.

Model M4, proposed by Schröder and Álvarez-González (2001), and model M2, proposed by Soares and Tomé (2002), performed adequately for all the pine species and accounted for approximately the same percentage of variance as model M5. However, problems of convergence were observed for *Quercus* and other broadleaved species with model M2, and for some pine species non-significant parameter estimates were obtained (Table 3).

In general, models M1 (Hui and von Gadow, 1993) and M6 (Temesgen and von Gadow, 2004) performed worst. The latter showed the second best results for *Pinus ayacahuite* and for the group of other broadleaved species, although in both cases, parameter b_2 (which affects quadratic mean diameter) was not significantly different from zero for a significance level of 5 per cent.

From the results of the model-fitting phase, it can be inferred that the stand variables considered in models M2, M3, M4 and M5 (quadratic mean diameter and dominant height) are sufficient to explain most of the variability in total tree height. Although the expressions of models M3 and M4 are very similar, the inclusion of a square root transformation of the diameter (rather than the diameter) in the denominator of the exponent substantially improves the accuracy and precision of the height estimations. Similar results were obtained by Schröder and Álvarez-González (2001) for fitting both models to data from stands of *Pinus pinaster* Ait. in north-west Spain.

The good performance of the above-mentioned four models may be partly due to the inclusion of the dominant height as an independent variable in their formulations. Many other authors (e.g. López *et al.*, 2003; Castedo-Dorado *et al.*, 2005, 2006; Trincado and Leal, 2006) have found that stand height is necessary in order to achieve acceptable individual tree height predictions.

Moreover, in most growth modelling studies, both tree and stand development are linked to the development of dominant height (e.g. Borders, 1989). This is justified because the dominant height indicates the site quality in terms of the stand growth and yield capacity (e.g. Eerikäinen, 2003; Trincado and Leal, 2006).

The inclusion of the quadratic mean diameter as an explanatory variable appeared to take into account the level of competition within the stand as there is a close relationship between this variable and the number of trees per hectare. As many authors have stated (e.g. Zhang *et al.*, 1997; Zeide and Vanderschaaf, 2002), stand density is the most obvious factor that affects the *h-d* relationship, particularly for trees grown in mixedwood stands (Huang and Titus, 1994): more dense stands tend to result in taller trees considering the same diameter (provided that the other conditions are the same).

On the other hand, the poor performance of model M6 proposed by Temesgen and von Gadow (2004) was not expected *a priori* since it includes the relative position of a tree through the BAL. Temesgen and von Gadow (2004) found this measurement of competition highly appropriate for improving the accuracy of height prediction for uneven-aged stands in British Columbia. The poor results of this model with the present data may be explained by the absence of dominant height in the original formulation of the model. The inclusion of this variable in model M6 (i.e. model M7) results in a better performance; nevertheless, in most cases, the goodness-of-fit statistics were worse than those obtained for models M2, M3, M4 and M5.

In general, considering all the species analysed, models M2, M3, M4 and M5 performed similarly and accurately. However, model M5 performed slightly better for five of the seven species (or group of species) and all the parameter estimates were significant at a probability level of 5 per cent. The percentage of the variability in tree heights explained by this model varied between 73 and 86 per cent for the pine species and was lower for the group of species of the genus *Quercus* and for other broadleaved species (72 and 64 per cent,

Table 3: Parameter estimates and goodness-of-fit statistics for the generalized *h-d* equations evaluated

	Estimated parameters					Goodness-of-fit statistics			
Species/model	b_0	b_1	b_2	b_3	b_4	R^2_{adj}	Bias	RMSE	AICd
Pinus cooperi									
M1	0.021*	1.56	1.83	−0.373		0.81	−0.056	2.341	395.5
M2	−0.841	0.129	−0.029	−0.059	−0.865	0.84	−0.041	2.167	144.0
M3	5.60	1.34	0.231	13.3		0.84	−0.031	2.146	120.2
M4	14.0	2.53	0.516	6.00		0.85	0.007	2.146	72.3
M5	1.09	0.990	0.380	0.490	1.87	0.85	−0.012	2.097	0
M6	0.811	0.184	−0.167	0.988		0.73	−0.022	2.806	967.4
M7	0.289	0.026	−0.402	0.783	0.870	0.84	−0.045	2.195	144.2
Pinus durangensis									
M1	0.055*	1.18	1.63	−0.299		0.78	−0.041	2.303	539.7
M2	0.118	0.053*	−0.026	−0.026	−0.824	0.81	−0.029	2.157	117.0
M3	12.1	1.09	0.316	12.6		0.81	0.032	2.157	116.2
M4	25.9	2.19	0.682	6.108		0.81	−0.014	2.147	86.6
M5	1.98	0.800	0.275	0.428	1.64	0.82	−0.001	2.118	0
M6	1.05	0.134	−0.119	0.897		0.77	−0.028	2.358	691.6
M7	0.422	0.034	−0.357	0.825	0.650	0.80	−0.039	2.200	244.5
Pinus leiophylla									
M1	0.022*	1.32	1.94	−0.297		0.83	−0.036	2.043	78.7
M2	0.691*	0.019*	−0.023*	0.012	−0.484	0.85	−0.013	1.913	14.1
M3	16.0	1.16	0.503	15.0		0.85	0.048	1.931	22.7
M4	38.3	2.69	1.23	7.23		0.86	−0.006	1.890	0.8
M5	2.46	0.767	0.261	0.506	1.57	0.86	−0.001	1.886	0
M6	0.569	0.075	−0.117*	1.06		0.80	−0.019	2.230	165.9
M7	0.358	0.017	−0.473	0.917	0.726	0.85	−0.034	1.937	26.8
Pinus teocote									
M1	0.015*	1.68	2.40	−0.462		0.81	−0.030	2.213	88.4
M2	0.151*	0.042	−0.017*	−0.021	−0.758	0.82	−0.006	2.117	36.7
M3	9.23	1.08	0.187	12.5		0.82	0.055	2.138	47.4
M4	18.8	2.16	0.447	5.95		0.83	−0.005	2.096	23.6
M5	1.99	0.807	0.224	0.405	1.47	0.83	0.004	2.053	0
M6	1.25	0.135	−0.125	0.826		0.75	−0.027	2.527	246.2
M7	0.305	−0.002*	−0.226	0.658	0.840	0.81	−0.048	2.176	69.3
Pinus ayacahuite									
M1	3.62*	−0.461*	0.354*	0.285*		0.72	−0.017	1.996	16.9
M2	0.460	−0.009	−0.006*	0.054	−0.581	0.73	0.028	1.979	11.6
M3	20.4	0.322	0.205*	13.1		0.70	0.056	2.075	44.3
M4	48.1	0.890	0.603	6.85		0.73	0.020	1.989	14.2
M5	12.0	0.334	0.067	0.357	1.04	0.73	−0.001	1.975	10.4
M6	1.14	0.074	−0.169*	0.892		0.73	−0.013	1.966	6.0
M7	0.962	0.054	−0.370	0.901	0.266	0.73	−0.010	1.946	0
Quercus sp.									
M1	0.156*	0.763	0.949	−0.153*		0.68	−0.049	2.2241	178.7
M2	NC	NC	NC	NC	NC	–	–	–	–
M3	8.89	0.756	0.171	13.5		0.71	0.054	2.135	69.4
M4	18.5	1.26	0.275	6.02		0.71	−0.009	2.105	32.5
M5	2.03	0.726	0.161	0.364	1.34	0.72	0.003	2.079	0
M6	0.742	0.069	0.043*	0.735		0.67	−0.041	2.251	210.8
M7	0.429	0.028	−0.176	0.697	0.468	0.69	−0.051	2.201	151.6

Table 3: Continued

Species/model	Estimated parameters					Goodness-of-fit statistics			
	b_0	b_1	b_2	b_3	b_4	R^2_{adj}	Bias	RMSE	AICd
Other broadleaved species									
M1	0.407*	0.215*	0.631	0.051*		0.63	−0.021	1.689	20.7
M2	NC	NC	NC	NC	NC	–	–	–	–
M3	8.08	0.173	−0.122	12.6		0.61	0.055	1.738	59.2
M4	17.2	0.376	−0.257	6.34		0.64	0.018	1.677	10.8
M5	8.44	0.257	0.013	−0.190*	0.963	0.64	−0.001	1.672	7.5
M6	0.297	0.031	0.148*	0.841		0.64	−0.010	1.666	1.2
M7	0.256	0.024	0.079*	0.831	0.140*	0.64	−0.011	1.663	0

NC = no convergence was achieved.
* No significant parameter estimate.

respectively). The poorer performance of model M5 for *Quercus* and other broadleaved species was expected, considering the wide range of heights within the same diameter class (see scatter plot of the dataset in Figure 2). Moreover, the results obtained for other allometric relationships (e.g. crown diameter–d.b.h., individual tree diameter increment–competition indices) also performed poorly for this type of species (Vargas-Larreta, 2006). The percentage of variability explained by model M5 was lower than that reported by other authors (>90 per cent in the studies of Peng *et al.*, 2001; Schröder and Álvarez-González, 2001; López-Sánchez *et al.*, 2003; Castedo-Dorado *et al.*, 2005; Diéguez-Aranda *et al.*, 2005), but similar to or even higher than those found by other authors (e.g. Lei and Parresol, 2001; Colbert *et al.*, 2002; Barrio-Anta *et al.*, 2006). It is important to emphasize that in all the cited studies, the data correspond to even-aged stands, in which the variability existing in the pairs of data of *h-d* within a stand is lower than that in uneven-aged stands.

The observed high variability in the *h-d* relationship could be explained by the high level of competition among species and by the different response of the species to the same conditions of local growth (e.g. Peng *et al.*, 2004). This would promote rapid differentiation of the height growth rates among the trees in the stand. Furthermore, the selective use of a few species (promoted by the forest management plans during various decades) has resulted in stands with a very variable spatial distribution and sociological and dimensional structure. According to Saunders and Wagner (2008), silvicultural practices and stand structure have a significant influence on modelling the *h-d* relationships.

Taking all these considerations into account, model M5 proposed by Sharma and Parton (2007) was selected as the best model for describing the *h-d* relationship for each of the species (or group of species) and was used for non-linear mixed-effects modelling. According to the classification proposed by López-Sánchez *et al.* (2003), the selected model can be considered as an intermediate sampling effort model because it requires measurement of the diameter of all the trees and of a sample of tree heights for its practical application.

Non-linear mixed-effects modelling

According to Fang and Bailey (2001), the following three steps are necessary for constructing a mixed-effects model: (1) determination of fixed- and random-effects parameters in the model, (2) specification of the within-unit variance–covariance structure for explaining variability among trees in the same unit and (3) definition of the structure of the among-unit variance–covariance matrix.

Initially, all parameters in model M5 were considered mixed, as was suggested by Pinheiro and Bates (1998), but the model failed to converge.

According to Sharma and Parton (2007), the number of random parameters was systematically reduced to achieve convergence. Only the models with parameters b_0 or b_1 (the two parameters determining the asymptote) as mixed converged, although the best values of the goodness-of-fit statistics were obtained by considering b_0 as mixed, composed of a fixed and a random varying part. Similar results were obtained by Sharma and Parton (2007) with the same model for different boreal tree species.

The expression of the mixed model M5 including b_0 as a mixed-effects parameter is the following:

$$h_{ij} = 1.3 + (b_0 + u_i) \cdot H_{0i}^{b_1} \cdot \left(1 - e^{-b_2 \cdot D_{gi}^{-b_3} \cdot d_{ij}}\right)^{b_4} + e_{ij} \quad \text{with } u_i \sim N[0, \sigma_u], \tag{10}$$

where b_1 to b_4 are considered fixed parameters, common to every unit; u_i is the random parameter, specific for unit i; d_{ij}, h_{ij} and e_{ij} are the diameter, height and estimation error for the jth observation in the ith unit, respectively, and H_{0i} and D_{gi} are the dominant height and quadratic mean diameter of unit i, respectively. Calibrated response (predicted height carried out over units with a prior height observation) for each tree in each unit was obtained with equation (10).

The general structure for the within-unit variance–covariance matrix $\boldsymbol{R}_i$ including correlation effects and weighting factors to balance error variance is given by

$$\boldsymbol{R}_i = \sigma_e{}^2 \boldsymbol{G}_i{}^{0.5} \Gamma_i \boldsymbol{G}_i{}^{0.5}, \tag{11}$$

where for a given unit i with n_i h-d measurements, $\boldsymbol{G}_i$ is an $n_i \times n_i$ diagonal matrix describing heterogeneous variance, since its components are standard deviations of the residual errors; Γ_i is an $n_i \times n_i$ matrix showing the structure of the correlation among observations for unit i and σ_e^2 is a scaling factor for the error dispersion (Gregoire *et al.*, 1995), given by the value of the residual variance of the model.

The variance–covariance matrix for the random effects, $\boldsymbol{D}$, common to all the units, defines variability existing among units. As there is only one random-effects parameter, the variance–covariance matrix of random effects is $D = \sigma_u^2$.

The parameter estimates and the goodness-of-fit statistics obtained for each species (or group of species) by the non-linear mixed-effects approach are shown in Table 4. The inclusion of the random parameter supposes a reduction in the RMSE of between 3.7 per cent in *P. cooperi* to 13.3 per cent in *P. ayacahuite*. The variance of the random parameter also varied across species from 0.001 to 0.469 for *P. cooperi* and *P. ayacahuite*, respectively.

The assumption of homogeneous variance was examined visually by plotting studentized residuals as a function of predicted heights obtained with mixed model M5 for each pine species and for the groups of *Quercus* and other broadleaved species (Figure 3). To analyse trends of underestimation

Table 4: Parameter estimates and goodness-of-fit statistics for the generalized h-d model M5 fitted by use of the non-linear mixed-effects modelling approach

	Estimated parameters							Goodness-of-fit statistics		
Species/model	b_0	b_1	b_2	b_3	b_4	σ_e^2	σ_u^2	R^2_{adj}	Bias	RMSE reduction (%)*
Pinus cooperi	1.015	1.017	0.306	0.429	1.816	4.126	0.001	0.86	−0.009	3.72
Pinus durangensis	1.992	0.799	0.241	0.403	1.555	3.997	0.022	0.84	−0.011	6.09
Pinus leiophylla	2.613	0.760	0.219	0.478	1.491	3.277	0.014	0.87	−0.005	5.36
Pinus teocote	2.943	0.699	0.138	0.308	1.291	3.499	0.057	0.86	−0.011	11.93
Pinus ayacahuite	25.71	0.422	0.016	0.374	0.894	3.101	0.469	0.79	−0.009	13.27
Quercus sp.	1.383	0.844	0.268	0.508	1.388	3.654	0.024	0.77	−0.005	9.23
Other broadleaved species	3.708	0.482	0.031		0.980	2.267	0.257	0.72	0.001	12.14

* Percentage of RMSE reduction with regards to RMSE obtained using an Ordinary Least Squares fitting method.

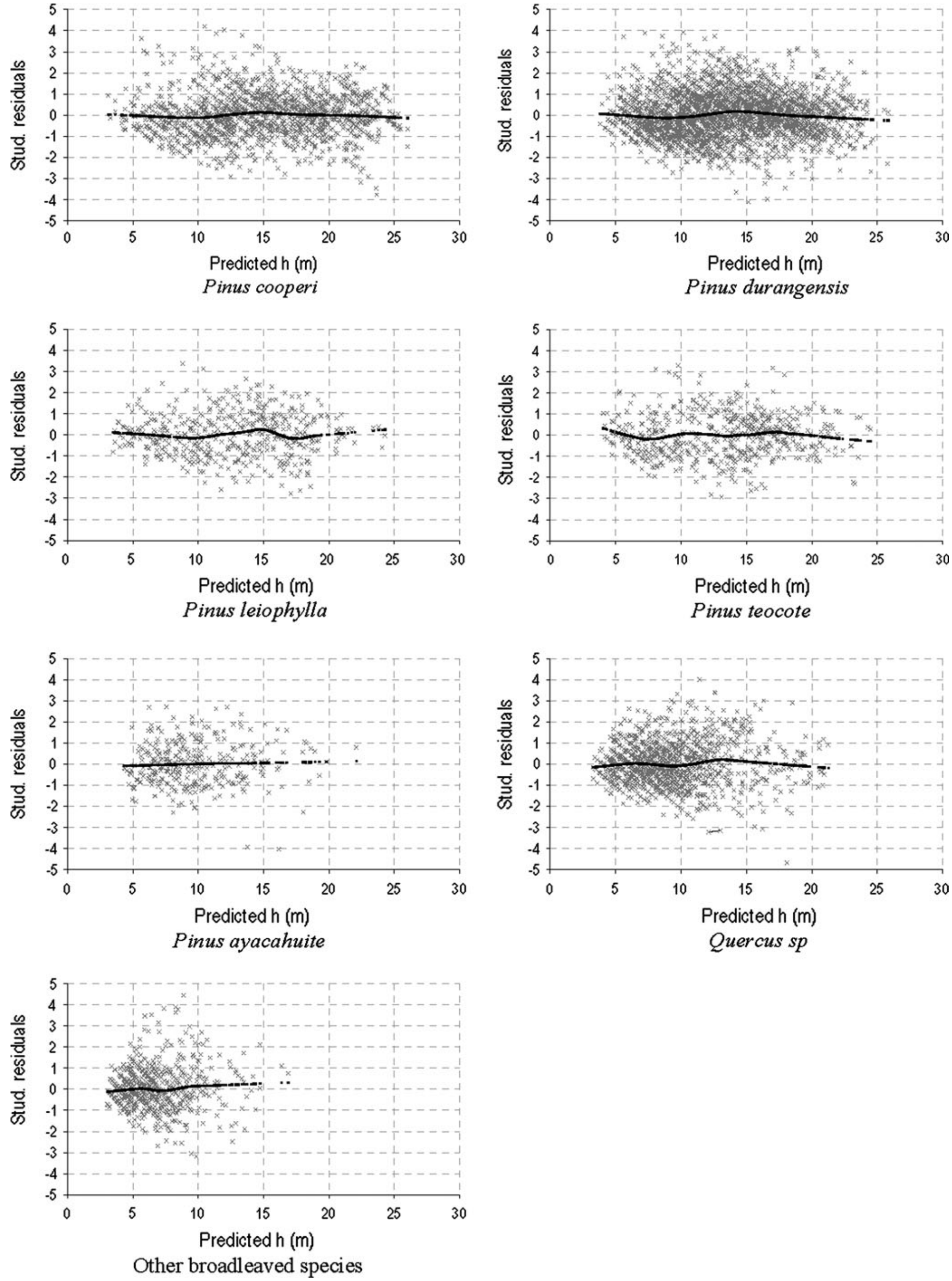

Figure 3. Plot of studentized residuals against predicted height values from model M5 for each species fitted using the non-linear mixed-effects approach together with a local regression loess smoothing curve.

or overestimation of tree height, a non-parametric curve was fitted by local regression with the LOESS procedure of SAS/STAT® statistical package (SAS Institute Inc., 2004). The smoothing parameter value used in the loess fit was obtained by means of a method that minimizes a bias-corrected Akaike's information criterion (Hurvich *et al.*, 1998).

The scatter plots showed homogeneous variance over the full range of the predicted values and no systematic pattern in the variation of the residuals. Random effects in equation (10) may help to remove some heterogeneity in the variance if it exists (Fang and Bailey, 2001).

As the database contains two measurements of each tree, it would be expected that these observations would be correlated, which violates the assumption of independent errors. To evaluate the presence of autocorrelation, a graph representing residuals versus age-lag residuals from the previous observations within each tree were examined visually (Figure 4). No trends in residuals as a function of age-lag residuals within the same tree were observed.

Since homogeneous variance and independence of error exist, the expression of within-unit variance–covariance matrix $\boldsymbol{R}_i$ is $\sigma_e^2 \cdot \boldsymbol{I}$, where $\boldsymbol{I}$ is the identity matrix $n_i \times n_i$.

Stochastic height estimation

Fixed-effects response prediction values of individual tree height (h_{ij}) can be obtained from model M5 (equation (10), without including the random parameter u_i specific for each unit). Calibrated response prediction values of tree height can be obtained from one randomly selected tree, with equation (10). A stochastic height $h_{ij(\mathrm{stoch})}$ for a given tree may be predicted by adding an unstructured stochastic component to the estimated height by use of the mixed or fixed-effect models. In the present study, the performance of the stochastic approach was evaluated for each species (or group of species) and diameter class, by use of the fitting dataset.

The box plots for comparison between the observed height frequencies and the height frequencies obtained with the calibrated function (equation (10) including the random parameter u_i specific for each unit), with and without the inclusion of an unstructured random component, are shown in Figure 5.

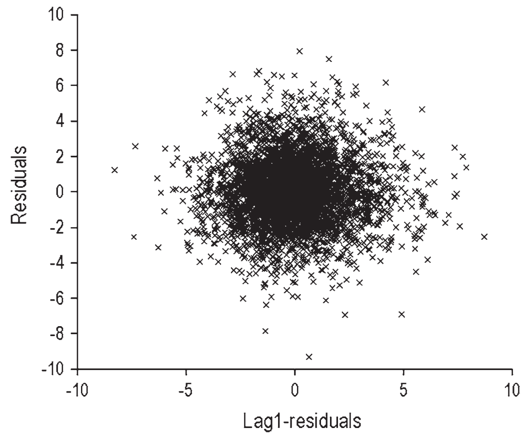

Figure 4. Residuals plotted against lag1 residuals for model M5 for all the species combined, fitted by the non-linear mixed-effects approach.

Comparison of the observed, calibrated estimates and estimates calibrated by adding an unstructured stochastic component height distribution per diameter class showed substantial differences between the three approaches, although the mean height values per diameter class obtained from all the approaches were very similar (Figure 5).

The observed height distribution was generally normal within diameter classes, with a high degree of variability in the datasets evaluated. The calibrated estimates provided a height distribution with low variability located around the observed mean value, whereas the random stochastic approach provided greater variability, consistent with the observed distribution.

The results shown indicate that the stochastic approach enables the natural variability in heights within diameter classes to be mimicked and therefore provides more realistic height predictions at stand level. This feature is considered very important since *h-d* models are used to fill in the missing heights of those trees whose heights were not measured. Stand-level or individual tree growth models (e.g. Corral-Rivas *et al.*, 2007b) also use *h-d* functions to estimate the height of the representative tree for each diameter class and the height of individual trees, respectively, which are subsequently incorporated in a taper function to estimate the total or merchantable stand volume. Therefore,

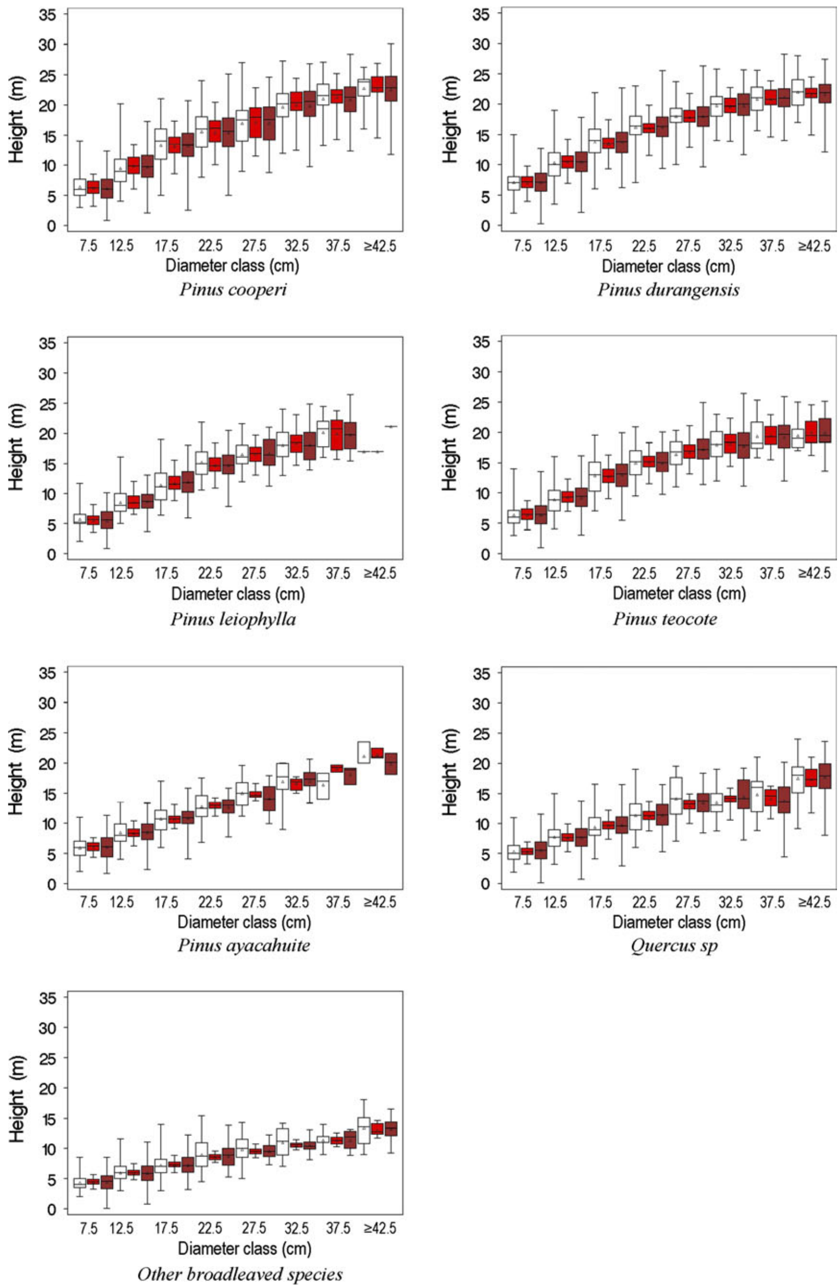

Figure 5. Box plots of height distributions (*y*-axis) against diameter classes (*x*-axis). The triangles represent the means of height estimates. The boxes represent the interquartile ranges. The maximum and minimum height estimates are represented by horizontal lines at the ends of the vertical lines. In white: observed distribution; in light grey: calibrated response predicted from one randomly selected tree per plot; in dark grey: calibrated response predicted from one randomly selected tree per plot with an unstructured stochastic component added.

the inclusion of a stochastic *h-d* relationship in this type of growth model enables more accurate predictions of stand volume to be obtained.

Conclusions

The study evaluates the contribution of stand variables to the estimation of individual tree heights in complex uneven-aged stands. The generalized *h-d* relationship provided by the model proposed by Sharma and Parton (2007) accounted for a high percentage of the variance in height, taking into account the structural complexity of these uneven-aged stands. Furthermore, this model was biologically reasonable beyond the range of the empirical observations used.

The model selected includes the d.b.h., the quadratic mean diameter and the dominant height as independent variables. Therefore, the variables needed to use the model require a low sampling effort since it is only necessary to measure the diameters and a small sample of heights corresponding to the dominant trees.

Mixed model techniques were used to estimate fixed- and random-effects parameters for the *h-d* function by Sharma and Parton. Techniques for calibration of this generalized model were used to obtain *h-d* functions tailored to individual units (each species or group of species within a sample plot) considering only the height of one randomly selected tree per plot.

An unstructured stochastic component was added to the mixed-effects model developed. Stochastic height predictions were tested with real observations and it was concluded that the model predictions performed acceptably well. The suggested approach enabled the natural variability in heights within diameter classes to be mimicked and therefore provided more realistic height predictions at stand level. This feature is considered very important since the *h-d* model developed in the present study will be used to fill in the missing heights corresponding to trees that were not measured.

Funding

Mexican National Council for Science and Technology (CONACyT, Agreement No. 103645).

Acknowledgements

The authors thank the two anonymous referees and the editor of the journal for the suggestions that helped us to improve the manuscript. We also thank Dr Christine Francis for correcting the English grammar of the text.
Conflict of Interest Statement
None declared.

References

Arabatzis, A.A. and Burkhart, H.E. 1992 An evaluation of sampling methods and model forms for estimating height-diameter relationships in loblolly pine plantations. *For. Sci.* **38**, 192–198.

Barrio-Anta, M., Diéguez-Aranda, U., Castedo-Dorado, F., Álvarez, J.G. and Rojo, A.A. 2006 Mimicking natural variability in tree height using a stochastic height-diameter relationship. *N. Z. J. For. Sci.* **36**, 21–34.

Borders, B.E. 1989 System of equations in forest stand modelling. *For. Sci.* **35**, 548–556.

Burkhart, H.E. and Strub, M.R. 1974 A model for simulation of planted loblolly pine stands. In *Growth Models for Tree and Stand Simulation*. J. Fries (ed). Royal College of Forestry, Stockholm, Sweden, pp. 128–135.

Burnham, K.P. and Anderson, D.R. 1998 *Model Selection and Inference: A Practical Information-Theoretic Approach*. Springer-Verlag, New York.

Calama, R. and Montero, G. 2004 Interregional nonlinear height-diameter model with random coefficients for stone pine in Spain. *Can. J. For. Res.* **34**, 150–163.

Canham, C.D., Finzi, A.C., Pacala, S.W. and Burbank, D.H. 1994 Causes and consequences of resource heterogeneity in forests: interspecific variation in light transmission by canopy trees. *Can. J. For. Res.* **24**, 337–349.

Castedo-Dorado, F., Barrio, M., Parresol, B.R. and Álvarez, J.G. 2005 A stochastic height-diameter model for maritime pine ecoregions in Galicia (northwestern Spain). *Ann. For. Sci.* **62**, 455–465.

Castedo-Dorado, F., Diéguez-Aranda, U., Barrio, M., Sánchez, M. and von Gadow, K. 2006 A generalized height-diameter model including random components for radiata pine plantations in northeastern Spain. *For. Ecol. Manage.* **229**, 202–213.

Colbert, K.C., Larsen, D.R. and Lootens, J.R. 2002 Height-diameter equations for thirteen midwestern bottomland harwood species. *North. J. Appl. For.* **19**, 171–176.

Corral-Rivas, J.J., Alvarez-González, J.G., Ruiz-González, A.D. and von Gadow, K. 2004 Compatible height and site index models for five pine species in El Salto, Durango (Mexico). *For. Ecol. Manage.* **201**, 145–160.

Corral-Rivas, J.J., Barrio-Anta, M., Aguirre-Calderon, O.A. and Diéguez-Aranda, U. 2007a Use of stump diameter to estimate diameter at breast height and tree volume for major pine species in El Salto, Durango (Mexico). *Forestry*. **80**, 29–40.

Corral-Rivas, J.J., Diéguez-Aranda, U., Corral-Rivas, S. and Castedo-Dorado, F. 2007b A merchantable volume system for major pine species in El Salto, Durango (Mexico). *For. Ecol. Manage.* **238**, 118–129.

Daniels, R.F. and Burkhart, H.E. 1975 *Simulation of Individual Tree Growth and Stand Development in Managed Loblolly Pine Plantations*. FWS 5–75. Virginia Polytechnic Institute and State University, Blacksburg, VA.

De Leeuw, J. 1992 Introduction to Akaike (1973) information theory and an extension of the maximum likelihood principle. In *Breakthroughs in Statistics*. S. Kotz and N.L. Johnson (eds). **Vol. 1**, Springer-Verlag, London, pp. 599–609.

Dennis, B., Brown, B.E., Stage, A.R., Burkhart, H.E. and Clark, S. 1985 Problems of modeling growth and yield of renewable resources. *Am. Stat.* **39**, 374–383.

Diéguez-Aranda, U., Barrio, M., Castedo, F. and Álvarez, J.G. 2005 Relación altura-diámetro generalizada para masas de Pinus sylvestris L. procedentes de repoblación en el noroeste de España. *Invest. Agrar. Sist. Recur. For.* **14**, 229–241.

Eerikäinen, K. 2003 Predicting the height-diameter pattern of planted *Pinus kesiya* stands in Zambia and Zimbabwe. *For. Ecol. Manage.* **175**, 355–366.

Fang, Z. and Bailey, R.L. 2001 Nonlinear mixed effects modelling for slash pine dominant height growth following intensive silvicultural treatments. *For. Sci.* **47**, 287–300.

Fox, J.C., Ades, P.K. and Bi, H. 2001 Stochastic structure and individual-tree growth models. *For. Ecol. Manage.* **154**, 261–276.

Gonda, H.E., Maguire, D.A., Cortés, G.O. and Tesch, S.D. 2004 Stand-level height-diameter equations for young ponderosa pine plantations in Neuquén, Patagonia, Argentina: evaluating applications of equations developed in the Western United States. *West. J. Appl. For.* **19**, 202–210.

Gregoire, T.G., Schabenberger, O. and Barrett, J.P. 1995 Linear modeling of irregularly spaced, unbalanced, longitudinal data from permanent plot measurements. *Can. J. For. Res.* **25**, 136–156.

Hirsch, R.P. 1991 Validation samples. *Biometrics*. **47**, 1193–1194.

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

Huang, S. and Titus, S.J. 1994 An age-independent individual tree height prediction model for boreal spruce-aspen stands in Alberta. *Can. J. For. Res.* **24**, 1295–1301.

Huang, S., Titus, S.J. and Wiens, D.P. 1992 Comparison of nonlinear height-diameter functions for major Alberta tree species. *Can. J. For. Res.* **22**, 1297–1304.

Hui, G.Y. and von Gadow, K. 1993 Zur Entwicklung von Einheitshöhenkurven am Beispel der Baumart Cunninghamia lanceolata. *Allg. Forst. Jagdztg.* **164** (12), 218–220.

Hurvich, C.M., Simonoff, J.S. and Tsai, C.L. 1998 Smoothing parameter selection in nonparametric regression using an improved Akaike information criterion. *J. R. Stat. Soc. B*. **60**, 271–293.

King, D.A. 1990 The adaptive significance of tree height. *Am. Nat.* **135**, 809–828.

Kozak, A. and Kozak, R. 2003 Does cross-validation provide additional information in the evaluation of the regression models?. *Can. J. For. Res.* **33**, 976–987.

Lappi, J. 1997 A longitudinal analysis of height/diameter curves. *For. Sci.* **43**, 555–570.

Larsen, D.R. and Hann, D.W. 1987 *Height-Diameter Equations for Seventeen Tree Species in Southwest Oregon*. Oregon State University, Forest Research Laboratory, Corvallis, Oregon. Research Paper 49.

Lei, Y. and Parresol, B.R. 2001 . *Remarks on height-diameter modelling*. Research Note SRS-10. USDA Forest Service, Southern Research Station, Asheville, NC.

Lindstrom, M.J. and Bates, D.M. 1990 Nonlinear mixed effects models for repeated measures data. *Biometrics*. **46**, 673–687.

López-Sánchez, C.A., Gorgoso, J.J., Castedo, F., Rojo, A., Rodríguez, R. and Álvarez González, J.G. *et al.* 2003 A height-diameter model for *Pinus radiata* D. Don in Galicia (Northwest Spain). *Ann. For. Sci.* **60**, 237–245.

Mirkovic, D. 1958 *Normale visinske krive za chrast kitnak i bukvu v NR Srbiji*. Zagreb: Glasnik sumarskog fakulteta 13. Cited in: Wenk, G., Antanaitis V., Smelko, S. 1990. Waldertragslehre. Deutscher Landwirtschaftsverlag, Berlin, 448 pp.

Nagel, J., Albert, M. and Schmidt, M. 2002 Das waldbaulich Prognose- und Entscheidungsmodell BWINPro 6.1. Neuparametrisierung und Modellerweiterungen. *Forst. Holz.* **57**, 486–492.

Neter, J., Kutner, M.H., Nachtsheim, C.J. and Wasserman, W. 1996 *Applied Linear Statistical Models*. 4th edn. McGraw-Hill, New York.

Oliver, C.D. and Larson, B.C. 1990 *Forest Stand Dynamics*. McGraw Hill, New York.

Peng, C., Zhang, L. and Liu, J. 2001 Developing and validating nonlinear height-diameter models for major tree species of Ontario's boreal forest. *North. J. Appl. For.* **18**, 87–94.

Peng, C., Zhang, L., Zhou, X. and Dang, Q. 2004 Developing and evaluating tree height-diameter models at three geographical scales for black spruce in Ontario. *North. J. Appl. For.* **21**, 83–92.

Pinheiro, J.C. and Bates, D.M. 1998 *Model Building for Nonlinear Mixed Effects Model*. Department of Statistics, University of Wisconsin, Madison, WI.

SAS Institute Inc. 2004 *SAS Online Doc© 9.1.2*. SAS Institute Inc, Cary, NC.

Saunders, M.R. and Wagner, R.G. 2008 Long-term spatial and structural dynamics in Acadian mixedwood stands managed under various silvicultural systems. *Can. J. For. Res.* **38**, 498–517.

Schröder, J. and Álvarez González, J.G. 2001 Comparing the performance of generalized diameter-height equations for maritime pine in Northwestern Spain. *Forstwiss. Centralbl.* **120**, 18–23.

Sharma, M. and Parton, J. 2007 Height-diameter equations for boreal tree species in Ontario using a mixed-effects modeling approach. *For. Ecol. Manage.* **249**, 187–198.

Soares, P. and Tomé, M. 2002 Height-diameter equation for first rotation eucalypt plantations in Portugal. *For. Ecol. Manage.* **166**, 99–109.

Stage, A.R. 1973 *Prognosis model for stand development*. Res. Pap. INT-137. USDA. Forest Service. Intermountain Forest and Range Experiment Station. Ogden, Utah.

Stage, A.R. 2003 How forest models are connected to reality: evaluation criteria for their use in decision support. *Can. J. For. Res.* **33**, 410–421.

Stage, A.R. and Wykoff, W.R. 1993 Calibrating a model of stochastic effects on diameter increment for individual-tree simulations of stand dynamics. *For. Sci.* **39**, 692–705.

Temesgen, H. and von Gadow, K. 2004 Generalized height-diameter models – an application for major tree species in complex stands of interior British Columbia. *Eur. J. For. Res.* **123**, 45–51.

Trincado, G. and Leal, D.C. 2006 Ecuaciones locales y generalizadas de altura-diámetro para pino radiata (*Pinus radiata*). *Bosque (Valdivia)*. **27**, 23–34.

Trincado, G., VanderSchaaf, C.L. and Burkhart, H.E. 2007 Regional mixed-effects height-diameter models for loblolly pine (Pinus taeda L.) plantations. *Eur. J. For. Res.* **126**, 253–262.

van Laar, A. and Akça, A. 2007 Forest mensuration. *Managing Forest Ecosystems*. **Vol. 13**, Springer (ed). Dordrecht, The Netherlands, 383 pp.

Vargas-Larreta, L.B. 2006 *Analyse und Prognose des Einzelbaumwachstums in Strukturreichen Mischbeständen in Durango Mexiko*. Dissertation. Universität Göttingen, Germany, 173 pp.

Vonesh, E.F. and Chinchilli, V.M. 1997 *Linear and Nonlinear Models for the Analysis of Repeated Measurements*. Marcel Dekker, Inc., New York.

von Gadow, K., Real, P. and Álvarez-González, J.G. 2001 Modelización del crecimiento y la evolución de los bosques. *IUFRO World Series vol. 12*. Vienna, Austria.

Wiedemann, E. 1936 Über die Vereifachung der Höhenermittlung bei der Vorratsaufnahme. *Mitt. Forstwirtsch. Forstwiss.* **7**, 387–412.

Wykoff, W.R., Crookston, N.L. and Stage, A.R. 1982 *User's Guide to the Stand Prognosis Model*. Gen. Tech. Rep. INT-133. USDA, Forest Service. Intermountain Forest and Range Experiment Station. Ogden, Utah.

Yang, Y., Monserud, R. and Huang, S. 2004 An evaluation of diagnostic tests and their roles in validating forest biometric models. *Can. J. For. Res.* **34**, 619–629.

Zeide, B. and Vanderschaaf, C. 2002 The effect of density on the height-diameter relationship. In *Proceedings of the 11th Biennial Southern Silvicultural Research Conference*. W. Outcalt Kenneth (ed). General Technical Report SRS-48. USDA, Forest Service. Southern Research Station, Asheville, NC, pp. 453–456.

Zhang, L. 1997 Cross-validation of non-linear growth functions for modeling tree height-diameter relationships. *Ann. Bot.* **79**, 251–257.

Zhang, S., Amateis, R. and Burkhart, H.E. 1997 The influence of thinning on tree height and diameter relationships in loblolly pine plantations. *South. J. Appl. For.* **21**, 199–205.

Received 8 October 2008