

Can random components explain differences in the height–diameter relationship in mixed uneven-aged stands?

Felipe Crecente-Campo · José Javier Corral-Rivas ·
Benedicto Vargas-Larreta · Christian Wehenkel

Received: 19 October 2012 / Accepted: 1 October 2013 / Published online: 16 October 2013
© INRA and Springer-Verlag France 2013

Abstract

• **Context** Tree height prediction is an important issue in forest management since tree heights are usually measured only in a sample of trees. Although numerous model approaches have been used for this purpose, no agreement on which one is more appropriate has been achieved.

Handling Editor: Barry Alan Gardiner

Contribution of the co-authors Felipe Crecente-Campo: analysing data, writing the manuscript.

José Javier Corral-Rivas: designing the experiment, field data collection, supervising the work, coordinating the research project.

Benedicto Vargas-Larreta: designing the experiment, field data collection, supervising the work.

Christian Wehenkel: designing the experiment, field data collection, supervising the work.

F. Crecente-Campo (✉)

Centro de Investigación Forestal (CIFOR), Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA),
Ctra. de la Coruña, km 7.5, 28040 Madrid, Spain
e-mail: felipe.crecente@gmail.com

J. J. Corral-Rivas

Facultad de Ciencias Forestales, Universidad Juárez del Estado de Durango, Río Papaloapan y Blvd. Durango s/n, Col. Valle del Sur,
34120 Durango, México
e-mail: jcorral@ujed.mx

B. Vargas-Larreta

Instituto Tecnológico de El Salto, Mesa del Tecnológico,
34942 El Salto, Durango, México
e-mail: bvargas@itf.edu.mx

C. Wehenkel

Instituto de Silvicultura e Industria de la Madera, Universidad Juárez del Estado de Durango, Carretera Mazatlán, km 5.5, 34120 Durango, México
e-mail: wehenkel@ujed.mx

• **Aims** To analyse the random effects of basic and generalised height–diameter ($h-d$) models fitted to multi-species uneven-aged forest stands, and to establish their ability to explain differences between ecoregions, plots and species.

• **Methods** Height and diameter measurements for 29,084 trees from 187 sample plots located in the state of Durango (Mexico) were used. Basic and generalised $h-d$ models were fitted in a mixed-models framework. The variability between ecoregions, plots and species was considered in the random effects definition. Model calibration for different height sampling designs and sampling sizes was also analysed.

• **Results** Random components performed well in explaining the differences in the $h-d$ relationship between the different plots and species; however, no significant variance for the random effects was found for the different ecoregions. A calibrated basic $h-d$ model produced similar results to a fixed-effects generalised $h-d$ model when a sufficiently large number of trees was used in the calibration process.

• **Conclusion** From a practical point of view, if no calibration is carried out, different models should be used for the different species, so that at least the variation among species is captured.

Keywords Ecoregion · Generalised height–diameter models · Mixed models · Calibration

1 Introduction

The forests of Durango (Mexico) cover an area of 4.9 million hectares, and timber resources account for about one quarter, in terms of volume, of the national forest resources in Mexico. Durango State is the top producer of pine

growing stock (27.7 %), coniferous roundwood (35.3 %) and other roundwood (37.6 %) nationwide (SRNyMA 2006). The pine-oak forests of Durango have been ranked as the most important forests in Mexico because of their extent and economic value. These forests are of particular interest, not only because they represent a unique ecosystem, but also because they are owned and managed by local communities, known as *Ejidors* (Thoms and Betters 1998).

Most of the forests in Durango are irregular, and conifer species occur as mixtures with hardwood species. This irregularity refers to the spatial arrangement of trees (vertical and horizontal irregularity) and the variation in the age structure of trees and stands. This structure is the result of the management history, which has depended on land ownership, as well as the economic and social changes that have taken place in the state, and also on natural conditions (Wehenkel et al. 2011).

The management of such mixed uneven-aged forests is more complex than the management of even-aged ones, and one of the most pressing research problems is the development of growth models to define sustainable harvests. However, few studies have attempted to estimate tree growth and stand development in Durango (e.g. Corral-Rivas et al. 2004; Vargas-Larreta et al. 2009).

In this sense, height–diameter (h – d) models (i.e. models relating individual tree height with individual tree diameter) are very useful for stand growth dynamics, yield, site index and dominant height estimation (Curtis 1967), stand structural analysis (Morrison et al. 1992), damage appraisal and stand stability (Parresol 1992), and product recovery and carbon budgeting models (Newton and Amponsah 2007).

In general, most height–diameter (h – d) models have been applied to pure even-aged stands or plantations (e.g. Soares and Tomé 2002; López Sánchez et al. 2003). Despite the homogeneous characteristics of this type of forest, a single h – d model is not usually adequate for all possible situations within a stand because the height curve does not remain constant and increases in steepness with age. Height curves for good quality sites have steeper slopes than those for poor quality sites, and for a particular height, trees growing in high density stands will have smaller diameters than those growing in less dense stands (Curtis 1967; Bailey and Brooks 1994; Lappi 1997; Zhang et al. 1997). This is even more evident in mixed and uneven-aged stands, in which different species, ages, sizes, crown types and levels of shade tolerance coexist (Vargas-Larreta et al. 2009).

To account for this level of variance, basic h – d relationships can be improved by taking into account stand variables that introduce the dynamics of each stand into the model

(Curtis 1967). Such models, called generalised h – d models, have been shown to be applicable to both even-aged (e.g. Soares and Tomé 2002; Castedo-Dorado et al. 2006) and uneven-aged stands (e.g. Sharma and Zhang 2004; Sharma and Parton 2007), and even to a mixture of both types of stands (e.g. Crecente-Campo et al. 2010). These models use diameter and stand-specific variables as regressors, accounting for the differences in the h – d relationship both within stands and over time (Curtis 1967; Soares and Tomé 2002; Sharma and Zhang 2004).

The hierarchical structure of the h – d data (i.e. trees within plots and plots within stands) usually results in a lack of independence among measurements, since observations from the same sampling unit are highly correlated (West et al. 1984). Development of the mixed-modelling methodology provided a statistical method capable of explicitly modelling this nested stochastic structure (Lappi 1997; Calama and Montero 2004; Castedo-Dorado et al. 2006). Mixed models are composed of a fixed functional part, common to the population, and random components acting at each sampling level. These models allow description of the variability in given phenomena among different sampling levels after defining a common fixed functional structure and a covariance structure for the random effects and residual terms (Lindstrom and Bates 1990; Calama and Montero 2005). Detailed information on nonlinear mixed-effects modelling for h – d relationships are provided by Calama and Montero (2004). General information and discussion on non-linear mixed models in the forestry context can be found in Hall and Bailey (2001). Finally, the multilevel case of mixed models has been discussed by several authors in a general context (e.g. Lindstrom and Bates 1990; Longford 1993; Goldstein 1995).

Mixed models may also improve the predictions obtained if it is possible to estimate the value of the random effects for an individual that has not been sampled. This approach is known as localisation or calibration and can be applied if supplementary observations of the dependent variable (total tree height in this case) are available (Lappi 1991; Jayaraman and Lappi 2001; Lynch et al. 2005; Calama and Montero 2005).

The h – d relationship is an important component of growth models. This model plus a compatible volume system (composed of a taper equation and a disaggregation equation) allows for stand volume classification by merchantable sizes, which are important tools for sustainable management of the species in the study area.

Generalised h – d models for uneven-aged stands were developed in a previous study for the region of El Salto (Durango, Mexico) (Vargas-Larreta et al. 2009). However, this previous study was very local, and results cannot be

extrapolated to the state of Durango. Moreover, the authors did not consider the different hierarchical levels (ecoregions, plots and species) in the random effects definition, and comparisons with non-generalised models were not made.

The objective of the present study was to fit basic and generalised $h-d$ models in a mixed-models framework and to analyse the random effects estimates of such models, in order to establish their ability to explain the differences in the $h-d$ relationships between ecoregions, plots and species observed in mixed uneven-aged stands in Durango, México. Models for practical use are planned for development on the basis of these analyses; therefore, the impact on predictability of the model of having a subsample of complementary trees where both variables (diameter and height) are measured was also analysed for different height sampling designs and sampling sizes within each sampling level.

2 Materials and methods

2.1 Study area and data description

The forests of Durango State are located in the Sierra Madre Occidental, at altitudes ranging between 363 and 3,190 m above sea level. The climate in the area is temperate to tropical, with most rainfall occurring in the summer, annual average precipitation of 443–1,452 mm and annual average temperature of 8.2–26.2 °C (Silva-Flores et al., submitted). The forests are mainly managed by selective removals, with only a small number of shelterwood harvests (3 % of the productive forest area). Clearfellings, which require special management expertise, are rarely applied (Wehenkel et al. 2011).

The data used in the study were obtained from a network of permanent sample plots used to monitor the growth and yield of Durango's forests. The plots used in this study were established between 2007 and 2010 and cover the main forest types and the current diameter distributions of commercial forests in Durango. The plots are 50×50 m in size and are distributed by systematic sampling (with some exceptions), with a variable grid ranging from 3 to 5 km, depending on the size of the *Ejidors*. The sampling plots are planned for re-measurement at 5-year intervals. Among other variables, tag number, species code, breast height diameter (d , centimetres, 1.3 m above ground level), total tree height (h , metres), height to the live crown (metres), azimuth (degrees Centigrade) and radius (metres) from the centre of the plot of all trees equal or larger than 7.5 cm in diameter were recorded. Currently, the database

includes measurement data for 29,084 trees from 187 sample plots.

These sample plots were located between latitudes 23° 10' N and 25° 20' N, and between longitudes 104° 45' W and 106° 40' W (Fig. 1). They were located at altitudes ranging from 2,006 to 2,939 m above sea level. The soils were very variable, with sand (particles over 2 mm) percentages varying from 18 to 85 % (average value of 52 %), silt (particles between 0.2 and 2 mm) percentages varying from 1 to 87 % (average value of 25 %) and clay (particles below 2 mm) percentages varying from 11 to 42 % (average value of 23 %). The number of stems per plot varied from 16 to 498 (average value of 156), and the number of species per plot varied from 2 to 12 (average value of 7).

Four different ecoregions, designated ecoregions I to IV, were included in the dataset: I (El Salto), II (San Dimas), III (Santiago y Topia) and IV (Tepehuanes) (Fig. 1). These ecoregions were defined on the basis of regional organisations previously designated by the National Forest Commission (CONAFOR), as Regional Forest Management Units (*Unidades de Manejo Forestal Regional* (UMAFORs)), which tend to focus on environmental services as a public policy goal. UMAFORs are defined as “representing” 30 % of the forest owners within a particular region, usually delineated along ecological zones, and thus include cross-community interdependence in a biological sense (Camille and García-López 2008).

Sixteen different species or groups of species (hereafter simply referred as species) were distinguished for posterior analysis as they presented similar growth patterns:

- 1: *Pinus arizonica*
- 2: *Pinus ayacahuite*
- 3: *Pinus cooperi*
- 4: *Pinus durangensis*
- 5: *Pinus engelmannii*
- 6: *Pinus herrerae*
- 7: *Pinus lumholtzii*
- 8: *Pinus teocote*
- 9: other pines: *Pinus chihuahuana*, *Pinus douglasiana*, *Pinus michoacana*, *Pinus oocarpa*, *Pinus tenuifolia*
- 10: *Quercus sideroxyla*
- 11: other *Quercus*: *Quercus candicans*, *Quercus coccolobifolia*, *Quercus conzattii*, *Quercus cordifolia*, *Quercus crassifolia*, *Quercus durifolia*, *Quercus fulva*, *Quercus grisea*, *Quercus microphylla*, *Quercus obtusata*, *Quercus resinosa*, *Quercus rugosa*, *Quercus rysophylla*, *Quercus urbanii*
- 12: *Abies durangensis*, *Picea chihuahuana*, *Pseudotsuga menziesii*

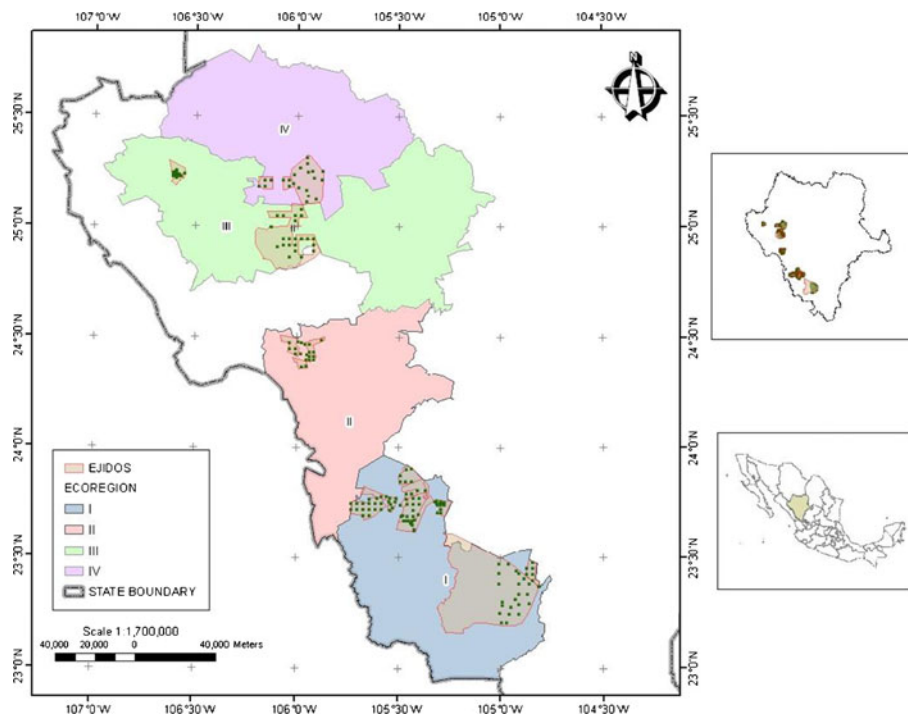


Fig. 1 Location, in the state of Durango (Mexico), of the research plots installed in the *Ejidots* in the four ecoregions analysed in this study: *I* (El Salto), *II* (San Dimas), *III* (Santiago y Topia) and *IV* (Tepehuanes)

- 13: *Cupressus lindleyi*
- 14: *Juniperus depeana*
- 15: *Alnus acuminata*
- 16: *Arbutus xalapensis*

The following stand variables were calculated from the tree data recorded in each plot: stems per hectare (N , trees per hectare), stand basal area (G , square metres per hectare), quadratic mean diameter (dg , centimetres), dominant height (h_{dom} , metres) and dominant diameter ($ddom$, centimetres). The last two variables were calculated from the proportion of the 100 trees with the largest diameter per hectare. Some of the trees with no species code, and those trees with broken or dead tops were then excluded from further analysis. Data from trees with broken or dead tops were not used to calculate h_{dom} .

The scatter plots of total tree height against diameter at breast height for each species were visually examined to detect possible anomalies in the data. Extreme data points (i.e. outside the overall picture of the $h-d$ pairs) were not observed. Some unusual data points (i.e. trees with quite large height and small diameters) were observed, but since we are attempting to develop a model for uneven-aged mixed forest (with a high variability), the unusual data points may be very valuable signals on $h-d$ relationships rather than outliers. We therefore decided to keep these points in model fitting and to analyse them via the residuals.

The database (Table 1) was finally randomly divided into a fitting part (80 % of the plots) and an evaluation part (20 % of the plots).

2.2 Models analysed

In a first step, 27 basic $h-d$ models (h as a function of only d) from Huang et al. (2000) were fitted to the dataset. In a second step, a total of 30 generalised $h-d$ models selected from previous studies (López Sánchez et al. 2003; Sharma and Zhang 2004; Castedo-Dorado et al. 2006; Sharma and Parton 2007) were fitted to the dataset. Some modifications to the models were also tested (i.e. dg and mean height were replaced by $ddom$ and h_{dom} , respectively). Models that require age or site index were not used or were modified to exclude these variables, as the stands were uneven-aged.

2.3 Model fitting and comparison

To make an initial selection from the above-mentioned models, they were fitted for the different species, by the *ordinary nonlinear least squares* (ONLS) method with the MODEL procedure of SAS/ETS® (SAS Institute Inc. 2008). Statistical and graphical analyses were used to compare the performance of the models. Four statistical criteria obtained from the residuals were examined: the root mean square error (RMSE), the model efficiency (EF) (similar to the coefficient of determination for

Table 1 Summary statistics for the fitting and the evaluation datasets

Variable	Sp.	Fitting dataset					Evaluation dataset				
		No. obs.	Mean	Min.	Max.	SD	No. obs.	Mean	Min.	Max.	SD
<i>d</i>	1	421	20.5	7.5	52.0	10.5	39	24.6	8.0	40.4	8.6
	2	992	16.7	6.5	57.0	8.2	189	15.7	6.2	91.3	10.4
	3	2,382	18.1	5.7	80.0	9.6	569	19.0	5.2	71.0	11.0
	4	6,143	17.8	6.6	99.5	10.4	1,919	16.8	6.6	62.0	9.5
	5	271	20.9	7.4	65.5	11.1	94	20.2	7.3	55.5	10.4
	6	601	22.3	7.5	78.5	13.1	132	26.6	8.0	67.5	12.3
	7	628	16.9	6.7	47.5	7.7	81	16.8	8.0	46.5	6.8
	8	1,842	17.8	6.9	68.0	9.7	866	14.8	7.0	55.4	7.7
	9	1,109	20.9	6.5	80.0	10.9	292	18.2	7.3	79.3	10.9
	10	3,576	17.8	6.0	120.0	11.9	800	18.4	7.1	90.7	11.2
	11	1,784	18.5	7.0	86.0	11.1	586	17.2	7.5	81.3	10.9
	12	289	25.2	5.0	103.1	18.5	55	22.1	5.9	57.0	14.2
	13	322	23.6	5.0	90.0	19.5	15	16.4	6.0	89.0	20.0
	14	1,000	14.3	5.7	104.0	7.7	255	14.8	5.0	62.0	8.5
	15	458	14.0	6.8	87.0	8.1	123	16.8	7.0	52.8	10.1
	16	921	14.9	6.7	65.8	7.7	330	14.9	7.0	70.0	9.1
<i>h</i>	1	421	12.7	2.8	29.3	5.6	39	17.4	4.8	25.9	6.3
	2	992	11.6	2.3	35.7	4.2	189	10.5	4.0	38.1	4.9
	3	2,382	12.6	2.5	38.8	5.5	569	12.5	3.1	36.5	5.9
	4	6,143	13.5	2.9	37.9	5.6	1,919	12.5	2.5	32.3	5.3
	5	271	14.1	2.9	31.4	5.8	94	13.4	2.9	26.5	5.0
	6	601	15.8	4.6	33.8	6.2	132	17.9	5.2	35.3	5.6
	7	628	10.1	1.6	27.0	4.8	81	10.0	3.2	19.8	3.5
	8	1,842	11.5	2.7	35.0	5.1	866	9.6	2.7	31.7	3.7
	9	1,109	13.9	2.7	45.3	5.9	292	11.6	2.7	25.1	4.9
	10	3,576	9.8	1.8	32.9	4.5	800	9.3	2.4	30.8	4.3
	11	1,784	9.1	1.9	40.1	4.2	586	9.1	2.1	28.8	4.2
	12	289	17.6	2.6	46.2	9.4	55	12.7	3.9	30.6	6.7
	13	322	13.2	2.3	34.4	8.1	15	7.2	3.7	22.7	4.6
	14	1,000	7.3	2.6	26.0	3.1	255	7.0	2.0	19.6	3.2
	15	458	7.5	2.4	29.8	3.3	123	7.3	2.1	24.3	3.8
	16	921	6.4	1.9	21.9	2.5	330	5.9	2.2	18.2	2.4
<i>N</i>		149	611	64	1992	312	38	668	196	1,616	310
<i>G</i>		149	21.1	3.3	58.4	9.3	38	20.9	7.4	36.3	7.2
<i>dg</i>		149	21.9	11.8	50.7	5.4	38	20.9	12.4	35.8	5.1
<i>hdom</i>		149	17.6	5.4	32.3	4.8	38	16.9	9.3	26.8	4.5
<i>ddom</i>		149	36.2	17.5	67.2	7.6	38	35.3	24.6	55.4	7.2

Species numbers are as described in the text

d (centimetres) diameter at breast height (1.3 m above ground level), *h* (metres) total tree height, *N* (trees per hectare) number of trees per hectare; *G* (square metres per hectare) stand basal area, *dg* (centimetres) quadratic mean diameter, *hdom* (metres) dominant height, *ddom* (metres) dominant diameter

linear regression), the mean bias (*E*) and Akaike’s information criterion (AIC) (Akaike 1974), summarised as follows:

$$RMSE = \sqrt{\frac{\sum_{l=1}^{l=n} (y_l - \hat{y}_l)^2}{n-p}} \tag{1}$$

$$EF = 1 - \frac{\sum_{l=1}^{l=n} (y_l - \hat{y}_l)^2}{\sum_{l=1}^{l=n} (y_l - \bar{y})^2} \tag{2}$$

$$E = \frac{\sum_{l=1}^{l=n} (y_l - \hat{y}_l)}{n} \tag{3}$$

$$AIC = n \ln \left(\frac{\sum_{l=1}^{l=n} (y_l - \hat{y}_l)^2}{n} \right) + 2(p + 1) \tag{4}$$

where y_l , \hat{y}_l and \bar{y} are the measured, estimated and average values of the dependent variable for the *l*th tree, respectively; *n* is the total number of observations used; *p* is the number of model parameters; and *ln* is the natural logarithm. Subscript *l*

was used to refer to tree for concordance with the mixed-models notation used later.

Model validation was not carried out because an independent dataset was not available, and other validation methods seldom provide any additional information compared with the fitting statistics (Kozak and Kozak 2003). However, the evaluation data (Table 1) were used in model calibration in order to provide informative results about the calibration process (see details later).

Once the best basic and generalised $h-d$ models were selected, a nonlinear mixed-effects modelling framework was used (e.g. Lappi 1997; Calama and Montero 2004; Castedo-Dorado et al. 2006), because there was high variability between ecoregions, plots and species. Basically, the parameter vector of the nonlinear model can be defined (Pinheiro and Bates 2000) as:

$$\Phi_{ijk} = \mathbf{A}_{ijk}\lambda + \mathbf{B}_{ijk}\mathbf{b}_{ijk} \quad (5)$$

where λ is the $p \times 1$ vector of fixed population parameters (where p is the number of fixed parameters in the model), \mathbf{b}_{ijk} is the $q \times 1$ vector of random effects associated with the i th ecoregion, the j th plot and the k th species (where q is the number of random effects in the model), and \mathbf{A}_{ijk} and \mathbf{B}_{ijk} are design matrices of size $r \times p$ and $r \times q$ (where r is the total number of parameters in the model) for the fixed parameters and random effects specific to each ecoregion, plot and species, respectively.

The SAS macro NLINMIX (Littell et al. 2006) was used to fit the models. This macro implements a two-step process for estimating the parameters. First, it performs a standard nonlinear regression by least squares, and in a second step it performs successive calls to the procedure of estimating parameters of linear mixed models on a linearised approximation function until convergence is achieved. The macro can use two methods of expansion: expansion around zero or the best linear unbiased predictor (*BLUP*), which is the expected value for the random effects (Beal and Sheiner 1982), or expansion around the empirical best linear unbiased predictor (*EBLUP*) of the random effects (Littell et al. 2006). Both approaches produce reliable estimates (Davidian and Giltinan 1993; Pinheiro and Bates 2000). The *EBLUP* method sometimes generates slightly better results but requires more computation time, is less stable, and is very sensitive to model specification (Hartford and Davidian 2000). Moreover, Vonesh (1996) showed that consistent estimates with the *EBLUP* expansion method could only be obtained if the number of individuals and observations per individual are infinite. For our data, most of the models with various combinations of random effects failed to converge when the *EBLUP* expansion method was used. Therefore, all results presented in this study are based on the expansion around zero method.

Different combinations of parameters were assumed to be mixed (composed of a fixed part, and a random part, specific to each ecoregion, plot and species). A variance components structure was selected for the covariance structure of the random effects, since convergence was easily achieved and better fitting statistics were obtained than with other random effects variance structures. This matrix contains the variance components in a diagonal structure. The significance of the variance of the random effects was tested using the *COVTEST* option in the *NLINMIX* macro, which displays asymptotic standard errors and Wald tests for covariance parameters (SAS Institute Inc. 2009).

Finally, heterocedasticity was analysed via the residuals. If necessary, a special structure for the within-plot variance-covariance matrix could be used to include weighting factors to balance error variance (Calama and Montero 2004).

2.4 Model calibration

An advantage of mixed-effects models is that, if a subsample of m tree heights is available, such data can be used to predict the random effects vector \mathbf{b}_{ijk} (i.e. to calibrate the model), with the following expression (Vonesh and Chinchilli 1997):

$$\hat{\mathbf{b}}_{ijk} \approx \hat{\mathbf{D}}\hat{\mathbf{Z}}_{ijk}^T \left(\hat{\mathbf{R}}_{ijk} + \hat{\mathbf{Z}}_{ijk}\hat{\mathbf{D}}\hat{\mathbf{Z}}_{ijk}^T \right)^{-1} \hat{\mathbf{e}}_{ijkl} \quad (6)$$

where $\hat{\mathbf{D}}$ is a $q \times q$ variance-covariance matrix for the among-ecoregion, among-plot and among-species variability, common to all plots and estimated in the general fitting of the model; $\hat{\mathbf{R}}_{ijk}$ is the $m \times m$ variance-covariance matrix for within-plot variability; $\hat{\mathbf{e}}_{ijkl}$ is the residual vector $m \times 1$, the components of which are given by the difference between the observed height of each tree included in the subsample, and the value predicted by the model including only fixed effects and $\hat{\mathbf{Z}}_{ijk}$ is the $m \times q$ matrix of partial derivatives with respect to the random effects evaluated at $\hat{\mathbf{b}}_{ijk}$.

For the best basic and generalised $h-d$ models, the calibrated response was evaluated using the evaluation dataset for different height sampling designs and sampling sizes within each ecoregion, plot and species. This dataset was used to show the calibrated results for plots that were not included in the fitting process. The alternatives selected according to previous studies (Calama and Montero 2004; Castedo-Dorado et al. 2006; Crecente-Campo et al. 2010) were:

1. Total height of one to four randomly selected trees per species and plot.
2. Total height of the one to four largest trees per species and plot.
3. Total height of the one to four smallest trees per species and plot.

4. Total height of the one to four medium-size trees per species and plot.
5. Total height of three trees per species and plot: the smallest, the largest and the medium-size trees.

Some combinations of plot and species with less than four trees were, therefore, eliminated, reducing the evaluation dataset from 6,345 to 6,225 trees in 38 plots, and resulting in a modal value of five species per plot.

The five calibration alternatives were evaluated in terms of the previously defined statistics (RMSE, *EF* and *E*) and compared with the estimates obtained by ONLS in the individual fit of each equation to each of the ecoregion, plot and species groups. For the randomly selected trees, mean values of the statistics after 100 simulations were obtained, as this was considered a large enough value to obtain representative results and was in accordance with previous studies (e.g. Castedo-Dorado et al. 2006; Crecente-Campo et al. 2010).

Finally, the selected models were also fitted independently for each species, and random effects specific to each ecoregion and plot were added, to determine whether the consideration of species as a fixed effect led to a significant improvement in the accuracy of the predictions.

3 Results

3.1 ONLS fits

Statistics for the fitting of the basic *h–d* equations (not shown) showed that the models of Bertalanffy–Richards (Bertalanffy 1949; Richards 1959), Weibull (Yang et al. 1978) and Schnute (Schnute 1981) were the most accurate, with very similar values for the fitting statistics. When the models were fitted to each species, the model of Schnute (1981) (Eq. 7) produced slightly better results (lower average AIC and RMSE values) and was, therefore, selected as a basic *h–d* model for posterior analysis:

$$h_l = \left(1.3^{a_0} + (a_1^{a_0} - 1.3^{a_0}) \frac{1 - \exp(-a_2 \cdot d_l)}{1 - \exp(-a_2 \cdot 100)} \right)^{\frac{1}{a_0}} \quad (7)$$

where h_l and d_l are, respectively, the total height and diameter at breast height of the l th tree and a_i are the parameters to be estimated.

From the total number of generalised *h–d* models analysed, a modification of the Bertalanffy–Richards (Bertalanffy 1949; Richards 1959) model (Eq. 8) showed the best results, as inferred from the fit statistics (Table 2). The same equation has previously been used by Sharma and Parton (2007) for several species in northern Ontario (Canada) and by Vargas-Larreta et al. (2009) for several species in El Salto (Durango, Mexico).

$$h_l = \left(1.3 + a_0 \cdot ddom_j^{a_1} \left(1 - \exp \left(-a_2 \cdot \left(\frac{N_j}{G_j} \right)^{a_3} d_l \right) \right) \right)^{a_4} \quad (8)$$

where subscript j refers to plot and subscript l refers to tree; the other variables are as previously defined.

3.2 Basic *h–d* mixed model

Independent fitting of Eq. 7 to each ecoregion, plot and species, by the ONLS method, showed that most of the variability in parameters was derived from the differences between plots and, secondly, species. The a_0 parameter showed the higher coefficient of variation for the different plots, while the a_2 parameter showed the higher coefficient of variation among species. Variability among ecoregions was small, as inferred from the differences between the values of the parameters for the different ecoregions. Finally, all possible combinations of random effects for each ecoregion, plot and species were tried. However, no significant variance for the ecoregion-specific random effects was found when fitting the model. The best results in the fitting of Eq. 7 as a mixed model were obtained when all the parameters were expanded to include plot-specific random effects, and some parameters were expanded to include species-specific random effects, resulting in the following mixed model (once parameters were scaled so that they were all around the same order of magnitude, to avoid instabilities in the fitting algorithm; Littell et al. 2006):

$$h_{jkl} = \left(1.3^{a_0 + v_{0j} + w_{0k}} + \left([20(a_1 + v_{1j} + w_{1k})]^{a_0 + v_{0j} + w_{0k}} - 1.3^{a_0 + v_{0j} + w_{0k}} \right) \frac{1 - \exp \left(-\frac{a_2 + v_{2j}}{50} \cdot d_{jkl} \right)}{1 - \exp \left(-\frac{a_2 + v_{2j}}{50} \cdot 100 \right)} \right)^{\frac{1}{a_0 + v_{0j} + w_{0k}}} \quad (9)$$

where subscript j refers to plot, subscript k refers to species and subscript l refers to tree; v_{ij} and w_{ik} are the random effects

specific to each plot and species, respectively, and the other variables are as previously defined.

Table 2 Statistics for the fitting of the basic and generalised mixed $h-d$ models

Equation	Type	RMSE	EF	E	AIC
(7)	Basic ONLS	3.767	0.5598	0.0024	60,317
(9)	Basic mixed	2.564	0.7960	0.5211	42,486
(11)	Basic mixed by sp	2.406	0.8203	-0.0462	39,959
(8)	Generalised ONLS	3.207	0.6809	0.0026	53,008
(10)	Generalised mixed	2.271	0.8400	0.1423	37,317
(12)	Generalised mixed by sp	2.225	0.8463	-0.0362	36,394

Results for Eqs. 11 and 12 are summarised for all the species but are not normalised by the number of trees of each species

Fit statistics for Eq. 9 (Table 2) showed an improvement over the basic ONLS model (Eq. 7). RMSE values were reduced by approximately 32 %, but E was greatly increased, although it kept at a reasonable level (approximately 0.5 m). This may suggest that differences have become biased either negative or positive. However, the plot of raw residuals against predicted values (Fig. 2a) showed no systematic trends and a homogeneous distribution. Trends in the residuals usually relate to missing covariates or heterogeneity of variance. Parameter estimates for Eq. 9 are shown in Table 3.

3.3 Generalised $h-d$ mixed model

The best results in the fitting of Eq. 8 as a mixed model were obtained when two parameters were expanded to include both plot and species-specific random effects. As previously, the ecoregion random effects variance was not significant. The resultant model (once parameters were scaled so that they were all around the same order of magnitude, to avoid instabilities in the fitting algorithm: Littell et al. 2006) was:

$$h_{jkl} = \left(1.3 + (a_0 + v_{0j} + w_{0k}) \cdot ddom_j^{a_1} \left(1 - \exp \left(-\frac{a_2}{50} \left(\frac{N_j}{G_j} \right)^{\frac{a_3}{4}} d_{jkl} \right) \right) \right)^{a_4 + v_{4j} + w_{4k}} \quad (10)$$

where all the variables are as previously defined.

$$h_{jkl} = \left(1.3^{a_{0k} + v_{0jk}} + \left([20(a_{1k} + v_{1jk})]^{a_{0k} + v_{0jk}} - 1.3^{a_{0k} + v_{0jk}} \right) \frac{1 - \exp \left(-\frac{a_{2k} + v_{2jk}}{50} \cdot d_{jl} \right)}{1 - \exp \left(-\frac{a_{2k} + v_{2jk}}{50} \cdot 100 \right)} \right)^{\frac{1}{a_{0k} + v_{0jk}}} \quad (11)$$

$$h_{jkl} = \left(1.3 + (a_{0k} + v_{0jk}) \cdot ddom_j^{a_{1k} + v_{1jk}} \left(1 - \exp \left(-\frac{a_{2k}}{50} \cdot \left(\frac{N_j}{G_j} \right)^{\frac{a_{3k}}{4}} d_{jl} \right) \right) \right)^{a_{4k} + v_{4jk}} \quad (12)$$

where subscript j refers to plot, subscript k refers to species and subscript l refers to tree; a_{ik} are the fixed-effects parameters for the k th species, v_{ijk} are the random effects specific to each j th plot and k th species and the other variables are as previously defined.

In this case, the estimate of the variance of the ecoregion random effects was again not significant. Fitting statistics

Fit statistics for Eq. 10 (Table 2) showed a reduction in RMSE by approximately 29 % as compared with the generalised ONLS model (Eq. 8). However, an increase in E was obtained, although it kept at a reasonable level (approximately 0.14 m). The plot of raw residuals against predicted values (Fig. 2b) showed no systematic trends and a homogeneous distribution. Parameter estimates for Eq. 10 are shown in Table 4.

While RMSE was only reduced by 11 % between Eqs. 9 and 10, E was reduced by 73 %. The differences between Eqs. 9 and 10 were also visually examined by species (Fig. 3). As confirmed, RMSEs were very similar, with only a slightly lower RMSE for Eq. 10. The greater RMSE differences were observed for species 9 and 12. E , however, was greater for all species for Eq. 9, particularly for species 6 and 12.

3.4 Fits to individual species

In an attempt to reduce the sources of variation, the variation between species was treated as a fixed effect (i.e. by fitting the models independently for each species and adding random effects for each plot and ecoregion), producing Eqs. 11 and 12.

were clearly superior to those obtained with the previous models, with a large reduction in E and a small reduction in RMSE (Table 2). Results for Eq. 12 were again only slightly better than those for Eq. 11 (Fig. 3). Parameter estimates for Eq. 11 are shown in Table 3, and those for Eq. 12 are shown in Table 4. In Eq. 11, the a_{1k} parameter (the model asymptote) showed clearly

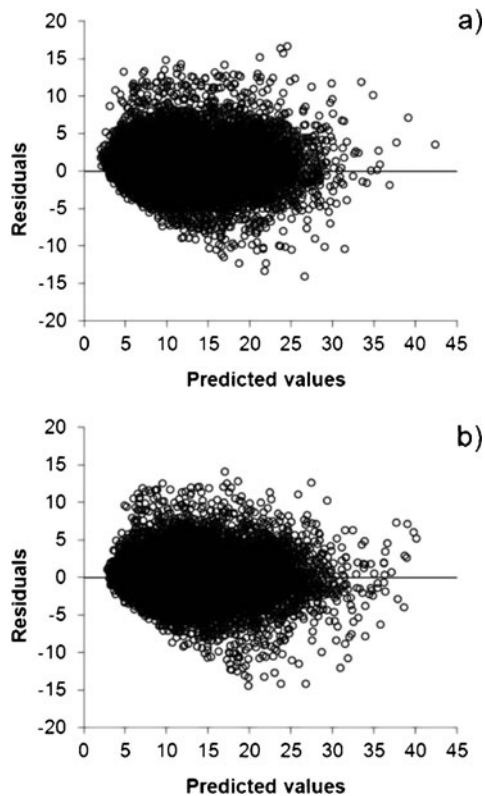


Fig. 2 Plots of raw residuals versus predicted values for Eqs. 9 (a) and 10 (b)

lower values for species 10, 11, 14, 15 and 16, which are the most shade tolerant species.

3.5 Model calibration

Results for the calibration process for all the mixed-models using the evaluation dataset showed small differences between Eqs. 9 and 10 and between Eqs. 11 and 12, using the random-trees selection method for calibration (Fig. 4). The results for all species together (Fig. 5) showed a similar trend, with quite different RMSEs for the fixed-effects model (i.e. the mixed model with 0 trees used in calibration) and quite similar RMSEs for the mixed-effects model using four trees for calibration. Similar results were found for EF and E (Fig. 5). The greatest reductions in RMSE between a fixed-effects model and a calibrated mixed-effects model were observed for the basic $h-d$ models, i.e. for Eqs. 9 and 11. As the evaluation dataset has a modal value of five species per plot, the calibrated results were aggregated in Figs. 4 and 5. So calibration with one tree per species means, a measurement of five trees per plot; calibration with two trees per species means a measurement of $5 \times 2 = 10$ trees per plot, and so on.

In a practical sense, when using Eq. 9, at least two trees of each species (i.e. a modal value of 10 trees per plot in the

evaluation dataset) must be measured to calibrate the model and to obtain the same results as the uncalibrated Eq. 12 (Fig. 5). For Eq. 11, the measurement of only one tree of each species per plot (a modal value of 5 trees per plot in the evaluation dataset) is enough to calibrate the model and obtain the same results as the uncalibrated Eq. 12. These general results are, however, influenced by the fact that Eq. 12 showed very poor results in the calibration process for species 12 and 13 (Fig. 4), maybe because there was a low number of observations included in the evaluation dataset (55 and 15 observations for each species, respectively) for these two species. For the rest of the species, Eq. 12 showed better calibrated results than Eq. 11 (Fig. 4).

The results obtained for the different calibration methods with Eq. 11 are shown in Fig. 6. This equation was selected as a reference for testing the different calibration methods because it produced comparable results to Eq. 12 for the total data set (Fig. 5) and needs fewer measurements to be applied. The medium-size tree selection method and the random tree selection method performed better than others in terms of RMSE and EF , followed very closely by the all-sizes tree selection method, which was superior in terms of E (Fig. 6). The other tree selection methods produced higher errors than the above-mentioned ones (Fig. 6).

4 Discussion

Several basic $h-d$ models were tested in this study for modelling the $h-d$ relationship in mixed uneven-aged stands in Durango (Mexico). Confirming the findings of Zhang (1997), the best models in this study were the Bertalanffy–Richards (Bertalanffy 1949; Richards 1959), Weibull (Yang et al. 1978) and Schnute (Schnute 1981) models. According to Lei and Parresol (2001), the Schnute function together with the Bertalanffy–Richards function are probably the most flexible and versatile functions available for modelling height–diameter relationships. Finally, the Schnute’s equation was selected for fitting as a basic $h-d$ mixed model because it showed superior fit statistics for almost all species groups.

From a total of 25 generalised $h-d$ models tested in this study, the equation that provided the best fit to the dataset was derived from the Bertalanffy–Richards model and included diameter at breast height, dominant height, stand basal area and trees per hectare in its formulation. The relationship of trees per hectare and dominant height to height growth has already been discussed by several authors (e.g. Calama and Montero 2004; Saunders and Wagner 2008; Vanclay 2009). Stand density (measured for example by stand basal area or trees per hectare) is the most important factor that affects the $h-d$ relationship

Table 3 Parameter estimates for the basic $h-d$ models selected

Model	Species	Parameter estimates								
		α_0	α_1	α_2	$\sigma_{wi,0}$	$\sigma_{wi,1}$	$\sigma_{vi,0}$	$\sigma_{vi,1}$	$\sigma_{vi,2}$	σ^2
Eq. 9	all	0.9761 (0.0481)	1.141 (0.072)	1.746 (0.108)	0.01090 (0.00476)	0.07487 (0.02759)	0.1206 (0.0220)	0.06554 (0.00913)	0.7793 (0.1530)	5.678 (0.054)
Eq. 11	1	0.2766 (0.1229)	1.127 (0.071)	3.585 (0.601)	–	–	0	0.05110 (0.01723)	0	5.308 (0.379)
	2	1.239 (0.041)	1.415 (0.038)	1* (0.601)	–	–	0.02074 (0.01038)	0.03518 (0.00977)	0	4.499 (0.214)
	3	0.9975 (0.0851)	1.353 (0.058)	1.803 (0.239)	–	–	0.05405 (0.02029)	0.07497 (0.01867)	0.2359 (0.1205)	4.968 (0.149)
	4	0.8965 (0.0529)	1.292 (0.030)	2.063 (0.151)	–	–	0.06456 (0.01375)	0.03652 (0.00696)	0.2031 (0.0756)	5.451 (0.101)
	5	0.4898 (0.2169)	1.197 (0.075)	3.188 (0.738)	–	–	0.04335 (0.02495)	0.0127 (0.0071)	0	6.886 (0.643)
	6	1.561 (0.160)	1.492 (0.071)	0.8689 (0.3436)	–	–	0	0	0.2100 (0.0805)	5.967 (0.350)
	7	0.6109 (0.1635)	1.137 (0.109)	2.506 (0.620)	–	–	0	0.06288 (0.02025)	0	4.461 (0.257)
	8	1.023 (0.092)	1.316 (0.063)	1.515 (0.257)	–	–	0.05052 (0.01421)	0.04562 (0.01043)	0	4.451 (0.152)
	9	0.8904 (0.1042)	1.309 (0.074)	1.834 (0.337)	–	–	0	0.04526 (0.01234)	0.2428 (0.0795)	5.321 (0.237)
	10	1.302 (0.088)	0.9409 (0.0338)	1.425 (0.202)	–	–	0	0.05652 (0.00842)	0	6.213 (0.149)
	11	0.8858 (0.1446)	0.8360 (0.0411)	2.423 (0.379)	–	–	0.1800 (0.0497)	0.05362 (0.01077)	0	4.909 (0.171)
	12	0.8037 (0.1044)	1.758 (0.091)	1.605 (0.259)	–	–	0	0.03758 (0.02075)	0	9.421 (0.803)
	13	1.038 (0.050)	1.213 (0.141)	1* (0.259)	–	–	0	0.2039 (0.0946)	0	8.258 (0.664)
	14	1.401 (0.068)	0.8523 (0.0359)	1* (0.259)	–	–	0	0.03514 (0.00696)	0	3.494 (0.164)
	15	1.290 (0.075)	0.9448 (0.0596)	1* (0.259)	–	–	0	0.06470 (0.01895)	0	3.060 (0.213)
	16	1.505 (0.086)	0.6993 (0.0284)	1* (0.259)	–	–	0	0.01394 (0.00336)	0	3.260 (0.161)

σ_{wi} , variance of the species-specific random effects; σ_{vi} , variance of the plot-specific random effects; σ^2 , error variance of the model; the rest of variables as previously defined. All parameters were significant at $p < 0.05$ except those marked with an asterisk; standard error in brackets

* $p < 0.05$: These parameters were not significant at this level and were fixed at 1 in the final model fitting. This produces slightly different models for each species; a value of 0 for the variance of a random effect indicates that this variance was not significant. A hyphen indicates that this parameter was not part of the model

in a stand (Zhang et al. 1997; Saunders and Wagner 2008); in dense stands, trees of the same diameter are usually taller than those in less dense stands. Dominant height has also previously been included in generalised $h-d$ models by many authors (e.g. Castedo-Dorado et al. 2006;

Vargas-Larreta et al. 2009; Amaral Paulo et al. 2011). In the present study, models that included dominant height provided more accurate results than those including mean height. This may be advantageous, since fewer trees need to be measured to estimate dominant height than to estimate mean height, and

Table 4 Parameter estimates for the generalised $h-d$ models selected (all parameters were significant at $p < 0.05$; standard errors in brackets)

Model	Species	Parameter estimates										
		α_0	α_1	α_2	α_3	α_4	σ_{w0}	σ_{w1}	σ_{v0}	σ_{v1}	σ_{v1}^2	
Eq. 10	All	2.164 (0.240)	0.8079 (0.0314)	1.478 (0.173)	0.2722 (0.1225)	1.035 (0.030)	0.2681 (0.0988)	0.006324 (0.002826)	0.02413 (0.00385)	0	0.03081 (0.00418)	5.780 (0.055)
Eq. 12	1	0.7750 (0.3382)	1.197 (0.155)	1.538 (0.529)	0.8350 (0.3838)	1.641 (0.177)	–	–	0.004639 (0.002198)	0	0	5.317 (0.379)
	2	5.865 (1.900)	0.6090 (0.0839)	0.4614 (0.1492)	0.6581 (0.3003)	0.8085 (0.0646)	–	–	0.2047 (0.1106)	0	0.002967 (0.001716)	4.516 (0.212)
	3	1.886 (0.476)	0.8951 (0.0850)	0.8445 (0.0831)	1* (0.059)	1.042 (0.059)	–	–	0.05336 (0.01402)	0	0.03282 (0.00920)	5.039 (0.150)
	4	2.740 (0.349)	0.7499 (0.0427)	0.9867 (0.0518)	1* (0.039)	1.108 (0.039)	–	–	0.04699 (0.010071)	0	0.04040 (0.00708)	5.511 (0.101)
	5	8.185 (3.067)	0.3801 (0.1339)	1.181 (0.2812)	1* (0.248)	1.381 (0.248)	–	–	0.6424 (0.3537)	0	0.06776 (0.03832)	6.915 (0.654)
	6	8.139 (2.306)	0.3893 (0.0901)	4.147 (1.992)	–1.048 (0.522)	0.8855 (0.0969)	–	–	0	0	0.3381 (0.0151)	5.843 (0.344)
	7	3.210 (0.977)	0.7436 (0.1040)	0.8383 (0.2214)	1* (0.1435)	1.203 (0.1435)	–	–	0.2229 (0.0802)	0	0	4.442 (0.256)
	8	2.875 (0.497)	0.7890 (0.0570)	0.6567 (0.0932)	1* (0.0591)	1.011 (0.0591)	–	–	0.04494 (0.01718)	0	0.01929 (0.00506)	4.490 (0.145)
	9	2.804 (0.602)	0.7421 (0.0705)	1.206 (0.279)	0.7623 (0.2570)	1.188 (0.087)	–	–	0.05294 (0.01813)	0	0.02158 (.00884)	5.329 (0.2236)
	10	0.5128 (0.1351)	1.203 (0.090)	0.5791 (0.1959)	1.548 (0.391)	0.9945 (0.0583)	–	–	0.007057 (0.001453)	0	0.05564 (0.01137)	5.624 (0.137)
	11	0.5844 (0.1790)	1.134 (0.105)	0.7255 (0.2887)	1.592 (0.470)	1.164 (0.106)	–	–	0.008623 (0.002161)	0	0.1147 (0.0303)	4.868 (0.169)
	12	1.391 (0.606)	0.9641 (0.1325)	0.3954 (0.1212)	2.617 (0.4356)	1.407 (0.112)	–	–	0.007239 (0.003348)	0	0	8.492 (0.719)
	13	0.6033 (0.2821)	1.231 (0.141)	0.2048 (0.09188)	2.293 (0.536)	0.9332 (0.0747)	–	–	0	0	0	8.491 (0.697)
	14	1.663 (0.612)	0.8710 (0.0961)	0.3988 (0.1881)	0.9059 (0.4100)	0.7803 (0.0681)	–	–	0.05772 (0.01480)	0	0	3.464 (0.161)
	15	1* (0.0334)	0.9750 (0.0797)	0.7061 (0.0797)	1* (0.0797)	1* (0.0797)	–	–	0	0.008882 (0.002635)	0	3.083 (0.214)
	16	1.376 (0.454)	0.7622 (0.0996)	0.6975 (0.2609)	1* (0.1172)	0.8648 (0.1172)	–	–	0.3680 (0.0145)	0	0.01428 (0.00646)	3.153 (0.158)

All the variables are as previously defined

* $p < 0.05$: These parameters were not significant at this level and were fixed at 1 in the final model fitting. This produces slightly different models for each species. A value of 0 for the variance of a random effect indicates that this variance was not significant. A hyphen indicates that this parameter was not part of the model

this greater sampling effort may limit the future use of models using mean height (López Sánchez et al. 2003). This generalised $h-d$ model showed better fitting statistics than the basic $h-d$ model (Table 2).

However, when the basic and generalised $h-d$ models were fitted in a mixed-models framework, fit statistics differed

little between them (Fig. 3, Table 2), confirming the expectation that random effects specific to each plot and species will capture most of the observed variability between these sampling units, and that the incorporation of additional predictor variables (i.e. stand variables) in a basic $h-d$ model has a major effect on the ability of the fixed-effects model in

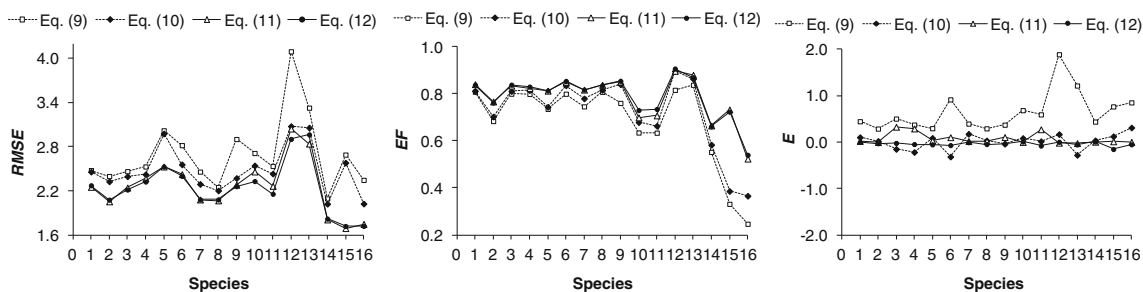


Fig. 3 Plots of the fitting statistics for Eqs. 9, 10, 11 and 12, for each species. *Lines* are used only for a better comparison of the differences between the models. *Numbers* on the *X*-axis refer to the species groups outlined in the text

explaining between-individual variability, but not on the mixed-effects model (Trincado et al. 2007). Another advantage usually assigned to mixed-models is that they take into account the correlations between observations that belong to the same sampling unit, thus dealing with the dependence of error terms which should be avoided in regression models (Calama and Montero 2004; Castedo-Dorado et al. 2006).

A result that seems somewhat counter intuitive is that bias increased for Eqs. 9 and 10 as compared with the respective *ONLS* Eqs. 7 and 8. This may be attributed to the fact that Eqs. 7 and 8 do not make any distinction between species and all the observations have the same importance in the calculation of fitting statistics. However, Eqs. 9 and 10 differentiate between species, assigning a random effect for each one, and in some cases, as with species 15 and 16, the explained variance was very low (Fig. 3) and the associated bias very high. The same happens for the bias obtained for species 6, 12 and 13 for Eq. 9, which showed large values, thereby increasing the overall bias (calculated over all species without normalising by the number of observations for each species).

Expansion of fixed parameters to account for differences between ecoregions led to non-significant variance of the random effects. This is not surprising, since only four ecoregions were sampled in this study. Several *t* tests were performed to compare the average values of the height-diameter ratios for the different ecoregions. These average values (i.e. 0.72, 0.68, 0.76 and 0.56 for ecoregions 1, 2, 3 and 4, respectively) were significantly different between the four ecoregions. However, group-level variance estimates of zero often arise when fitting multilevel or hierarchical linear models, especially when the number of groups is small (Chung et al. 2013). The same may be expected for non-linear models.

Equations 11 and 12, i.e. the species-specific ones, showed better precision (RMSE) and bias (*E*) than Eqs. 9 and 10, respectively (Table 2, Fig. 3), indicating that a general mixed-model for the whole population may not be sufficient to explain the observed variation among species. However,

the differences between the models were greatly reduced when they were calibrated with only one tree by species. Only a marginal gain in accuracy was observed when two, three or four sample trees per species were used for calibration (Fig. 5). Other studies have reported similar behaviour using linear or nonlinear mixed-effects *h-d* equations for other tree species (e.g. Jayaraman and Lappi 2001; Calama and Montero 2004; Trincado et al. 2007).

For the whole dataset (fitting plus evaluation), there was a modal value of six species present in each plot ($SD=1.75$). This means that if one tree per species was used for calibration, an average of six trees would have to be measured in each plot to calibrate Eq. 11. This number is substantially lower than the number of trees that should be measured to calculate dominant height (a variable included in Eq. 12) (25 dominant trees per plot in this study, as all the plots had an area of 2,500 m²). One problem is whether we should focus our efforts on measuring only the height of dominant trees, or measuring heights for all species. Foresters usually prefer to measure dominant height, since it is one of the key driving variables in growth and yield models and can, sometimes, be obtained as an output of growth simulators. However, if the purpose of the *h-d* relationship is only to know missing heights, the calibrated version of Eq. 11 showed better results in terms of prediction accuracy (as showed by *RMSE* and *E* values) and measurement effort. In any case, as the area of the sample plot decreases, the number of trees used to calculate dominant height also decreases, thus reducing the measurement effort required for the application of the fixed-effects of Eq. 12.

When fitting Eqs. 11 and 12, some parameters were not significant at the 0.05 level (Table 2). If these parameters were assigned a value of 0, then some variables would be eliminated from the model, resulting in a less accurate model. We therefore decided to keep these variables in the model and assign a value of 1 to the parameters (note that for parameters that are multiplicative or in the exponent of a power term, assigning a value of 1 means that this parameter is dropped

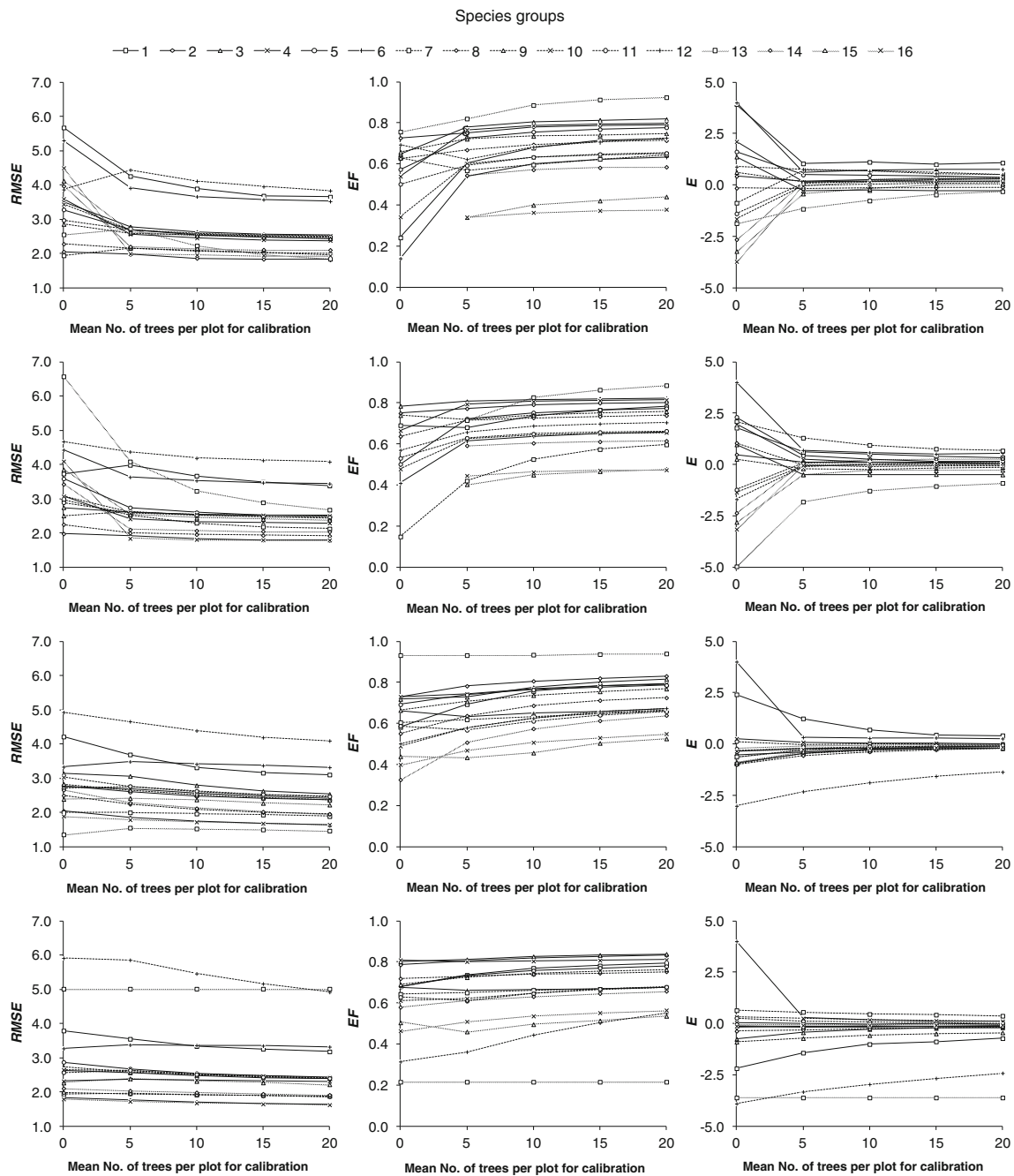


Fig. 4 Plots of the statistics obtained in the calibration process including the random tree selection method, for Eq. 9 (first row), Eq. 10 (second row), Eq. 11 (third row) and Eq. 12 (fourth row), for each species. Zero trees for calibration represent the fixed-effects model. The X-axis shows

aggregate values to an average number of trees per plot, because the evaluation dataset has a modal value of five species per plot. Numbers in the legend refer to the species groups outlined in the text

from the model). Therefore, slightly modified models were used in these cases. Convergence was achieved in all cases with the simpler models, and the accuracy of the models (shown by the *RMSE* and *E* values) was higher than that obtained by simply eliminating the parameters and their associated variables.

Calibration of mixed-effects models has been shown to be a good alternative for height prediction even when using a basic *h–d* model and only one height measurement per sampling level is available (e.g. Trincado et al. 2007), as in the current study. Other studies report similar results for height prediction (e.g. Calama and Montero 2004; Castedo-Dorado

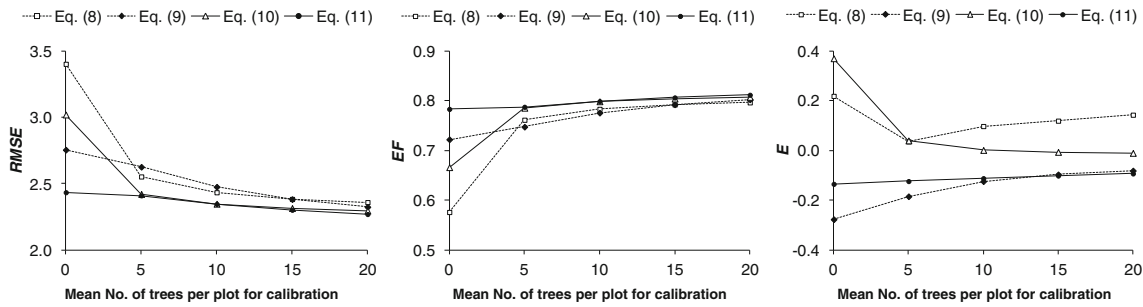


Fig. 5 Plots of the statistics obtained in the calibration process including the random tree selection method, for Eqs. 9, 10, 11 and 12, for all species together. Zero trees for calibration represent the fixed-effects model. The

X-axis shows aggregate values to an average number of trees per plot, because the evaluation dataset has a modal value of five species per plot

et al. 2006; Sharma and Parton 2007) as well as prediction of other growth variables (e.g. Calama and Montero 2005). Different calibration methods were tested for Eq. 11 in this study, because it was the equation that showed the best trade-off between fit statistics and measurement effort. The medium-size tree selection method, the random tree selection method and the all-size tree selection method performed better than others. This can be attributed to the fact that most forests in Durango are irregular, and the smaller and larger diameters show more variation among stands, whereas the mean diameters are much more similar between stands. Some authors (e.g. Castedo-Dorado et al. 2006; Crecente-Campo et al. 2010) obtained the best results when the smaller trees in the plot were used for calibration and attributed this to the fact that their models included dominant height as a fixed effect, and therefore heights corresponding to the larger trees did not provide much additional information for calibration. In the current study, Eq. 11 does not contain stand variables as regressors, but the large-tree method gave the worst results in calibration (Fig. 6). However, the smaller trees were not the

best possible sample in this study, perhaps because the selected model was not restricted to pass through any point. In the aforementioned studies, the models were restricted to pass through the point ($ddom$, $hdom$), implying that the models cannot change much in the upper part of the $h-d$ relationship, and small trees provided much more information for calibration than larger trees (Crecente-Campo et al. 2010). The greater the number of measurements included in the subsample, the greater the decrease in RMSE and increase in EF (Fig. 5). However, a large sample is often not justifiable, because of the higher cost of sampling (Castedo-Dorado et al. 2006). This was particularly important in this study, since calibration should be carried out independently for each sampling unit.

For the species-specific Eqs. 11 and 12, calibration can be done for each plot independently, as the plot-specific random effect associated with each species is different. A macro program may be very useful in this case. However, when calibrating Eqs. 9 and 10, calibration should be done for the whole dataset for which we want to obtain calibrated

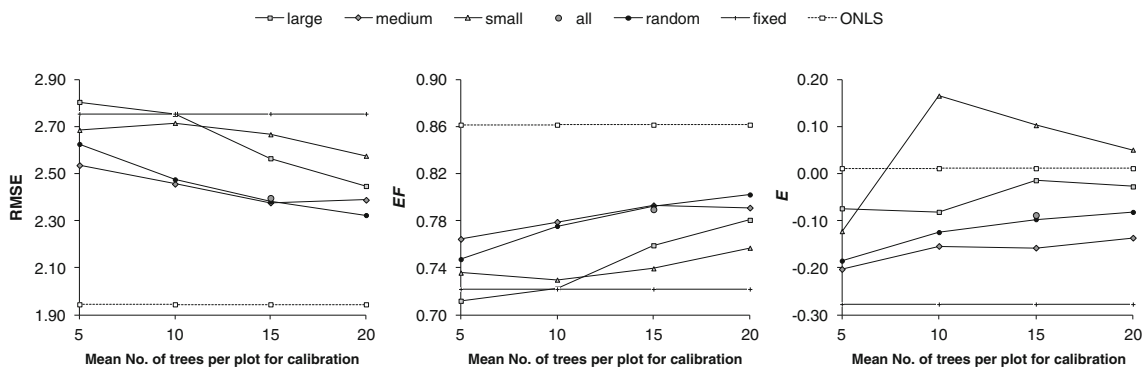


Fig. 6 Plots of the fitting statistics obtained with the different tree selection options for calibration of Eq. 11. The X-axis shows aggregate values to an average number of trees per plot, because the evaluation

dataset has a modal value of five species per plot. Lines and points in the legend refer to size of trees used in calibration

predictions (see Appendix A). This is because each species-specific random effect is the same for all the plots in which one species is present, and each plot-specific random effect is the same for all the species included in a particular plot. Therefore, each plot–species combination has a different random effect, obtained by summing the corresponding plot-specific and species-specific random effects.

Height–diameter models have previously been developed for other multi-species, multi-height and multi-age forests (e.g. Sharma and Parton 2007; Vargas-Larreta et al. 2009). However, the developed models were always species-specific (as Eqs. 11 and 12 in this study), and the utility of developing a general model like Eqs. 9 and 10 had not been shown. The approach used in this study could be applied in other irregular stands, where a general model with species-specific and plot-specific random effects could be calibrated by measuring only one tree of each species per plot.

However, the specific models developed in this study can only be used for the forests in Durango, because the variance and covariance parameters were calculated with the variability observed in the research plots, which were all located in this state. Variability in other areas, such as the bordering states of Chihuahua, Nayarit and Jalisco, where the Sierra Madre Occidental is located, may be different and therefore any predictions in these areas using these models should be taken with caution until the developed models are tested.

Following these results, we therefore recommend measuring one tree height for each species in the stands in order to apply the species-specific Eq. 11. We recommend the measurement of medium-sized trees for this purpose, because the random tree selection method can give worse results than the ones showed in this study if only one repetition is carried out (note that the values showed for the statistics of the random tree selection method are mean values after 100 simulations). If heights are not measured, the uncalibrated Eq. 11 should be used. If dominant height is available, and assuming that trees per hectare and basal area are available because all diameters would have been measured, then the uncalibrated Eq. 12 should be used. Finally, Eq. 12 can also be calibrated if dominant height and additional height measurements are available.

5 Conclusions

Random components of a basic h – d mixed model performed very well in explaining the observed differences in the h – d relationship between different plots and species in mixed uneven-aged stands in Durango (Mexico). However, the variance of the random effects for the different ecoregions was

non-significant. We likely do not have enough information to estimate such variance since only four ecoregions were included in the dataset. A generalised h – d model is not necessary if a calibrated basic h – d model is used, as similar results are obtained when only one tree per species is used in the calibration process. If the models are not being calibrated, species-specific models should be used because they at least capture the variation between species.

Funding Fondo Sectorial CONAFOR–CONACYT (Project CONAFOR–CONACYT 115900) (México); Fondo de Cooperación Internacional en Ciencia y Tecnología Unión Europea–México (Project FONCICYT-92739). Consellería de Economía e Industria (Galicia, Spain).

Appendix. Height estimation—an example

This Appendix shows an example of the height estimation using the calibrated Eqs. 9 and 11. Equations 10 and 12 should be calibrated in the same way as Eqs. 9 and 11, respectively.

We suppose that the following data belong to three new plots in which several species are present, but we are interested in calibrating the models for species 1, 2 and 3, which are the commercial ones. Thus, we measure one tree of these species in each one of the plots, resulting in:

Plot	Species	d (cm)	h (m)
1	1	15	12
1	2	17	12
1	3	14	12
2	1	20	13
2	2	19	13
2	3	16	13
3	1	17	11
3	2	15	11
3	3	18	11

Case 1: use of Eq. 9

In this case, calibration should be done for the whole dataset for which we want to calibrate Eq. 9, using the expression shown in Eq. 6.

Note that subscript i in Eq. 6 is finally not necessary because there were no differences between ecoregions, so hereafter it was omitted.

The estimated variances and the covariance of the random effects (Table 3) are the elements of the variance–covariance matrix $\hat{\mathbf{D}}$, which should be arranged according to the

dataset in the following way: three species, with two random parameters by species and three plots with

three random parameters by plot, resulting in a 15×15 diagonal matrix:

$$\hat{\mathbf{D}} = \begin{bmatrix} 0.01090 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.01090 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.01090 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.07487 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.07487 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.07487 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.1206 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.1206 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.1206 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.06554 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.06554 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.06554 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.7793 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.7793 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.7793 \end{bmatrix}$$

The variance–covariance matrix for the random error term is determined by assuming that all estimations have

constant variance (σ^2 , Table 3) and that the errors are not correlated:

$$\hat{\mathbf{R}}_{jk} = \sigma^2 \times \mathbf{I}_9 = \begin{bmatrix} 5.678 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 5.678 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 5.678 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 5.678 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5.678 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5.678 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 5.678 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 5.678 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 5.678 \end{bmatrix}$$

where \mathbf{I}_9 is the identity matrix with dimension (9×9) equal to the number of data used for calibration.

Eq. 9. Using $z = [1 - \exp(-\frac{a_2}{50}d_{jkl})] / [1 - \exp(-\frac{a_2}{50} \cdot 100)]$ for simplification, the partial derivatives are:

The partial derivatives matrix $\hat{\mathbf{Z}}_{jk}$ with respect to the random effects is calculated with the partial derivatives of

$$\frac{\partial h_{jkl}}{\partial w_{0k}} = \left(\frac{1}{a_0}\right)^2 \left[((20a_1)^{a_0} - 1.3^{a_0})z + 1.3^{a_0} \right]^{\frac{1}{a_0}} \cdot \left[\frac{a_0((\ln 20 + \ln a_1)(20a_1)^{a_0}z - 1.3^{a_0}z \ln 1.3 + 1.3^{a_0} \ln 1.3)}{((20a_1)^{a_0} - 1.3^{a_0})z + 1.3^{a_0}} - \ln(((20a_1)^{a_0} - 1.3^{a_0})z + 1.3^{a_0}) \right]$$

$$\frac{\partial h_{jkl}}{\partial w_{1k}} = 20^{a_0} a_1^{a_0-1} z [((20a_1)^{a_0} - 1.3^{a_0})z + 1.3^{a_0}]^{\left(\frac{1}{a_0}-1\right)}$$

$$\frac{\partial h_{jkl}}{\partial v_{0j}} = \frac{\partial h_{jkl}}{\partial w_{0k}}$$

$$\frac{\partial h_{jkl}}{\partial v_{1j}} = \frac{\partial h_{jkl}}{\partial w_{1k}}$$

$$\frac{\partial h_{jkl}}{\partial v_{2j}} = \frac{(1.3^{a_0} - (20a_1)^{a_0}) \exp\left[\left(2 - \frac{d_{jkl}}{50}\right)a_2\right]}{(50a_0(\exp(2a_2) - 1)^2)} \cdot \left(d_{jkl}(-\exp(2a_2)) + 100\left(\exp\left(\frac{a_2}{50}d_{jkl}\right) - 1\right) + d_{jkl} \right) \cdot [((20a_1)^{a_0} - 1.3^{a_0})z + 1.3^{a_0}]^{\left(\frac{1}{a_0}-1\right)}$$

The values for the partial derivatives are:

Plot	Species	<i>d</i> (cm)	<i>h</i> (m)	$\frac{\partial h_{jkl}}{\partial w_{0k}}$	$\frac{\partial h_{jkl}}{\partial w_{1k}}$	$\frac{\partial h_{jkl}}{\partial v_{0j}}$	$\frac{\partial h_{jkl}}{\partial v_{1j}}$	$\frac{\partial h_{jkl}}{\partial v_{2j}}$
1	1	15	12	6.0907	8.2505	6.0907	8.2505	3.3777
1	2	17	12	6.0354	9.0770	6.0354	9.0770	3.5535
1	3	14	12	6.0887	7.8158	6.0887	7.8158	3.2712
2	1	20	13	5.8419	10.2163	5.8419	10.2163	3.7346
2	2	19	13	5.9183	9.8494	5.9183	9.8494	3.6844
2	3	16	13	6.0720	8.6707	6.0720	8.6707	3.4716
3	1	17	11	6.0354	9.0770	6.0354	9.0770	3.5535
3	2	15	11	6.0907	8.2505	6.0907	8.2505	3.3777
3	3	18	11	5.9835	9.4697	5.9835	9.4697	3.6243

The partial derivatives should be arranged in the matrix into different columns, to reflect how the data is arranged in

the dataset, to obtain a 9×15 matrix (9 observations used in calibration and 15 random parameters) as follows:

$$\widehat{\mathbf{Z}}_{jk} = \begin{bmatrix} \frac{\partial h_{jkl}}{\partial w_{0k}} & 0 & 0 & \frac{\partial h_{jkl}}{\partial w_{1k}} & 0 & 0 & \frac{\partial h_{jkl}}{\partial v_{0j}} & 0 & 0 & \frac{\partial h_{jkl}}{\partial v_{1j}} & 0 & 0 & \frac{\partial h_{jkl}}{\partial v_{2j}} & 0 & 0 \\ 0 & \frac{\partial h_{jkl}}{\partial w_{0k}} & 0 & 0 & \frac{\partial h_{jkl}}{\partial w_{1k}} & 0 & \frac{\partial h_{jkl}}{\partial v_{0j}} & 0 & 0 & \frac{\partial h_{jkl}}{\partial v_{1j}} & 0 & 0 & \frac{\partial h_{jkl}}{\partial v_{2j}} & 0 & 0 \\ 0 & 0 & \frac{\partial h_{jkl}}{\partial w_{0k}} & 0 & 0 & \frac{\partial h_{jkl}}{\partial w_{1k}} & \frac{\partial h_{jkl}}{\partial v_{0j}} & 0 & 0 & \frac{\partial h_{jkl}}{\partial v_{1j}} & 0 & 0 & \frac{\partial h_{jkl}}{\partial v_{2j}} & 0 & 0 \\ \frac{\partial h_{jkl}}{\partial w_{0k}} & 0 & 0 & \frac{\partial h_{jkl}}{\partial w_{1k}} & 0 & 0 & 0 & \frac{\partial h_{jkl}}{\partial v_{0j}} & 0 & 0 & \frac{\partial h_{jkl}}{\partial v_{1j}} & 0 & 0 & \frac{\partial h_{jkl}}{\partial v_{2j}} & 0 \\ 0 & \frac{\partial h_{jkl}}{\partial w_{0k}} & 0 & 0 & \frac{\partial h_{jkl}}{\partial w_{1k}} & 0 & 0 & \frac{\partial h_{jkl}}{\partial v_{0j}} & 0 & 0 & \frac{\partial h_{jkl}}{\partial v_{1j}} & 0 & 0 & \frac{\partial h_{jkl}}{\partial v_{2j}} & 0 \\ 0 & 0 & \frac{\partial h_{jkl}}{\partial w_{0k}} & 0 & 0 & \frac{\partial h_{jkl}}{\partial w_{1k}} & 0 & \frac{\partial h_{jkl}}{\partial v_{0j}} & 0 & 0 & \frac{\partial h_{jkl}}{\partial v_{1j}} & 0 & 0 & \frac{\partial h_{jkl}}{\partial v_{2j}} & 0 \\ \frac{\partial h_{jkl}}{\partial w_{0k}} & 0 & 0 & \frac{\partial h_{jkl}}{\partial w_{1k}} & 0 & 0 & 0 & 0 & \frac{\partial h_{jkl}}{\partial v_{0j}} & 0 & 0 & \frac{\partial h_{jkl}}{\partial v_{1j}} & 0 & 0 & \frac{\partial h_{jkl}}{\partial v_{2j}} \\ 0 & \frac{\partial h_{jkl}}{\partial w_{0k}} & 0 & 0 & \frac{\partial h_{jkl}}{\partial w_{1k}} & 0 & 0 & 0 & \frac{\partial h_{jkl}}{\partial v_{0j}} & 0 & 0 & \frac{\partial h_{jkl}}{\partial v_{1j}} & 0 & 0 & \frac{\partial h_{jkl}}{\partial v_{2j}} \\ 0 & 0 & \frac{\partial h_{jkl}}{\partial w_{0k}} & 0 & 0 & \frac{\partial h_{jkl}}{\partial w_{1k}} & 0 & 0 & \frac{\partial h_{jkl}}{\partial v_{0j}} & 0 & 0 & \frac{\partial h_{jkl}}{\partial v_{1j}} & 0 & 0 & \frac{\partial h_{jkl}}{\partial v_{2j}} \end{bmatrix}$$

resulting in:

$$\widehat{\mathbf{Z}}_{jk} = \begin{bmatrix} 6.0907 & 0 & 0 & 8.2505 & 0 & 0 & 6.0907 & 0 & 0 & 8.2505 & 0 & 0 & 3.3777 & 0 & 0 \\ 0 & 6.0354 & 0 & 0 & 9.0770 & 0 & 6.0354 & 0 & 0 & 9.0770 & 0 & 0 & 3.5535 & 0 & 0 \\ 0 & 0 & 6.0887 & 0 & 0 & 7.8158 & 6.0887 & 0 & 0 & 7.8158 & 0 & 0 & 3.2712 & 0 & 0 \\ 5.8419 & 0 & 0 & 10.2163 & 0 & 0 & 0 & 5.8419 & 0 & 0 & 10.2163 & 0 & 0 & 3.7346 & 0 \\ 0 & 5.9183 & 0 & 0 & 9.8494 & 0 & 0 & 5.9183 & 0 & 0 & 9.8494 & 0 & 0 & 3.6844 & 0 \\ 0 & 0 & 6.0720 & 0 & 0 & 8.6707 & 0 & 6.0720 & 0 & 0 & 8.6707 & 0 & 0 & 3.4716 & 0 \\ 6.0354 & 0 & 0 & 9.0770 & 0 & 0 & 0 & 0 & 6.0354 & 0 & 0 & 9.0770 & 0 & 0 & 3.5535 \\ 0 & 6.0907 & 0 & 0 & 8.2505 & 0 & 0 & 0 & 6.0907 & 0 & 0 & 8.2505 & 0 & 0 & 3.3777 \\ 0 & 0 & 5.9835 & 0 & 0 & 9.4697 & 0 & 0 & 5.9835 & 0 & 0 & 9.4697 & 0 & 0 & 3.6243 \end{bmatrix}$$

The matrix $\hat{\mathbf{e}}_{jkl}$ containing the errors obtained with the fixed-effects model (Eq. 9 with w_{0k} , w_{1k} , v_{0j} , v_{1j} and v_{2j} equal 0) is:

$$\hat{\mathbf{e}}_{jkl} = \begin{bmatrix} 1.7949 \\ 0.9066 \\ 2.2621 \\ 0.6825 \\ 1.0767 \\ 2.3432 \\ -0.09339 \\ 0.7949 \\ -0.5154 \end{bmatrix}$$

Therefore, the random parameters calculated using Eq. 6 were:

$$\hat{\mathbf{b}}_{jk} = \begin{bmatrix} w_{01} \\ w_{02} \\ w_{03} \\ w_{11} \\ w_{12} \\ w_{13} \\ v_{01} \\ v_{02} \\ v_{03} \\ v_{11} \\ v_{12} \\ v_{13} \\ v_{21} \\ v_{22} \\ v_{23} \end{bmatrix} = \begin{bmatrix} 0.01561 \\ 0.02045 \\ 0.05229 \\ 0.002747 \\ 0.009959 \\ 0.03455 \\ 0.05549 \\ 0.04252 \\ -0.0002459 \\ 0.03248 \\ 0.02262 \\ -0.01373 \\ 0.1756 \\ 0.1310 \\ -0.03699 \end{bmatrix}$$

This estimation of the random effects can be used to calibrate the model to obtain a specific response for the plots and species of this example, using Eq. 9.

The estimated heights and associated errors using the fixed effects Eq. 9 and the calibrated Eq. 9 are, therefore:

Plot	Species	d (cm)	h (m)	$h_{jkl} \text{ fixed}$	$e_{jkl} \text{ fixed}$	$h_{jkl} \text{ cal.}$	$e_{jkl} \text{ cal.}$
1	1	15	12	10.21	1.79	11.54	0.46
1	2	17	12	11.09	0.91	12.58	-0.58
1	3	14	12	9.74	2.26	11.55	0.45
2	1	20	13	12.32	0.68	13.41	-0.41
2	2	19	13	11.92	1.08	13.11	-0.11
2	3	16	13	10.66	2.34	12.22	0.78
3	1	17	11	11.09	-0.09	10.95	0.05
3	2	15	11	10.21	0.79	10.17	0.83
3	3	18	11	11.52	-0.52	11.89	-0.89
				$\sum \hat{\mathbf{e}}_{jkl} $	10.47		4.54

It can be observed how the prediction errors are reduced from using a fixed-effects prediction to using a calibrated prediction.

Case 2: use of Eq. 11

This is the usual way of calibrating a mixed-effects model with only one sampling level (i.e. the plot level), as done in many previous studies (e.g. Calama and Montero 2004; Castedo-Dorado et al. 2006; Sharma and Parton 2007). In this case, the calibration process has to be done for each species independently, because each species has a different model (Eq. 11 with the parameters shown in Table 3).

Hereafter, we include an example for species 1 on plot 1. The other species and plots would be calibrated in the same way.

Not that for Eq. 11 and species 1, only the variance for the random effect v_1 was significant, therefore only one random effect has to be calculated.

The estimated variances and the covariance of the random effects (Table 3) are the elements of the variance-covariance matrix $\hat{\mathbf{D}}$, conforming a simple (1×1) diagonal matrix (1 random parameter):

$$\hat{\mathbf{D}} = [0.05110]$$

The variance-covariance matrix for the random error term is determined by assuming that all estimations have constant variance (σ^2 , Table 3) and that the errors are not correlated. In this case, it only contains one observation:

$$\hat{\mathbf{R}}_{jk} = \sigma^2 \times \mathbf{I}_1 = [5.308]$$

where \mathbf{I}_1 is the identity matrix with dimension (1×1) equal to the number of data points used in calibration.

The partial derivatives matrix $\hat{\mathbf{Z}}_{jk}$ with respect to the random effects is calculated with the partial derivatives of Eq. 11, which have the same expression as those from Eq. 9.

The values for the partial derivatives are:

Plot	Species	d (cm)	h (m)	$\frac{\partial h_{jkl}}{\partial v_{1j}}$
1	1	15	12	7.7013

Therefore the corresponding matrix is:

$$\hat{\mathbf{Z}}_{jk} = \left[\frac{\partial h_{jkl}}{\partial v_{1j}} \right] = [7.7013]$$

The matrix $\hat{\mathbf{e}}_{jkl}$ containing the errors obtained with the fixed-effects model (Eq. 11 with v_{0jk} , v_{1jk} and v_{2jk} equal 0) is:

$$\hat{\mathbf{e}}_{jkl} = [1.2840]$$

Therefore, the random effect estimated using Eq. 6 was:

$$\hat{\mathbf{b}}_{jk} = [v_{11}] = [0.06060]$$

This estimation of the random effect can be used to calibrate the model to obtain a specific response for the plot and species of this example, using Eq. 11.

The estimated heights and associated errors using the fixed effects of Eq. 11 and the calibrated Eq. 11 are, therefore:

Plot	Species	d (cm)	h (m)	$h_{jkl} \text{ fixed}$	$e_{jkl} \text{ fixed}$	$h_{jkl} \text{ cal.}$	$e_{jkl} \text{ cal.}$
1	1	15	12	10.71	1.28	11.18	0.82

References

- Akaike H (1974) A new look at the statistical model identification. *IEEE T Automat Contr* AC-19:716–723
- Amaral Paulo J, Tomé J, Tomé M (2011) Nonlinear fixed and random generalized height-diameter models for Portuguese cork oak stand. *Ann For Sci* 68:295–309
- Bailey RL, Brooks JR (1994) Determining site index and estimating timber volumes without measuring heights. *South J Appl For* 18:15–18
- Beal SL, Sheiner LB (1982) Estimating population kinetics. *Crit Rev Biomed Eng* 8:195–222
- Bertalanffy L (1949) Problems of organic growth. *Nature* 163:156–158
- Calama R, Montero G (2004) Interregional nonlinear height–diameter model with random coefficients for stone pine in Spain. *Can J Forest Res* 34:150–163
- Calama R, Montero G (2005) Multilevel linear mixed model for tree diameter increment in stone pine (*Pinus pinea*): a calibrating approach. *Silva Fenn* 39:37–54
- Camille A, García-López GA (2008) Cross-scale linkages in common-pool resource management: the evolution of forest associations in the Mexican forest commons. Prepared for the 12th IASC 2008 biennial conference. University of Gloucester, Cheltenham
- Castedo-Dorado F, Diéguez-Aranda U, Barrio Anta M, Sánchez Rodríguez M, Kv G (2006) A generalized height–diameter model including random components for radiata pine plantations in north-western Spain. *Forest Ecol Manage* 229:202–213
- Chung Y, Rabe-Hesketh S, Dorie V, Gelman A, Liu J (2013). A nondegenerate penalized likelihood estimator for variance parameters in multilevel models. *Psychometrika* (in press). doi:10.1007/S11336-013-9328-2
- Corral-Rivas JJ, Álvarez-González JG, Ruiz-González AD, Kv G (2004) Compatible height and site index models for five pine species in El Salto, Durango (Mexico). *Forest Ecol Manage* 201:145–160
- Crecente-Campo F, Tomé M, Soares P, Diéguez-Aranda U (2010) A generalized nonlinear mixed-effects height-diameter model for *Eucalyptus globulus* L. in northwestern Spain. *Forest Ecol Manage* 259:943–952
- Curtis RO (1967) Height–diameter and height–diameter–age equations for second-growth Douglas-fir. *Forest Sci* 13:365–375
- Davidian M, Giltinan D (1993) Some general estimation methods for nonlinear mixed effects models. *J Biopharm Stat* 3:23–55
- Goldstein H (1995) *Multilevel statistical models*, 2nd edn. Halstead Press, New York, 178 pp
- Hall DB, Bailey RL (2001) Modeling and prediction of forest growth variables based on multilevel nonlinear mixed models. *Forest Sci* 47:311–321
- Hartford A, Davidian M (2000) Consequences of misspecifying assumptions in nonlinear mixed effects models. *Comput Stat Data Anal* 34:139–164
- Huang S, Price D, Titus SJ (2000) Development of ecoregion-based height-diameter models for white spruce in boreal forests. *Forest Ecol Manage* 129:125–141
- Jayaraman K, Lappi J (2001) Estimation of height–diameter curves through multilevel models with special reference to even-aged teak stands. *Forest Ecol Manage* 142:155–162
- Kozak A, Kozak R (2003) Does cross validation provide additional information in the evaluation of regression models? *Can J Forest Res* 33:976–987
- Lappi J (1991) Calibration of height and volume equations with random parameters. *Forest Sci* 37:781–801
- Lappi J (1997) A longitudinal analysis of height/diameter curves. *Forest Sci* 43:555–570
- Lei Y, Parresol BR (2001) Remarks on height-diameter modelling. USDA For. Serv. Research Note SRS-10. Southern Research Station, Asheville
- Lindstrom MJ, Bates DM (1990) Nonlinear mixed-effects models for repeated measures data. *Biometrics* 46:673–687
- Littell RC, Milliken GA, Stroup WW, Wolfinger RD, Schabenberger O (2006) *SAS® for mixed models*, 2nd edn. SAS Institute Inc, Cary
- Longford NT (1993) *Random coefficient models*. Clarendon Press, Oxford, 270 pp
- López Sánchez CA, Gorgoso Varela J, Castedo Dorado F, Rojo Alboreca A, Rodríguez Soalleiro R, Alvarez González JG, Sánchez Rodríguez F (2003) A height-diameter model for *Pinus radiata* D. Don in Galicia (Northwest Spain). *Ann Forest Sci* 60:237–245
- Lynch TB, Holley AG, Stevenson DJ (2005) A random-parameter height–dbh model for cherrybark oak. *South J Appl For* 29:22–26
- Morrison ML, Marcot BG, Mannan RW (1992) *Wildlife habitat relationships: concepts and applications*. Univ. Wisconsin Press, Madison
- Newton PF, Amponsah IG (2007) Comparative evaluation of five height–diameter models developed for black spruce and jack pine stand-types in terms of goodness-of-fit, lack-of-fit and predictive ability. *Forest Ecol Manage* 247:149–166
- Parresol BR (1992) Baldcypress height–diameter equations and their prediction confidence interval. *Can J Forest Res* 22:1429–1434
- Pinheiro JC, Bates DM (2000) *Mixed-effects models in S and S-PLUS*. Springer Verlag, New York
- Richards FJ (1959) A flexible growth function for empirical use. *J Exp Bot* 10:290–300
- SAS Institute Inc (2008) *SAS/ETS® 9.2 User’s Guide*. SAS Institute Inc, Cary
- SAS Institute Inc (2009) *SAS/STAT® 9.2 User’s guide*, 2nd edn. SAS Institute Inc, Cary
- Saunders MR, Wagner RG (2008) Height-diameter models with random coefficients and site variables for tree species in central Maine. *Ann For Sci* 65:203p1–203p10
- Schnute J (1981) A versatile growth model with statistically stable parameters. *Can J Fish Aquat Sci* 38:1128–1140
- Sharma M, Parton J (2007) Height–diameter equations for boreal tree species in Ontario using a mixed-effects modeling approach. *Forest Ecol Manage* 249:187–198
- Sharma M, Zhang SY (2004) Height–diameter models using stand characteristics for *Pinus banksiana* and *Picea mariana*. *Scand J Forest Res* 19:442–451
- Soares P, Tomé M (2002) Height–diameter equation for first rotation eucalypt plantations in Portugal. *Forest Ecol Manage* 166:99–109
- SRNyMA (2006) Programa estratégico forestal 2030. Secretaría de recursos naturales y medio ambiente del Estado de Durango. Dgo, Victoria de Durango
- Thoms CA, Betters DR (1998) The potential for ecosystems management in Mexico’s forest ejidos. *Forest Ecol Manage* 103:149–1179
- Trincado G, VanderSchaaf CL, Burkhart HE (2007) Regional mixed-effects height–diameter models for loblolly pine (*Pinus taeda* L.) plantations. *Eur J Forest Res* 126:253–262

- Vanclay JK (2009) Tree diameter, height and stocking in even-aged forest. *Ann For Sci* 66:702p1–702-p7
- Vargas-Larreta B, Castedo-Dorado F, Álvarez-González JG, Barrio-Anta M, Cruz-Cobos F (2009) A generalized height-diameter model with random coefficients for uneven-aged stands in El Salto, Durango (Mexico). *Forestry* 82:445–462
- Vonesh EF (1996) A note on the use of Laplace's approximation for nonlinear mixed effects models. *Biometrika* 82:447–452
- Vonesh EF, Chinchilli VM (1997) Linear and nonlinear models for the analysis of repeated measurements. Marcel Dekker Inc, New York
- Wehenkel C, Corral-Rivas JJ, Hernández-Díaz JC, Gadov Kv (2011) Estimating balanced structure areas in multi-species forests on the Sierra Madre Occidental, Mexico. *Ann For Sci* 68:385–394
- West PW, Ratkowsky DA, Davis AW (1984) Problems of hypothesis testing of regressions with multiple measurements from individual sampling units. *Forest Ecol Manage* 7:207–224
- Yang RC, Kozak A, Smith JH (1978) The potential of Weibull-type functions as a flexible growth curve. *Can J Forest Res* 8:424–431
- Zhang L (1997) Cross-validation of nonlinear growth functions for modeling tree height-diameter distributions. *Ann Bot* 79:251–257
- Zhang SA, Burkhart HE, Amateis RL (1997) The influence of thinning on tree height and diameter relationships in loblolly pine plantations. *South J of Appl For* 21:199–205