Analysis and comparison of nonlinear tree height prediction strategies for Douglas-fir forests

H. Temesgen, V.J. Monleon, and D.W. Hann

Abstract: Using an extensive Douglas-fir data set from southwest Oregon, we examined the (1) performance and suitability of selected prediction strategies, (2) contribution of relative position and stand-density measures in improving tree height (*h*) prediction values, and (3) effect of different subsampling designs to fill in missing *h* values in a new stand using a regional nonlinear model. Nonlinear mixed-effects models (NMEM) substantially improved the accuracy and precision of height prediction over the conventional nonlinear fixed-effects model (NFEM) that assumes the observations are independent, particularly when a few trees are subsampled for height. The predictive performance of a correction factor on a NFEM with relative position and stand-density measures was comparable to that of a NMEM when four or more trees were subsampled for height. When two or more heights were randomly subsampled, the NMEM efficiently explained the differences in the height–diameter relationship because of the variations in relative position of trees and stand density without having to incorporate them into the model. When only one height was subsampled, selecting the largest diameter tree in the stand would result in a lower predicted root mean square error (RMSE) than randomly selecting the height, regardless of the model form or fitting strategy used.

Résumé : À l'aide d'une banque de données exhaustive sur le sapin Douglas du sud-ouest de l'Oregon, nous avons examiné (1) la performance et la pertinence des stratégies de prédiction sélectionnées, (2) la contribution de la position relative de l'arbre et de la densité du peuplement pour améliorer la prédiction de la hauteur des arbres et (3) l'effet de différents dispositifs d'échantillonnage pour imputer la hauteur manquante dans un nouveau peuplement à l'aide d'un modèle non linéaire régional. Les modèles non linéaires à effets mixtes (MNLEM) améliorent substantiellement l'exactitude et la précision des prédictions de la hauteur comparativement au modèle non linéaire à effets fixes conventionnel (MNLEF). Ce dernier suppose que les observations sont indépendantes, particulièrement lorsque peu d'arbres sont échantillonnés pour évaluer la hauteur. La performance prédictive d'un facteur de correction pour le MNLEF basé sur la mesure de la position relative de l'arbre et de la densité du peuplement est comparable à celle du MNLEM lorsque quatre arbres ou plus sont échantillonnés pour évaluer la hauteur. Lorsque deux hauteurs ou plus sont échantillonnées aléatoirement, le MNLEM explique efficacement les différences dans la relation hauteur-diamètre dues aux variations de la position relative des arbres et de la densité sans avoir à les incorporer formellement dans le modèle. Lorsqu'une seule hauteur est échantillonnée, le choix du plus gros arbre dans le peuplement pourrait entraîner une erreur de prédiction plus faible que lorsque la hauteur est sélectionnée au hasard, peu importe la forme du modèle ou la stratégie d'ajustement utilisée.

[Traduit par la Rédaction]

Introduction

Modelling stand development over time relies on accurate estimates of tree height (*h*) and diameter (*d*). Accurate height measurements are required for describing vertical stand structure and estimating stand development over time (e.g., Dubrasich et al. 1997), stand volume, and site quality (Clutter et al. 1983). However, height is costly to measure and, as a result, trees are frequently subsampled for height. Often the subsample is concentrated in the trees of greatest diameter; for example, those used to estimate site index. Subsampled heights can also be used to localize regional height–diameter (*h*–*d*) functions (e.g., Wykoff et al. 1982; Robinson and Wykoff 2004; Hann 2005).

Many growth and yield models require height and diameter as basic input variables, with all or part of the heights predicted from measured diameters using regional height–diameter functions (Wykoff et al. 1982, Huang et al. 1992, Hann 2005). Regional height diameter functions can also be used to indirectly predict height growth (Larsen and Hann 1987). For example, in the southwestern Oregon version of the ORGANON growth and yield model (SWO-ORGANON; Hann 2005), missing heights are directly predicted using the species specific height–diameter equations of Hanus et al. (1999*a*), and these equations are also used to estimate height growth from diameter growth for minor species. Tree height is also a critical variable in many process and hybrid models, such as Biome-BGC (Running and Coughlan 1988)

Received 1 January 2007. Accepted 20 May 2007. Published on the NRC Research Press Web site at cjfr.nrc.ca on 25 February 2008.

H. Temesgen[1] and D. Hann. Department of Forest Resources, Oregon State University, 280 Peavy Hall, Corvallis, OR 97331-5703, USA.
V. Monleon. Forest Inventory and Analysis, Pacific Northwest Research Station, USDA Forest Service. 620 SW Main Street, Suite 400, Portland, OR 97205, USA.

[1]Corresponding author (e-mail: hailemariam.temesgen@oregonstate.edu).

 doi:10.1139/X07-104

and 3-PG (Landsberg and Waring 1997), as height provides a numerical link between processes that are based on leaf area (e.g., light capture, photosynthesis, and transpiration) and foliage mass (Temesgen and Weiskittel 2006).

The relationship between tree height and diameter varies from stand to stand owing to differences in site quality (Larsen and Hann 1987, Wang and Hann 1988), stand density (Larsen and Hann 1987, Zeide and Vanderschaaf 2001, Temesgen and von Gadow 2004), and stand age (Zeide and Vanderschaaf 2001). Even within the same stand, the relationship varies over time (Curtis 1967); by relative position of trees in a stand (Temesgen and von Gadow 2004); and spatial distribution pattern (Aguirre et al. 2003). Thus, localization of the regional height–diameter relationship is an important step in obtaining accurate growth and yield estimates.

Conventionally, regional height–diameter models have used only diameter as a predictor variable (Wykoff et al. 1982; Huang et al. 1992; Lappi 1997). However, several studies have explored the use of additional variables that might influence the height–diameter relationship both within and between stands. As examples, stand-density measures have been used by Larsen and Hann (1987), Temesgen and von Gadow (2004), and Castedo et al. (2006); relative tree position variables have been used by Temesgen and von Gadow (2004); site quality variables have been used by Larsen and Hann (1987) and Wang and Hann (1988); and the average h and d of the top height trees have been used by Krumland and Wensel (1978) and Hanus et al. (1999*b*) for even-aged stands. The inclusion of these variables improved the precision of height estimates, as expected. However, Sharma and Zhang (2004) found that site index did not improve the precision of height estimates.

Data for development of regional height–diameter equations are often collected on trees sampled in plots selected from multiple stands. Because trees from the same plot tend to be more similar to each other than to trees from different plots, the classical regression assumption that observations are independent does not hold (Neter et al. 1990, Gregoire et al. 1995). Conventional regression models do not account for the clustered and nested structure of the data, nor do they identify the different sources of variation. Recognizing this lack of independence, several recent studies have explored the use of linear mixed-effects models (LMEM; Monleon 2003; Robinson and Wykoff 2004; Mehtätalo 2004), and nonlinear mixed-effects models (NMEM; Calama and Montero 2004; Castedo et al. 2006) to predict height as a function of diameter and to account for a stand effect.

From a practical perspective, it has been reported that LMEM and NMEM allow a more accurate and precise estimation of the height–diameter relationship than conventional linear and nonlinear regression models that assume that observations are independent (we shall abbreviate these latter approaches as LFEM and NFEM). The LMEM and NMEM approaches also account for the correlation structure of the data and provide realistic variance estimates for stochastic simulation and for modelling natural variability (e.g., Lappi 1997; Castedo et al. 2006). The LMEM and NMEM estimates are effectively empirical Bayes estimates and can be motivated from a Bayesian perspective. The prediction of height for trees from a new stand is based on the prior information from the training data set, actualized with new data collected as a subsample of heights from the stand.

Despite the growing research interest in LMEM and NMEM for predicting heights, detailed analyses that quantify the gains obtained by using these methods over LFEM or NFEM are lacking. Also lacking are analyses that examine the ability, efficiency, and suitability of any of these approaches for predicting missing height measurements when the trees subsampled for height in the new stand are selected using different subsampling designs and subsample sizes.

The primary objective of this study was to examine the performance of NMEM for predicting tree height when a subsample of heights from the new stand is available. Using an extensive data set of Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco var. *glauca* (Beissn.) Franco) in southwestern Oregon, we examined the performance of NFEM and NMEM for predicting missing heights from a regional height–diameter equation that has been localized using a subsample of trees measured for height in the new stand. The subsample of heights was selected using different designs and subsample sizes. We also evaluated the contribution that stand-level variables, in particular relative position and stand-density measures, can provide in improving the prediction of height from regional height–diameter equations fitted using these alternative strategies. In comparing the different strategies, we examined three questions concerning the predictive performance, efficiency, and suitability of these methods to fill-in missing heights.

1. Will the accuracy and precision of height prediction vary by the type of prediction strategy? Is there substantial improvement in accuracy and precision by predicting heights using NMEM over the conventional NFEM strategy?
2. How does the predictive ability of localizing NMEM and NFEM from a subsample of heights in the new stand vary by subsampling design and subsample size?
3. Will the prediction of the random effects in NMEM efficiently explain the differences in the height–diameter relationship due to stand-related effects, such as variations in relative position of trees and stand density?

Data

The data were collected in two studies associated with the development of the growth and yield model SWO-ORGANON (Hann 2005). The first set of data was collected between 1981 and 1983, as part of the southwest Oregon Forestry Intensified Research (FIR) Growth and Yield Project. This study included 391 plots in an area extending from near the California border (42°10′N) in the south to Cow Creek (43°00′N) in the north, and from the Cascade crest (122°15′W) in the east to approximately 15 mi. (1 mi. = 1.609 km) west of Glendale, California (123°50′W). Elevation of the sample plots ranged from 250 to 1600 m. Selection was limited to stands under 120 years of age and with 80% of its basal area consisting of conifer species. The second study's data was collected between 1992 and 1996. It covered about the same area, but extended the selection criteria to include stands with trees over 250 years old and younger stands with a greater component of hardwoods. An additional 138 plots were measured in the second study. Stands treated in the previous 5 years were not sampled in either study.

Table 1. Summary of tree- and stand-level attributes used in the study.

	Min.	Mean	Max.	SD
Tree-level attributes				
Diameter (cm)	0.3	35.7	178.9	21.5
Height (m)	1.4	25.9	62.1	11.9
Basal area in larger trees (m^2/ha)	0.0	25.8	93.4	16.7
Crown competition factor in larger trees	0.0	112.4	485.6	79.1
Stand-level attributes				
Basal area (m^2/ha)	8.1	52.1	101.0	15.7
Crown competition factor	51.6	263.8	536.7	79.1
Site index (m)	12.6	30.7	44.8	5.4

Note: No. of trees used in the study was 4948 and no. of stands in the study was 142.

In both studies, each stand was sampled with 4–25 sample points spaced 45.73 m apart. The sampling grid was established in a manner such that all sample points were at least 30.5 m from the edge of the stand. Therefore, the specific parameter estimates derived from this study may not be applicable to edge trees in southwestern Oregon. At each sample point, trees were sampled with a nested plot design composed of four subplots: trees having a $d < 10.2$ cm were selected on a circular subplot with a fixed radius of 2.37 m; trees having a d = 10.3–20.3 cm were selected on a circular subplot with a fixed radius of 4.74 m; trees having a d = 20.4–91.4 cm were selected on a 4.592-BAF variable radius subplot; and trees having a $d > 91.4$ cm were selected on a 13.776-BAF variable radius subplot.

Measurements of height and diameter were taken on all sample trees. Tree diameter was measured to the nearest 2.54 mm with a diameter tape. Tree height was measured on all trees either directly with a 25–45 ft (1 ft = 0.3048 m) telescoping fiberglass pole or, for taller trees, indirectly using the pole–tangent method (Larsen and Hann 1987) and recorded to the nearest 0.03 m. For trees with broken or dead tops, height was measured to the top of the live crown.

A total of 30 tree species were found on 529 plots. The number of species found on a single plot ranged from 1 to 12 and averaged almost five species. Douglas-fir was the most common species, found on 339 plots, and will be the focus of this study. From each untreated plot, heights and diameters of undamaged trees were extracted to assess and evaluate selected height prediction strategies. Since one of the main objectives of this study was to evaluate the predictive performance of the models as a function of the number of heights subsampled from the stand, only stands with at least 25 Douglas-fir sample trees were included. The data set covered a wide array of stand densities, with the basal area (BA) ranging from 8.1 to 101.0 m^2/ha; the crown competition factor (CCF) ranging from 51.6 to 536.7; d ranging from 0.3 to 178.9 cm; and h ranging from 1.4 to 62.1 m (Table 1).

Methods

Models and prediction strategies

In a related study, five sets of four nonlinear regional height–diameter equations (for a total of 20 alternative equations) were evaluated (Temesgen et al. 2007). The first set included four base equations for estimating height as a function of diameter alone: a Weibull-based equation applied by Yang et al. (1978) to tree species in British Columbia; a Chapman–Richards (Richards 1959) equation applied by Garman et al. (1995) to 24 western Oregon tree species; a function proposed by Ratkowsky (1990) and used by Flewelling and de Jong (1994) for western hemlock in the coastal region of the Pacific Northwest; and an equation used by Larsen and Hann (1987), Wang and Hann (1988), and Hanus et al. (1999*a*, 1999*b*) for 26 tree species in the Pacific Northwest. Two functional forms were selected based on their superior predictive performance: the Chapman–Richards equation and the equation used by Hanus et al. (1999*a*, 1999*b*).

These two base equations were then modified to include various tree- and stand-level variables. The crown length of a tree affects its form and, as a result, the height–diameter relationship (Larson 1963). Factors affecting crown length include the relative position of the tree within the stand and the stand's density (Ritchie and Hann 1987; Zumrawi and Hann 1989; Hann et al. 2003). It is expected that a decrease in either relative position or an increase in stand density would result in an increase in predicted height for a given diameter. Thus, two stand-density measures (basal area per hectare in m^2/ha (BA) and the crown competition factor (CCF) of Krajicek et al. 1961), and two relative position measures (crown competition factor in larger trees (CCFL) and basal area in larger trees in m^2/ha (BAL)) were evaluated for potential improvement in the predictive abilities of each of the base equations. Computation of all four of these variables requires the measurement of only diameter for all sample trees in the stand and knowledge of the sampling probability for each sample tree. Both CCF and CCFL were calculated using the maximum crown width equations of Paine and Hann (1982). The Chapman–Richards equation with CCFL and BA resulted in the best overall predictive performance (Temesgen et al. 2007) and, therefore, was selected for further examination in this study.

In our previous study, all regional height–diameter models were fitted and evaluated under the assumption that each tree was an independent observation and did not include random stand effects. Here, the predictive performance of the Chapman–Richards equation will be further examined under a combination of model-fitting strategies and tree- and stand-level variables. Briefly, the four modelling strategies examined are NFEM under the traditional assumption that all observations are independent; a NMEM with a random stand effect on the asymptote, thus explicitly incorporating the clustered structure of the sample; a NMEM with two random stand effects, the model asymptote and curvature parameters; and a stand-level height correction factor calculated using ordinary least-squares regression (OLS) and applied to the predicted heights from the NFEM (Hanus et al. 1999*a*). Within each of the four model-fitting strategies, we estimated h as a function of d alone and as a function of d, CCFL, and BA. The predictive performance of the Hanus et al. (1999*a*, 1999*b*) equation was also evaluated. The results were almost identical to those of the Chapman–Richards equation and will not be reported.

Strategy 1 — NFEM assuming independent observations

The Chapman–Richards equation used in this study is expressed as

$$[1]\qquad h_{ij} = 1.37 + \beta_0(1 - e^{\beta_1 d_{ij}})^{\beta_2} + \varepsilon_{ij}$$

where β_0, β_1, and β_2 are parameters to be estimated; d_{ij} is the diameter of tree j in stand i, and h_{ij} is the height of tree j in stand i; and ε_{ij} is an error term, assumed to be independent between observations and $N(0, \sigma_\varepsilon^2)$. In this equation, $\beta_0 > 0$, β_1, and β_2 represent the asymptotic height, steepness, and curvature, respectively.

The base equation was enhanced by making the asymptotic parameter dependent on the stand-level variables CCFL and BA as follows

$$[2]\qquad h_{ij} = 1.37 + (\beta_{00} + \beta_{01}\mathrm{CCFL}_{ij} + \beta_{02}\mathrm{BA}_i) \times (1 - e^{\beta_1 d_{ij}})^{\beta_2} + \varepsilon_{ij}$$

where CCFL_{ij} is the crown competition factor on larger trees for the jth tree in the ith stand and BA_i is the basal area of the ith stand. Note that CCFL depends on both the value of d for the tree and the value of d for the other trees in the stand. Therefore, it can be considered both a stand- and tree-level variable. The estimation models considered in this study explicitly account for this type of variable. The asymptote depends on stand- and tree-level variables, but all the parameters are fixed effects. The prediction of height for trees in a new stand under this strategy is a straightforward application of either eq. 1 or eq. 2 to the diameter measurements taken in the stand.

Strategy 2 — NMEM with one random stand effect

The relationship between height and diameter is reported to vary by stand (Larsen and Hann 1987; Temesgen and von Gadow 2004), indicating the need for varying parameters among stands. Mixed-model approaches formally incorporate the between-stand variability of the h–d relationship into the model. Estimation of h using NMEM has been previously reported by Lynch et al. (2005) and Castedo et al. (2006).

Regardless of the number of random effects, the NMEM can be motivated as a hierarchical model (Pinheiro and Bates 2000). The h of the jth tree from the ith stand is modeled as

$$[3]\qquad h_{ij} = f(\boldsymbol{\beta}_{ij}, d_{ij}) + \varepsilon_{ij}, \quad i = 1, ..., M, \quad j = 1, ..., n_i, \quad \varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2)$$

where $\boldsymbol{\beta}_{ij}$ is a vector of p stand and possibly tree-specific parameters, f is a real-valued function that relates the height of a tree to its diameter, and ε_{ij} is a within-stand error term that is assumed to be independent and normally distributed. In this study, f is the Chapman–Richards equation (i.e., eqs. 1 and 2).

The kth element of the parameter vector $\boldsymbol{\beta}_{ij}$ is then modeled as a linear function of fixed and random effects

$$[4]\qquad \beta_{ijk} = \boldsymbol{x}'_{ijk}\boldsymbol{\beta}_k + \boldsymbol{z}'_{ijk}\boldsymbol{b}_{ik}, \quad k = 0, ..., p-1, \quad \boldsymbol{b}_{ik} \sim N(\mathbf{0}, \boldsymbol{\Psi}_\mathbf{k})$$

where $\boldsymbol{\beta}_k$ is a vector of fixed effects and $\boldsymbol{b}_{ik}$ is a vector of random effects associated with the ith stand. The number of fixed and random effects does not have to be the same. It is assumed that $\boldsymbol{b}_{ik}$ is normally distributed, with mean $\mathbf{0}$ and covariance matrix $\boldsymbol{\Psi}_\mathbf{k}$, and independent of ε_{ij}. $\boldsymbol{x}_{ijk}$ and $\boldsymbol{z}_{ijk}$ are vectors of covariates associated with the ith stand and possibly with the jth tree. Allowing for these vectors to depend on both the stand and tree within stand permits the use of variables such as CCFL.

In matrix form, eqs. 3 and 4 can be written as

$$\boldsymbol{h}_i = \mathbf{f}(\boldsymbol{\beta}_\mathbf{i}, \boldsymbol{d}_i) + \boldsymbol{\varepsilon}_\mathbf{i}$$

$$\boldsymbol{\beta}_i = \mathbf{X_i}\boldsymbol{\beta} + \mathbf{Z_i}\boldsymbol{b}_i, \quad \boldsymbol{b}_i \sim N(\mathbf{0}, \boldsymbol{\Psi})$$

These two equations can be combined, so that

$$[5]\qquad \boldsymbol{h}_i = \mathbf{f}(\boldsymbol{\beta}, \boldsymbol{b}_i, \mathbf{X_i}, \mathbf{Z_i}) + \boldsymbol{\varepsilon}_\mathbf{i}$$

where $\boldsymbol{h}_i$ is a vector of tree heights from stand i, $\boldsymbol{\beta}_\mathbf{i}$ is a vector of stand-level parameters, $\boldsymbol{d}_i$ is a vector of tree diameters from stand i, $\boldsymbol{\varepsilon}_\mathbf{i}$ is a vector of within-stand error terms, $\boldsymbol{\beta}$ is a vector of fixed parameters that do not depend on the stand, $\boldsymbol{b}_i$ is a vector of stand-level random effects, and $\mathbf{X_i}$ and $\mathbf{Z_i}$ are matrices of explanatory variables. The random effects for the stand enter into eq. 5 nonlinearly, making the model a NMEM.

After examining the between stand variability of the coefficients of the base equations, we added a single random effect to the parameter that controls the model asymptote. When a single random effect is included in the asymptote of the Chapman–Richards function, eq. 4 becomes

$$[6]\qquad \beta_{0ij} = \beta_0 + b_i, \quad b_i \sim N(0, \sigma_b^2)$$

if only diameter is included in the model or

$$[7]\qquad \beta_{0ij} = \beta_{00} + \beta_{01}\mathrm{CCFL}_{ij} + \beta_{02}\mathrm{BA}_i + b_i, \quad b_i \sim N(0, \sigma_b^2)$$

if relative position and stand-density variables are included. Inserting eq. 6 as the asymptote of eq. 1 and eq. 7 as the asymptote of eq. 2 yields

$$[8]\qquad h_{ij} = 1.37 + (\beta_0 + b_i)(1 - e^{\beta_1 d_{ij}})^{\beta_2} + \varepsilon_{ij}$$

and

$$[9]\qquad h_{ij} = 1.37 + (\beta_{00} + \beta_{01}\mathrm{CCFL}_{ij} + \beta_{02}\mathrm{BA}_i + b_i) \times (1 - e^{\beta_1 d_{ij}})^{\beta_2} + \varepsilon_{ij}$$

Predicting heights of trees for a new stand

Suppose that the heights of a subsample of n_m trees from a new stand, not included in the original training data set, are known. Let $\boldsymbol{h_m}$ be the height vector and $\mathbf{X_m}$ and $\mathbf{Z_m}$ be the covariate matrices from those trees. Then, the height of another tree from the same stand can be estimated by first predicting the random effects of the stand, $\boldsymbol{b_m}$, based on the subsample of n_m trees of known height, and then calculating the new height as

$$\widehat{h}_{m,\text{new}} = f(\widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{b}}_{\boldsymbol{m}}, \boldsymbol{x}_{\boldsymbol{m},\mathbf{new}}, \boldsymbol{z}_{\boldsymbol{m},\mathbf{new}})$$

where $\widehat{\boldsymbol{\beta}}$ is the estimated fixed-effects parameter vector, $\widehat{\boldsymbol{b}}_{\boldsymbol{m}}$ the predicted random effects for stand m, and $\boldsymbol{x}_{\boldsymbol{m},\mathbf{new}}$ and $\boldsymbol{z}_{\boldsymbol{m},\mathbf{new}}$ are the vectors of covariates for the new tree from stand m that does not have a measured height. The fixed-effects parameters are estimated from the training data set.

Prediction of the random effects follows directly from Lindstrom and Bates' algorithm for estimating the model parameters (Lindstrom and Bates 1990; Pinheiro and Bates 2000). The NMEM is linearized using a Taylor expansion about the current estimate of $\boldsymbol{\beta}$ and $\boldsymbol{b}$. At each iteration, the model is approximated by the following LMEM:

$$\widehat{\mathbf{h}}_{\mathbf{i}} = \widehat{\mathbf{X}}_{\mathbf{i}}\widehat{\beta} + \widehat{\mathbf{Z}}_{\mathbf{i}}\widehat{\mathbf{b}}_{\mathbf{i}}$$

where the transformed variables $\widehat{\mathbf{h}}_{\mathbf{i}}$, $\widehat{\mathbf{X}}_{\mathbf{i}}$, and $\widehat{\mathbf{Z}}_{\mathbf{i}}$ are defined as

$$[10] \qquad \widehat{\mathbf{h}}_{\mathbf{i}} = \mathbf{h_i} - \mathbf{f}(\mathbf{X_i}, \mathbf{Z_i}, \widehat{\beta}, \widehat{\mathbf{b}}_{\mathbf{i}}) + \widehat{\mathbf{X}}_{\mathbf{i}}\widehat{\beta} + \widehat{\mathbf{Z}}_{\mathbf{i}}\widehat{\mathbf{b}}_{\mathbf{i}}$$

$$[11] \qquad \widehat{\mathbf{X}}_{\boldsymbol{i}} = \frac{\partial f(\mathbf{X_i}, \mathbf{Z_i}, \beta, \mathbf{b_i})}{\partial \beta'}\bigg|_{\widehat{\boldsymbol{\beta}},\widehat{\boldsymbol{b}}_{\mathbf{i}}}$$

$$\widehat{\mathbf{Z}}_{\boldsymbol{i}} = \frac{\partial f(\mathbf{X_i}, \mathbf{Z_i}, \beta, \mathbf{b_i})}{\partial \boldsymbol{b}'_{\boldsymbol{i}}}\bigg|_{\widehat{\boldsymbol{\beta}},\widehat{\boldsymbol{b}}_{\mathbf{i}}}$$

The random effects for a new stand are approximated using an empirical best linear unbiased predictor (BLUP) (Goldberger 1962) on this LMEM approximation at convergence

$$[12] \qquad \widehat{\mathbf{b}}_{\mathbf{m}} \approx \boldsymbol{\Psi}_{\widehat{\boldsymbol{\sigma}}^2}\widehat{\mathbf{Z}}'_{\mathbf{m}}\mathbf{V}^{-1}_{\widehat{\boldsymbol{\sigma}}^2}(\widehat{\mathbf{h}}_{\mathbf{m}} - \widehat{\mathbf{X}}_{\mathbf{m}}\widehat{\beta})$$

where $\widehat{\boldsymbol{\beta}}$ is the estimate of the fixed parameters and $\boldsymbol{\Psi}_{\hat{\boldsymbol{\sigma}}^2}$ is the matrix $\boldsymbol{\Psi}$ evaluated at the estimated variance components $\hat{\boldsymbol{\sigma}}^2$ and

$$\mathbf{V}^{-1}_{\widehat{\boldsymbol{\sigma}}^2} = [\text{Var}(\widehat{\mathbf{h}}_{\mathbf{m}})|_{\widehat{\boldsymbol{\sigma}}^2}]^{-1} = (\widehat{\sigma}^2_{\varepsilon}\mathbf{I_m} + \widehat{\mathbf{Z}}_{\mathbf{m}}\boldsymbol{\Psi}_{\widehat{\boldsymbol{\sigma}}^2}\widehat{\mathbf{Z}}'_{\mathbf{m}})^{-1}$$

where $\mathbf{I_m}$ is the $m \times m$ identity matrix. Both $\widehat{\beta}$ and $\widehat{\boldsymbol{\sigma}}^2$ are estimated from the training data set.

Equation 10 is applied to stand m to obtain $\widehat{\mathbf{h}}_{\mathbf{m}}$. Substituting into eq. 12 yields:

$$[13] \qquad \widehat{\mathbf{b}}_{\mathbf{m}} \approx \boldsymbol{\Psi}_{\widehat{\boldsymbol{\sigma}}^2}\widehat{\mathbf{Z}}'_{\mathbf{m}}\mathbf{V}^{-1}_{\widehat{\boldsymbol{\sigma}}^2}(\boldsymbol{h_m} - \mathbf{f}(\mathbf{X_m}, \widehat{\beta}, \widehat{\mathbf{b}}_{\mathbf{m}}) + \widehat{\mathbf{Z}}_{\mathbf{m}}\widehat{\mathbf{b}}_{\mathbf{m}})$$

Note that the terms involving $\widehat{\mathbf{X}}_{\mathbf{new}}$ cancel out, so that eq. 11 does not have to be calculated. In general, a solution to eq. 13 has to be obtained iteratively.

We will illustrate this process with the Chapman–Richards model with only one random effect and d as the sole explanatory variable (eq. 8). Because there is only one random effect, both b_m and $\Psi = \sigma^2_b$ are scalars, and the calculations simplify considerably. $\widehat{\boldsymbol{Z}}_{\boldsymbol{m}}$ is a vector with jth element.

$$\widehat{z}_{mj} = \frac{\partial f(d_{mj}, \boldsymbol{\beta}, b_m)}{\partial b_m}\bigg|_{\widehat{\boldsymbol{\beta}},\widehat{b}_m} = \frac{\partial(\beta_0 + b_m)(1 - e^{\beta_1 d_{mj}})^{\beta_2}}{\partial b_m}\bigg|_{\widehat{\beta},\widehat{b}_m} = (1 - e^{\widehat{\beta}_1 d_{mj}})^{\widehat{\beta}_2}$$

$$\mathbf{V}^{-1}_{\widehat{\boldsymbol{\sigma}}^2} = (\widehat{\sigma}^2_{\varepsilon}\mathbf{I_m} + \widehat{\sigma}^2_b\widehat{\mathbf{z}}_{\mathbf{m}}\widehat{\mathbf{z}}'_{\mathbf{m}})^{-1}$$

Then, after simplification, eq. 13 becomes

$$[14] \qquad \widehat{b}_m \approx \widehat{\sigma}^2_b\widehat{\mathbf{z}}'_{\mathbf{m}}(\widehat{\sigma}^2_{\varepsilon}\mathbf{I_m} + \widehat{\sigma}^2_b\widehat{\mathbf{z}}_{\mathbf{m}}\widehat{\mathbf{z}}'_{\mathbf{m}})^{-1} \times \begin{bmatrix} h_{m1} - \widehat{\beta}_0(1 - e^{\widehat{\beta}_1 d_{m1}})^{\widehat{\beta}_2} \\ \vdots \\ h_{mn_m} - \widehat{\beta}_0(1 - e^{\widehat{\beta}_1 d_{mn_m}})^{\widehat{\beta}_2} \end{bmatrix}$$

In this case, there exists a closed form solution and $\widehat{b}_m$ can be calculated directly. The height of a new tree from the stand is predicted to be

$$[15] \qquad h_{m,\text{new}} = 1.37 + (\widehat{\beta}_0 + \widehat{b}_m)(1 - e^{\widehat{\beta}_1 d_{m,\text{new}}})^{\widehat{\beta}_2}$$

When relative position and stand-density variables are included, eqs. 14 and 15 have the same form but with $\widehat{\beta}_0$ substituted by $\widehat{\beta}_{00} + \widehat{\beta}_{01}\text{CCFL}_{mj} + \widehat{\beta}_{02}\text{BA}_m$.

Strategy 3 — NMEM with two random stand effects

In addition to the random asymptote, one may consider a random steepness (β_1) or curvature parameter (β_2) or both. The correlation between the steepness parameter and the asymptote is very high, 0.97. As a consequence, modeling all three parameters as random effects results in serious convergence problems and is not likely to result in a significant improvement in predictive performance. We will consider the case of a random asymptote and curvature parameter (correlation = –0.81).

Model formulation follows directly from the general description for strategy 2. Equation 3 with random asymptote and curvature parameters is

$$[16] \qquad h_{ij} = 1.37 + \beta_{0ij}(1 - e^{\beta_1 d_{ij}})^{\beta_{2ij}} + \varepsilon_{ij}$$

If only random effects, no relative position, and stand-density variables are added, eq. 4 becomes

$$\beta_{0i} = \beta_0 + b_{0i}$$

$$\beta_{2i} = \beta_2 + b_{2i}, \quad \mathbf{b_i} \sim N(\mathbf{0}, \boldsymbol{\Psi})$$

where $\boldsymbol{\Psi} = \begin{bmatrix} \sigma^2_{b_0} & \sigma_{b_0 b_2} \\ \sigma_{b_0 b_2} & \sigma^2_{b_2} \end{bmatrix}$ and $\sigma_{b_0 b_2} = \text{Cov}(b_{0i}, b_{2i})$. These equations can be combined to yield the final model.

$$h_{ij} = 1.37 + (\beta_0 + b_{0i})(1 - e^{\beta_1 d_{ij}})^{(\beta_2 + b_{2i})} + \varepsilon_{ij}$$

When the relative position and stand-density variables are added, eq. 4 becomes

$$\beta_{0ij} = \beta_{00} + \beta_{01}\mathrm{CCFL}_{ij} + \beta_{02}\mathrm{BA}_i + b_{0i}$$

$$\beta_{2i} = \beta_2 + b_{2i}, \quad \mathbf{b_i} \sim N(\mathbf{0}, \mathbf{\Psi})$$

Combining those two equations with eq. 16 yields

$$h_{ij} = 1.37 + (\beta_{00} + \beta_{01}\mathrm{CCFL}_{ij} + \beta_{02}\mathrm{BA}_i + b_{0i}) \times (1 - e^{\beta_1 d_{ij}})^{(\beta_2 + b_{2i})} + \varepsilon_{ij}$$

Predicting heights of trees for a new stand

Prediction of the height of a tree from a new stand proceeds as described in the previous section. In this case, there are two random effects (b_{0m}, b_{2m}), and the calculations are more complex. $\widehat{\mathbf{Z}}_\mathbf{m}$ is a $n_m \times 2$ matrix, with the jth row composed of:

$$\widehat{z}_{0mj} = \frac{\partial f(d_{mj}, \boldsymbol{\beta}, \boldsymbol{b_m})}{\partial b_{0m}}\bigg|_{\widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{b}}} = \frac{\partial(\beta_0 + b_{0m})(1 - e^{\beta_1 d_{mj}})^{(\beta_2 + b_{m2})}}{\partial b_{0m}}\bigg|_{\widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{b}}} = (1 - e^{\widehat{\beta}_1 d_{mj}})^{(\widehat{\beta}_2 + \widehat{b}_{2m})}$$

$$\widehat{z}_{2mj} = \frac{\partial f(d_{mj}, \boldsymbol{\beta}, \boldsymbol{b_m})}{\partial b_{m2}}\bigg|_{\widehat{\boldsymbol{\beta}}, \widehat{\mathbf{b}}} = (\widehat{\beta}_0 + \widehat{b}_{0m})(1 - e^{\widehat{\beta}_1 d_{mj}})^{(\widehat{\beta}_2 + \widehat{b}_{2m})}\log(1 - e^{\widehat{\beta}_1 d_{mj}})$$

$$\widehat{\mathbf{Z}}_\mathbf{m} = \begin{bmatrix} (1 - e^{\widehat{\beta}_1 d_{m1}})^{(\widehat{\beta}_2 + \widehat{b}_{2m})} & (\widehat{\beta}_0 + \widehat{b}_{0m})(1 - e^{\widehat{\beta}_1 d_{m1}})^{(\widehat{\beta}_2 + \widehat{b}_{2m})}\log(1 - e^{\widehat{\beta}_1 d_{m1}}) \\ \vdots & \vdots \\ \left(1 - e^{\widehat{\beta}_1 d_{mn_m}}\right)^{(\widehat{\beta}_2 + \widehat{b}_{2m})} & (\widehat{\beta}_0 + \widehat{b}_{0m})\left(1 - e^{\widehat{\beta}_1 d_{mn_m}}\right)^{(\widehat{\beta}_2 + \widehat{b}_{2m})}\log(1 - e^{\widehat{\beta}_1 d_{mn_m}}) \end{bmatrix}$$

$\mathbf{V}^{-1}_{\hat{\boldsymbol{\sigma}}^2}$ is a $n_m \times n_m$ matrix.

$$\mathbf{V}^{-1}_{\widehat{\boldsymbol{\sigma}}^2} = \left[\widehat{\sigma}^2_\varepsilon \mathbf{I_m} + \widehat{\mathbf{Z}}_\mathbf{m} \begin{bmatrix} \widehat{\sigma}^2_{b_0} & \widehat{\sigma}_{b_0 b_2} \\ \widehat{\sigma}_{b_0 b_2} & \widehat{\sigma}^2_{b_2} \end{bmatrix} \widehat{\mathbf{Z}}'_\mathbf{m}\right]^{-1}$$

After some calculation, the random effects ($\widehat{b}_{0m}, \widehat{b}_{2m}$) can be approximated by iteratively solving the following equation.

[17]
$$\begin{bmatrix} \widehat{b}_{0m} \\ \widehat{b}_{2m} \end{bmatrix} \approx \mathbf{\Psi}_{\widehat{\boldsymbol{\sigma}}^2} \widehat{\mathbf{Z}}'_\mathbf{m} \boldsymbol{V}^{-1}_{\widehat{\boldsymbol{\sigma}}^2} \begin{bmatrix} h_{m1}\left[\widehat{\beta}_0 - \widehat{b}_{2m}\left(\widehat{\beta}_0 + \widehat{b}_{0m}\right)\log\left(1 - e^{\widehat{\beta}_1 d_{m1}}\right)\right]\left(1 - e^{\widehat{\beta}_1 d_{m1}}\right)^{\left(\widehat{\beta}_2 + \widehat{b}_{2m}\right)} \\ \vdots \\ h_{mm}\left[\widehat{\beta}_0 - \widehat{b}_{2m}\left(\widehat{\beta}_0 + \widehat{b}_{0m}\right)\log\left(1 - e^{\widehat{\beta}_1 d_{mm}}\right)\right]\left(1 - e^{\widehat{\beta}_1 d_{mm}}\right)^{\left(\widehat{\beta}_2 + \widehat{b}_{2m}\right)} \end{bmatrix}$$

The height of a new tree from this stand is predicted to be

[18]
$$h_{m,\mathrm{new}} = 1.37 + (\widehat{\beta}_0 + \widehat{b}_{0m})(1 - e^{\widehat{\beta}_1 d_{m,\mathrm{new}}})^{(\widehat{\beta}_2 + \widehat{b}_{2m})}$$

When relative position and stand-density variables are included, eqs. 17 and 18 are the same but with $\widehat{\beta}_0$ above substituted by $\widehat{\beta}_{00} + \widehat{\beta}_{01}\mathrm{CCFL}_{mj} + \widehat{\beta}_{02}\mathrm{BA}_m$.

Strategy 4 — Adjusting the NFEM to measurements from a new stand

Suppose again that the height of a subsample of n_m trees from a new stand is known. Let $\boldsymbol{h_m}$ be the height and $\mathbf{X_m}$ and $\mathbf{Z_m}$ be the matrices of covariates from those trees. Then, the height of another tree from the same stand can be adjusted with the following OLS correction factor on the regional height–diameter equations (Draper and Smith 1998, p. 225):

$$k^*_m = \frac{\sum_{j=1}^{n_{ij}}\left[\left(\widehat{h}_{ij} - 1.37\right)\left(h_{ij} - 1.37\right)\right]}{\sum_{j=1}^{n_{ij}}\left(\widehat{h}_{ij} - 1.37\right)^2}$$

where k^*_m is the correction factor, $\hat{h}_{ij}$ is the predicted height from eq. 1 or 2, and h_{ij} is the observed height. Then, the ad-

Table 2. Cross-validation bias and root mean square error by prediction method and model form.

	Diameter model only		Diameter, CCFL, and BA models	
Prediction method	Bias (m)	RMSE (m)	Bias (m)	RMSE (m)
NFEM (strategy 1)	–0.13	3.76	–0.04	3.25
NMEM with one random effect (strategy 2)	–0.01	2.50	0.00	2.45
NMEM with one random effect, fixed component*	0.88	4.11	0.18	3.31
NMEM with two random effects (strategy 3)	0.01	2.36	0.00	2.36
NMEM with two random effects, fixed component*	0.88	4.20	0.23	3.32
Correction factor (strategy 4)	0.25	2.72	0.06	2.45

Note: NFEM, nonlinear fixed effects model; NMEM, nonlinear mixed effects model; CCFL, crown competition factor on larger trees; and BA, basal area of the stand.

*The fixed component of the NMEM was calculated by setting the random effects to 0.

justed predicted height for a tree from the stand can be calculated to be

$$[19]\qquad \widehat{h}_{m,\text{new}} = 1.37 + k_m^* \widehat{\beta}_0(1 - e^{\widehat{\beta}_1 d_{m,\text{new}}})^{\widehat{\beta}_2}$$

If relative position and stand-density variables are included, the equation becomes

$$[20]\qquad \widehat{h}_{m,\text{new}} = 1.37 + k_m^*(\widehat{\beta}_{00} + \widehat{\beta}_{01}\text{CCFL}_{mj} + \widehat{\beta}_{02}\text{BA}_m) \times (1 - e^{\widehat{\beta}_1 d_{m,\text{new}}})^{\widehat{\beta}_2}$$

Model fitting and evaluation of predictive performance

All strategies were fitted in the standard R software language (available from www.r-project.org), using the nlme package for fitting NMEMs (strategies 2 and 3), and the nlreg package for fitting NFEMs (strategies 1 and 4). To find a global minimum, the starting value of each parameter was varied and several runs were obtained.

The predictive performance of the different strategies was evaluated as follows.

1. A stand was selected for evaluation and the remaining 141 stands were used to estimate the regional parameters of the alternative models using either NFEM or NMEM.
2. A subsample of heights from the evaluation stand was selected following different criteria. The random stand effects (strategies 2 and 3 and eqs. 14 and 17) and the OLS correction factor (strategy 4 and eq. 19) were calculated from this subsample.
3. The height of the remaining trees in the evaluation stand was predicted and the prediction error, $(h_{ij} - \widehat{h}_i)$, calculated. Note that the predicted height was obtained from a set of trees different from those used to fit the model.
4. The process was repeated for all stands to estimate the prediction RMSE and prediction bias. Because the main interest is in predicting heights from a stand, estimates of the RMSE and bias were obtained first at the stand level and then averaged over the n stands.

$$\text{Bias} = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{m_i}\sum_{j=1}^{m_i}(h_{ij} - \widehat{h}_{ij})\right]$$

$$\text{RMSE} = \frac{1}{n}\sum_{i=1}^{n}\sqrt{\frac{1}{m_i}\sum_{j=1}^{m_i}(h_{ij} - \widehat{h}_{ij})^2}$$

where h_{ij} is the measured height of the jth tree from the ith stand, $\widehat{h}_{ij}$ is its predicted height, m_i is the number of trees from the ith stand for which height was predicted, and n is the number of stands.

5. The process was repeated 200 times. The prediction RMSE and bias calculated at each iteration were then averaged across all 200 iterations.

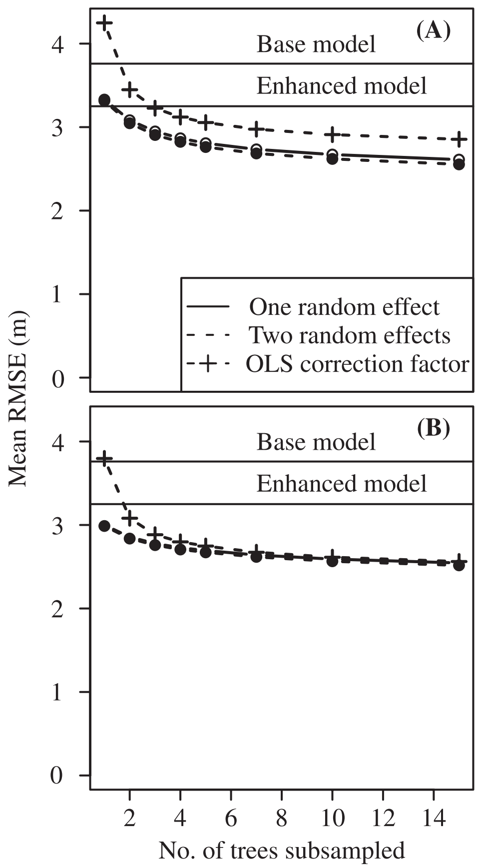
The following subsample selection procedures were used in step 2:

1. Random selection of heights: the subsample sizes evaluated varied from 1 to 15 heights. Because we did not correct for the different inclusion probabilities associated with the use of nested fixed and variable radius subplots, the random-subsampling approach we used will tend to include larger trees from the stand more frequently than smaller trees.
2. Selection of dominant heights, often collected for site index estimation: the heights of the largest one, two, three, or four trees in the stand, based on diameter, were selected and the prediction RMSE calculated for each of the four subsample sizes. In this case, only steps 1–4 above were used, since there was no possibility for repetition when large trees were selected. In selecting the largest one, two, three, or four trees in the stand, there were very few stands in which two candidate trees exhibited the same diameter. When this occurred, the height of one tree was randomly chosen for the subsample.

If an estimate of the height of the trees from a new stand is desired but a subsample of heights from that stand is not available, then neither the random effects nor the OLS correction factor can be calculated. As a result, the only strategies available are to use the regional height–diameter equations from strategy 1 or to use the equations from strategy 2 or 3 with only their fixed effects parameters — that is, setting the random stand effects, $\mathbf{b_m}$, equal to zero.

The performance and appropriateness of these approaches was evaluated using leave-one-out cross-validation (Stone 1974). At each iteration, one stand was excluded from the data set and models were fitted to the remaining stands. The various height–diameter models were then used to predict the height of all the trees in the excluded stand. The same process was repeated for every stand in the data set. Because trees from the same stand tend to be correlated, excluding an entire stand to examine the performance of the models provides stronger model evaluation if prediction of heights for trees from a new stand not included in the original data set is desired (Monleon et al. 2004). The leave-one-

Fig. 1. Predictive performance (root mean square error, RMSE, (*m*)) for the four strategies as a function of the number of heights subsampled at random to predict the stand random effects (strategies 2 and 3) or ordinary least-squares (OLS) correction factor (strategy 4). For comparison, the cross-validation results from strategy 1 are included as two horizontal lines. (A) Models using diameter only and (B) models using diameter, crown competition factor in larger trees (CCFL), and basal area (BA).



out cross-validation statistics was summarized and used to evaluate the performance of these approaches.

Results and discussion

Q.1. Will the accuracy and precision of height prediction vary by the type of prediction strategy? Is there substantial improvement in accuracy and precision by predicting height using NMEM over the conventional NFEM strategy?

Substantial differences were found among the predictive abilities of the alternative strategies examined for developing regional height–diameter equations. For the NFEM, the cross-validation RMSE of the base model form (i.e., with diameter only) was 3.76 m (Table 2). The enhanced model, which included the stand-level variables, resulted in a decrease in RMSE of 0.51 m (13.3%). The bias was also reduced from –0.13 to –0.04 m. The results are in agreement with those of several recent studies that have included relative position and (or) stand density in the base regional height–diameter equation (e.g., Zeide and Vanderschaaf 2001; Temesgen and von Gadow 2004; Sharma and Zhang 2004; Castedo et al. 2006).

Table 3. Mean bias (m) as a function of the number of heights subsampled to predict the stand random effects (strategies 2 and 3) or ordinary least-squares (OLS) correction factor (strategy 4). The subsample of heights from the stand was selected at random.

	Mean bias (diameter model only) (m)			Mean bias (diameter, CCFL, and BA models) (m)		
	Strategy			Strategy		
Subsample size (*hs*)	2	3	4	2	3	4
1	0.00	0.01	–0.54	0.05	0.06	–0.34
2	–0.07	–0.06	–0.05	0.01	0.01	–0.03
3	–0.08	–0.07	0.07	–0.01	–0.01	0.02
4	–0.08	–0.07	0.12	–0.02	–0.02	0.02
5	–0.06	–0.05	0.15	–0.01	–0.01	0.04
7	–0.05	–0.05	0.18	–0.02	–0.02	0.04
10	–0.04	–0.03	0.21	–0.02	–0.02	0.04
15	–0.03	–0.03	0.22	–0.01	–0.02	0.05

Note: CCFL, crown competition factor in larger trees; and BA, basal area of the stand.

When a subsample of heights was available to predict the random effects, the predictive performance of the NMEM was substantially better than that of the NFEM (Table 2). Although a more detailed analysis follows, when the heights of all the sample trees in a stand were used to estimate a single random stand effect using eq. 13, the RMSE decreased by 1.36 m (36%) compared with the NFEM. Adding a second random effect to the curvature parameter of the Chapman–Richards equation, and therefore increasing the complexity of the model, only produced a marginal additional gain of 0.04 m. The bias of the NMEM was negligible in both cases. The OLS correction factor, calculated with the heights of all the available sample trees from the stand, also improved the precision of the height prediction, although not as much as the NMEM (1.04 m, 28%). However, it resulted in substantial bias. Note that using the heights of all the sample trees in a stand to calibrate the model and to assess its predictive performance is not reasonable from an application point of view. However, the values provided in Table 2 for NFEM, NMEM, and the OLS correction factor provide limiting values to evaluate the gains in performance when using a subsample of heights. They also allow for comparison between methods when using all the information available.

One explanation for the improvement in accuracy and precision of NMEM over NFEM with the use of a subsample is that NFEM ignores the clustered structure of the data, assuming that all trees are independent, and produces an overall regional height–diameter equation, while NMEM acknowledges this structure of the data and produces a regional equation that also characterizes individual stand height–diameter curves. This, in turn, allows NMEM greater flexibility in describing the variance and covariance structure and account for the within and between stand height–diameter variations.

If a subsample of heights is not available to predict the stand random effects, then the random effects are usually set to zero and the model with only the fixed effect parame-

Fig. 2. Frequency distribution of the root mean square error (RMSE) after 200 simulations for the mixed model with one random effect and diameter alone (strategy 2), as a function of the number of randomly selected heights used to predict the random effect. For comparison, the cross-validation RMSE for the nonlinear fixed effects model (NFEM, grey lines) and nonlinear mixed effects model (NMEM) based on all heights (black lines) (Table 2) are included. Broken lines represent the RMSE of the base model (diameter only) and light solid lines represent the RMSE of the enhanced model (diameter, crown competition factor in larger trees (CCFL), and basal area of the stand (BA)).

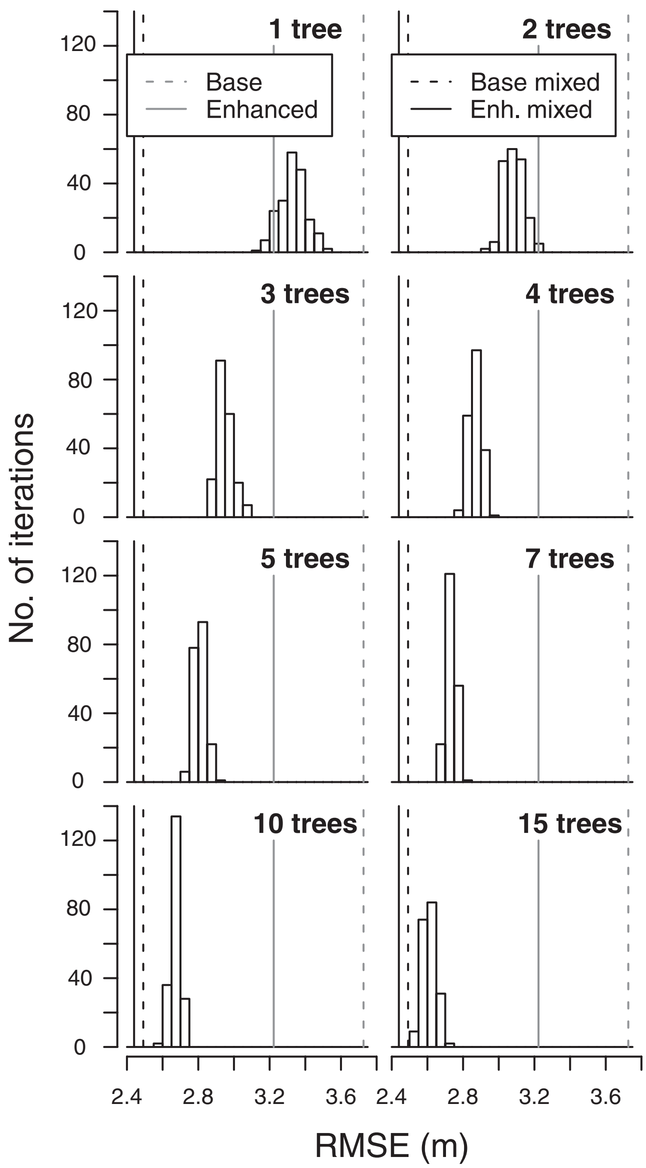
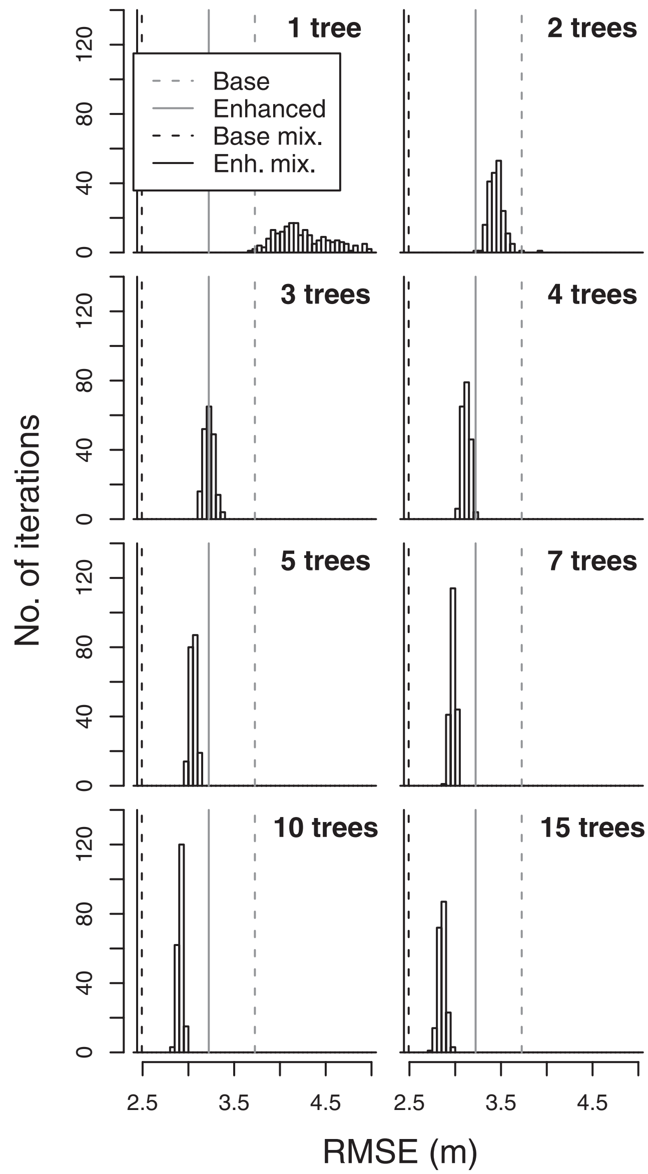


Fig. 3. Frequency distribution of the root mean square error (RMSE) after 200 simulations for the nonlinear fixed effects model (NFEM) with the ordinary least-squares (OLS) correction factor and diameter alone (strategy 4), as a function of the number of randomly selected heights used to predict the random effect. For comparison, the cross-validation RMSE for the NFEM (grey lines) and nonlinear mixed effects model (NMEM) based on all heights (black lines) (Table 2) are included. Broken lines represent the RMSE of the base model (diameter only) and solid lines represent the RMSE of the enhanced model (diameter, crown competition factor in larger trees (CCFL), and basal area of the stand (BA)).



ters is used. Using a mixed-effects model in this case can result in a substantial decrease in predictive performance (Monleon 2003). In this study, the bias increased to 0.88 m and the RMSE to more than 4 m (Table 2). A similar decrease in performance was reported by Monleon (2003) and Robinson and Wykoff (2004) and noted by Vonesh and Chinchilli (1997; p. 295). Adding relative position and stand-density measures improved the performance of the NMEM with random effects set to zero, but it remained worse than that of the simpler NFEM with relative position and stand-density measures. Therefore, we would not recommend the use of mixed models when a subsample of heights from the stand is not available.

Q.2. How does the predictive ability of localizing NMEM and NFEM from a subsample of height in the stand vary by subsampling design and subsample size?

For the NMEM base model, increasing the number of subsampled heights used to estimate the random stand ef-

Fig. 4. Frequency distribution of the root mean square error (RMSE) after 200 simulations for the mixed model with one random effect and diameter, crown competition factor in larger trees (CCFL), and basal area of the stand (BA) (strategy 2), as a function of the number of randomly selected heights used to predict the random effect. For comparison, the cross-validation RMSE for the nonlinear fixed effects model (NFEM, grey lines) and the nonlinear mixed effects model (NMEM) based on all heights (black lines) (Table 2) are included. Broken lines represent the RMSE of the base model (diameter only) and solid lines represent the RMSE of the enhanced model (diameter, CCFL, and BA).

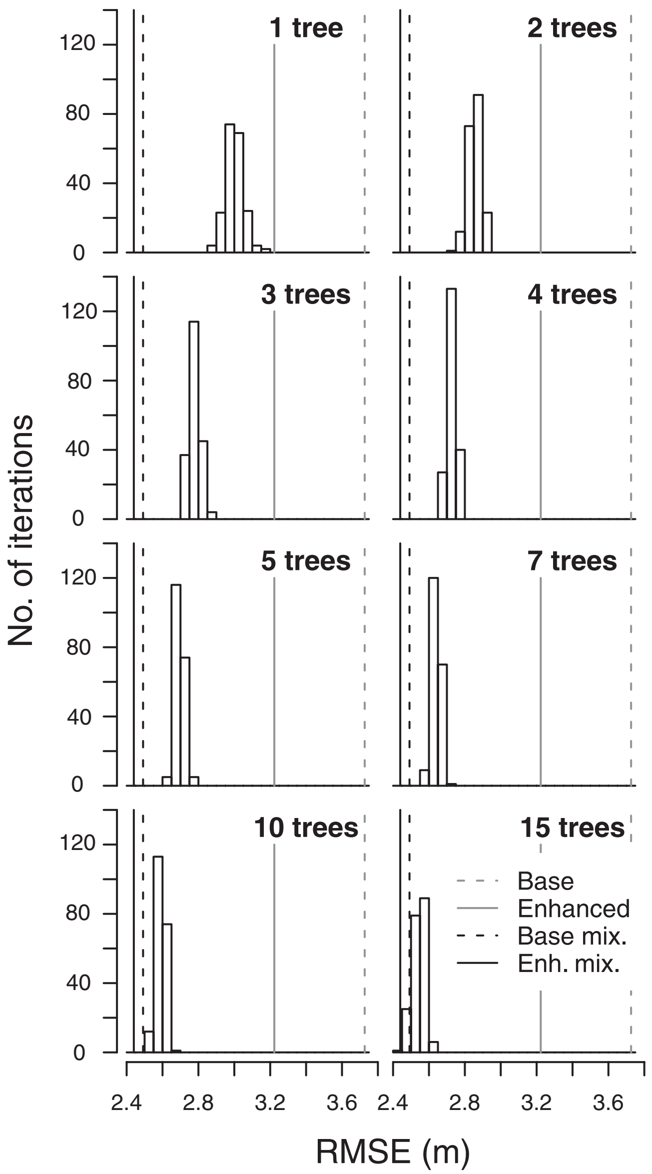
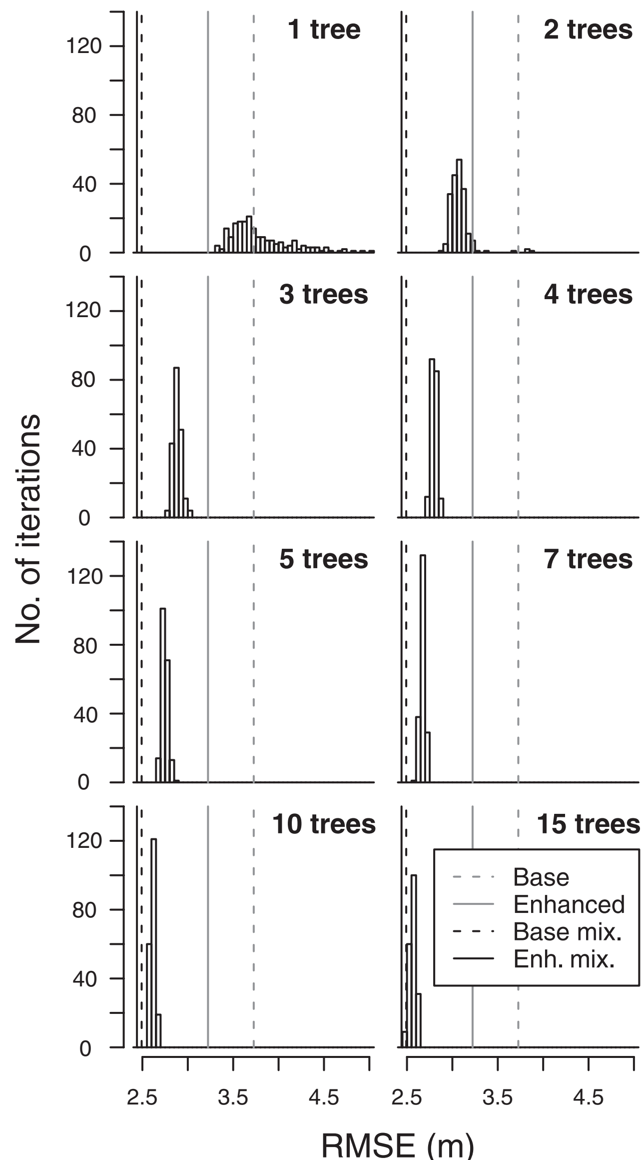


Fig. 5. Frequency distribution of the root mean square error (RMSE) after 200 simulations for the the nonlinear mixed effects model (NFEM) with the ordinary least-squares (OLS) correction factor and diameter, crown competition factor in larger trees (CCFL), and and basal area of the stand (BA) (strategy 4), as a function of the number of randomly selected heights used to predict the random effect. For comparison, the cross-validation RMSE for the NFEM (grey lines) and the nonlinear mixed effects model (NMEM) based on all heights (black lines) (Table 2) are included. Broken lines represent the RMSE of the base model (diameter only) and solid lines represent the RMSE of the enhanced model (diameter, CCFL, and BA).



fects from one height up to 15 heights reduced prediction RMSE (Fig. 1), while the bias remained negligible (Table 3). The largest decrease in prediction RMSE between the NFEM and NMEM strategies was when only one height from the stand was subsampled: on average, the RMSE decreased from 3.76 to 3.33 m (11%). More importantly, after 200 simulations, the RMSE for the NMEM was never greater than that of the NFEM (Fig. 2): even when only one height was subsampled, the RMSE remained below 3.55 m, compared with 3.76 m for the NFEM. The relative gains declined as the number of heights subsampled increased. However, after 15 heights were subsampled from each stand, the RMSE was reduced to 2.61 m, on average, an improvement of 30% over the NFEM strategy. This value was very close to 2.5 m, the RMSE obtained when the heights of all the trees in the sample were used to both estimate the random effects and calculate the RMSE (Table 2). Adding a second random effect did not significantly improve the predictive performance of the model over that of a single random effect (Fig. 1).

Table 4. Predictive performance (root mean square error, RMSE, m) as a function of the number of heights from the largest diameter trees used to predict the random effects (strategies 2 and 3) or the ordinary least-squares (OLS) correction factor (strategy 4).

	Diameter model only			Diameter, CCFL, and BA models		
	Strategy			Strategy		
No. of heights	2	3	4	2	3	4
1	3.10	3.07	3.68	2.93	2.92	3.18
2	2.96	2.93	3.50	2.81	2.81	2.97
3	2.87	2.85	3.42	2.74	2.73	2.86
4	2.81	2.80	3.36	2.69	2.68	2.77

Note: CCFL, crown competition factor in larger trees; and BA, basal area of the stand.

The substantial improvement in performance with just one subsampled height is in accordance with the results reported by Monleon (2003), who found an even greater improvement for a LMEM (20%). However, Castedo et al. (2006) found a much smaller reduction in RMSE, even after several heights were used to estimate the random effects. This can be ascribed to the fact that a small proportion of the total variability in their data was between stands rather than within stands. Further, those authors included the height of the dominant tree as a fixed effect in their model and, therefore, less additional information may be obtained from a subsample of heights from the stand. Obviously, application of their models to a new stand always requires a measurement of at least one height, that of the dominant tree.

Strategy 4 applied to the base model resulted in biased height estimates (Table 2), though the size of the bias for subsamples of two or larger are within the measurement precision for most tree sizes found for this data set (Larsen et al. 1987). For the same number of subsampled heights used to estimate the random effects, the OLS correction factor to the base model resulted in higher average prediction RMSE and greater variability among the 200 simulations (Figs. 1 and 3). Strategy 4 performed poorly when only one or two heights were subsampled. For the expanded model, strategy 4 produced nearly unbiased height estimates if more than one height was subsampled. While the prediction RMSE was always higher with the OLS correction factor versus the random effect models, the difference was close to 0.1 m for the expanded model when four or more heights were subsampled, again within the measurement precision for most tree sizes found by Larsen et al. 1987. Furthermore, the NFEM with OLS correction factor on the expanded model was almost as precise and accurate as the NMEM with one random effect and diameter only when subsample sizes were four or larger.

The frequency distributions of the prediction RMSE from the 200 simulations for different random subsample sizes are given in Figs. 2 and 4 for the base and enhanced model forms fit using strategy 2 and in Figs. 3 and 5 using strategy 4. The histograms for strategy 3 were very similar to those of strategy 2, so they are not included. For the 200 simulations, the empirical distributions narrowed as the size of the subsample increased.

Examination of Figs. 3 and 5 shows that the application of strategy 4 when only one height has been randomly subsampled produces prediction RMSE values that are nearly always larger than the residual RMSE from the uncorrected NFEM of strategy 1. While application of strategy 4 when two heights have been randomly selected often produces prediction RMSE values lower than the residual RMSE, the strategy at times still produces values larger than the residual RMSE. We conclude that strategy 4 should not be applied when only one or two heights have been randomly selected from a stand for height measurement.

Table 5. Estimated parameters and associated standard errors for models with diameter (d) only and with stand-level variables.

Strategy 1		Strategy 2		Strategy 3	
Estimate	SE	Estimate	SE	Estimate	SE
Model with diameter only					
β_0	51.9954	β_0	40.4218	β_0	41.8199
β_1	–0.0208	β_1	–0.0276	β_1	–0.0241
β_2	1.0182	β_2	0.936	β_2	0.8604
σ_ε	4.029	σ_b	6.544	σ_{b_0}	6.409
		σ_ε	2.693	σ_{b_2}	0.165
				$\sigma_{b_0b_2}$	0.31
				σ_ε	2.574
Model with stand-level variables					
β_{00}	43.7195	β_{00}	32.4635	β_{00}	35.7419
β_{01}	0.0644	β_{01}	0.0363	β_{01}	0.0431
β_{02}	0.128	β_{02}	0.2585	β_{02}	0.2447
β_1	–0.0194	β_1	–0.021	β_1	–0.0184
β_2	1.0805	β_2	0.9906	β_2	0.961
σ_ε	3.519	σ_b	4.635	σ_{b_0}	4.743
		σ_ε	2.641	σ_{b_2}	0.098
				$\sigma_{b_0b_2}$	0.125
				σ_ε	2.566

Selecting the heights of the largest trees in the stand when predicting random effects, as opposed to a random subsample, was advantageous (Table 4). For the base model and both strategies 2 and 3, the prediction RMSE when the height of the largest tree was selected was approximately 0.2 m less than when a single height was selected at random, but this difference was reduced to 0.07 m when the enhanced models were used. Given that the original sample was selected with an angle gauge, thus increasing the proportion of larger trees, we would expect that the difference would have been greater if a random subsample had been conducted on a sample tree list that had been selected proportional to tree frequency. As the number of heights of the largest trees subsampled increased, the difference in performance between large-tree and random selection strategies decreased. The difference was negligible when the heights of more than four large trees were subsampled. Both Calama and Montero (2004) and Castedo et al. (2006) found that for their models, selecting the heights of the largest trees resulted in a greater RMSE than selecting heights at random. However, those authors included the height of the dominant tree as a fixed effect, thus making their use for estimating the random effects redundant. For strategy 4, the advantage of the large-tree selection strategy was not that clear. While there was a substantial decrease in the mean RMSE when the height of the largest tree was selected, compared with a random height, this was not the case when the heights of two or more large trees were selected for the subsample (Table 4).

For the same model form and subsample size, the OLS correction factor of strategy 4 produced mean prediction RMSEs and ranges in the prediction RMSEs about their means that were consistently larger than the values derived from using the random effects of strategy 2 (Tables 2 and 4). In particular, a subsample size of one resulted in notably large differences. Increasing the number of randomly selected heights used to estimate the OLS correction factor did substantially reduce the prediction RMSE for subsamples up to four heights.

Q.3. Will the prediction of the random effects in NMEM efficiently explain the differences in the height–diameter relationship due to stand-related effects, such as variations in relative position of trees and stand density?

Predicting a random stand effect from a subsample of known heights can be seen as an empirical surrogate for measuring stand-level variables that affect the height–diameter relationship. Therefore, we can compare the performance of the mixed effects models with that of the model with CCFL and BA to study the effectiveness of those two approaches. When only one height was selected at random from the stand, the mean RMSE was 3.33 m, compared to 3.25 m for the enhanced model (strategy 1). However, if the single height was from a large (dominant) tree, the RMSE was 3.10 m, substantially less than that of the model that included relative position and stand-density measures. More interestingly, when two heights were subsampled at random, the average RMSE from all 200 simulations was less that that of the enhanced model (Fig. 1). This result seems to suggest that predicting a random effect parameter is a very effective way of incorporating stand effects. Many of the variables that could influence the height–diameter relationship may not be known or may not be practical to measure, but their effect may be captured by measuring a small subsample of heights from the stand, thus substantially improving the predictive performance of the models.

Including both random effects and stand-level variables had a relatively small effect on the RMSE, compared with including random effects and diameter alone (Tables 2 and 5 and Fig. 1). However, inclusion of the stand variables was advantageous when a small number of heights were subsampled, especially when only one random effect was included (Fig. 1). Table 5 provides the parameter estimates obtained using the three strategies.

When three or fewer heights were randomly subsampled, the prediction of the random effects parameter efficiently explained the differences in the height–diameter relationship due to variations in relative position of trees and stand density (Fig. 1). The use of NMEM is more robust and practical than including relative position and stand-density measures in estimating height.

Conclusions and recommendations

We recommend the following strategy when either developing a new regional height–diameter equation that will be applied to stands not in the modelling data set or when applying a previously parameterized regional height–diameter equation to a new stand.

1. The NFEM with the expanded (including BA and CCFL) model form should be used for predicting missing heights of trees in a new stand for which a subsample of height measurements is not available.
2. When a subsample of heights from the new stand is available, the use of modelling strategies that incorporate random stand effects is advantageous. The predictive performance was superior to that of fixed-effects models, even when relative position and stand-density measures were included in the model. When the subsample is small, it is beneficial to select the height of dominant trees to estimate the random effects. The difference in predictive performance between one or two random parameters was negligible.
3. In general, mixed-effects models performed better than an OLS adjustment (strategy 4). However, the predictive performance of strategy 4 with relative position and stand-density measures is comparable to that of a mixed model with only one random effect when four or more trees were subsampled for height.

The results of this study are sufficiently convincing to advocate the use of NMEM estimates for predicting missing heights when a subsample of them is available. In some circumstances, the use of NFEM and the OLS correction factor of strategy 4 is a viable option for reducing prediction RMSE, particularly when only a NFEM is available for application to a new stand with a subsample of four or more heights.

Acknowledgements

We thank Drs. D. Marshall and D. Maguire for organizing a workshop on the mixed-effects model in 2004 and Professor T. Gregoire for leading the workshop. We also thank Dr. James Flewelling, Bianca Eskelson, and Aaron Weiskittel for their insights and comments on an early draft. We are grateful to three anonymous referees and the Associate Editor for their constructive comments and suggestions.

References

Aguirre, O., Hui, G., von Gadow, K., and Jiménez, J. 2003. An analysis of spatial forest structure using neighbourhood-based variables. For. Ecol. Manage. **183**: 137–145. doi:10.1016/S0378-1127(03)00102-6.

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. doi:10.1139/x03-199.

Castedo, F., Dieguez-Aranda, U., Barrio, M., Sanchez, M.R., and von Gadow, K. 2006. A generalized height–diameter model including random components for radiata pine plantations in northwestern Spain. For. Ecol. Manage. **229**: 202–213. doi:10.1016/j.foreco.2006.4.028.

Clutter, J.L., Fortson, J.C., Pienaar, L.V., Brister, G.H., and Bailey, R.L. 1983. Timber management: a quantitative approach. John Wiley & Sons, New York.

Curtis, R.O. 1967. Height–diameter and height–diameter–age equations for second-growth Douglas-fir. For. Sci. **13**: 365–375.

Dubrasich, M.E., Hann, D.W., and Tappeiner, J.C. 1997. Methods for evaluating crown area profiles of forest stands. Can. J. For. Res. **27**: 385–392. doi:10.1139/x96-206.

Draper, N.R., and Smith, H. 1998. Applied regression analysis. 3rd ed. John Wiley & Sons, New York.

Flewelling, J.W., and de Jong, R.D. 1994. Considerations in simultaneous curve fitting for repeated height-diameter measurements. Can. J. For. Res. **24**: 1408–1414. doi:10.1139/x94-181.

Garman, S.L., S.A Acker, J.L. Ohmann, and T.A. Spies. 1995. Asymptotic height–diameter equations for twenty-four tree species in Western Oregon. Oregon State Univ. For. Res. Lab. Res. Cont. 10.

Goldberger, A.S. 1962. Best linear unbiased prediction in the generalized linear regression model. J. Am. Stat. Assoc. **57**: 369–375. doi:10.2307/2281645.

Gregoire, T.G., Schabenberger, O., and Barrett, J.P. 1995. Linear modelling of irregularly spaced, unbalanced, longitudinal data from permanent-plot measurements. Can. J. For. Res. **25**: 137–156. doi:10.1139/x95-017.

Hann, D.W. 2005. ORGANON [computer program] user's manual. Version 8.0. Department of Forest Resources, Oregon State University, Corvallis, Oregon.

Hann, D.W., D.D. Marshall, and M.L. Hanus. 2003. Equations for predicting height-to-crown-base, 5-year diameter-growth rate, 5-year height-growth rate, 5-year mortality rate, and maximum size-density trajectory for Douglas-fir and western hemlock in the coastal region of the Pacific Northwest. Oregon State Univ. For. Res. Lab. Res. Cont. 40.

Hanus, M.L., D.W. Hann, and D.D. Marshall. 1999*a*. Predicting height for undamaged and damaged trees in southwest Oregon. Oregon State Univ. For. Res. Lab. Res. Cont. 27.

Hanus, M.L., D.D. Marshall, and D.W. Hann. 1999*b*. Height–diameter equations for six species in the coastal regions of the Pacific Northwest. Oregon State Univ. For. Res. Lab. Res. Cont. 25.

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. doi:10.1139/x92-172.

Krajicek, J., Brinkman, K.A., and Gingrich, S.F. 1961. Crown competition – a measure of density. For. Sci. **7**: 35–42.

Krumland, B.E., and Wensel, L.C. 1978. Generalized height–diameter equations for coastal conifers. CO-OP Redwood Yield Research Project, Department of Forestry and Conservation, University of California, Berkeley, California. Res. Note 8.

Landsberg, J.J., and Waring, R.H. 1997. A generalized model of forest productivity using simplified concepts of radiation use efficiency, carbon balance, and partitioning. For. Ecol. Manage. **95**: 209–228. doi:10.1016/S0378-1127(97)00026-1.

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

Larsen, D.R., and D.W. Hann. 1987. Height–diameter equations for seventeen tree species in southwest Oregon. Oregon State Univ. For. Res. Lab. Res. Pap. 49.

Larsen, D.R., Hann, D.W., and Stearns-Smith, S.C. 1987. Accuracy and precision of the tangent method of measuring tree height. West. J. Appl. For. **2**: 26–28.

Larson, P. 1963. Stem form development of forest trees. For. Sci. Mongr. No. 5.

Lindstrom, M.J., and Bates, D.M. 1990. Nonlinear mixed effects models for repeated measurements data. Biometrics, **46**: 673–687. doi:10.2307/2532087. PMID:2242409.

Lynch, T.B., Holley, A.G., and Stevenson, D.G. 2005. A random-parameter height–diameter model for cherrybark oak. South. J. Appl. For. **29**: 22–26.

Mehtätalo, L. 2004. A longitudinal height-diameter model for Norway spruce in Finland. Can. J. For. Res. **34**: 131–140. doi:10.1139/x03-207.

Monleon, V.J. 2003. A hierarchical linear model for tree height prediction. *In*: Proceedings of the 2003 Meeting of the American Statistical Association, Section on Statistics and the Environment, 3–7 August 2003, San Franciscon, California. The American Statiscal Association, Alexandria, Virginia. pp. 2865–2869.

Monleon, V.J., Azuma, D., and Gedney, D. 2004. Equations for predicting uncompacted crown ratio based on compacted crown ratio and tree attributes. West. J. Appl. For. **19**: 260–267.

Neter, J., Wasserman, W., and Kutner, M.H. 1990. Applied linear regression models. Richard D. Irwin, Boston.

Paine, D.P., and D.W. Hann. 1982. Maximum crown width equations for southwest Oregon tree species. Oregon State Univ. For. Res. Lab. Res. Pap. 46.

Pinheiro, J.C., and Bates, D.M. 2000. Mixed effects models in S and S-PLUS. Springer, New York.

Ratkowsky, D.A. 1990. Handbook of nonlinear regression. Marcel and Dekker Inc., New York.

Richards, F.J. 1959. A flexible growth function for empirical use. J. Exp. Bot. **10**: 290–300. doi:10.1093/jxb/10.2.290.

Ritchie, M.W., and D.W. Hann. 1987. Equations for predicting height to crown base for fourteen tree species in southwest Oregon. Oregon State Univ. For. Res. Lab. Res. Pap. 50.

Robinson, A., and Wykoff, W.R. 2004. Imputing missing height measures using a mixed-effects modeling strategy. Can. J. For. Res. **34**: 2492–2500. doi:10.1139/x04-137.

Running, S.W., and Coughlan, J.C. 1988. A general model of forest ecosystem processes for regional applications. I. Hydrologic balance, canopy gas exchange, and primary production processes. Ecol. Model. **41**: 125–154.

Sharma, M., and Zhang, S.Y. 2004. Height–diameter models using stand characteristics for *Pinus banksiana* and *Picea mariana*. Scand. J. For. Res. **19**: 442–451.

Stone, M. 1974. Cross-validatory choice and assessment of statistical predictions. J. R. Stat. Soc. [Ser A], **36**: 111–133.

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

Temesgen, H., and Weiskittel, A.R. 2006. Leaf mass per area relationships across light gradients in hybrid spruce crowns. Trees, **20**: 522–530. doi:10.1007/s00468-006-0068-0.

Temesgen, H., Hann, D.W., and Monleon, V.J. 2007. Regional height–diameter equations for major tree species of southwest Oregon. West. J. Appl. For. **22**: 213–219.

Vonesh, E.F., and Chinchilli, V.M. 1997. Linear and nonlinear models for the analysis of repeated measurements. Marcel Dekker Inc., New York.

Wang, C.H., and D.W. Hann. 1988. Height–diameter equations for sixteen tree species in the central Willamette Valley of Oregon. Oregon State Univ. For. Res. Lab. Res. Pap. 51.

Wykoff, W.R., Crookston, N.L., and Stage, A.R. 1982. User's guide to the Stand Prognosis Model. USDA For. Serv. Gen. Tech. Rep. INT-GTR-133.

Yang, R.C., Kozak, A., and Smith, J.H.G. 1978. The potential of Weibull-type functions as flexible growth curves. Can. J. For. Res. **8**: 424–431.

Zeide, B., and Vanderschaaf, C. 2001. The effect of density on the height–diameter relationship. *In*: Proceedings of the Eleventh Biennial Southern Silvicultural Research Conference, 20–22 March, Knoxville, Tenn. *Edited by* K.W. Outcalt. USDA For. Serv. Gen. Tech. Rep. SRS-GTR-48, pp. 453–456.

Zumrawi, A.A., and Hann, D.W. 1989. Equations for predicting the height to crown base of six tree species in the central western Willamette Valley of Oregon. Oregon State Univ. For. Res. Lab. Res. Pap. 52.