Forest Ecology and Management 258 (2009) 1235–1247

Contents lists available at ScienceDirect

Forest Ecology and Management

journal homepage: www.elsevier.com/locate/foreco

Assessing the impacts of species composition, top height and density on individual tree height prediction of quaking aspen in boreal mixedwoods

Shongming Huang [a,*], Douglas P. Wiens [b], Yuqing Yang [a], Shawn X. Meng [a], Curtis L. Vanderschaaf [c]

[a] *Biometrics Unit, Forest Management Branch, Alberta Department of Sustainable Resource Development, 8th Floor, 9920-108 Street, Edmonton, Alberta, Canada T5K 2M4*
[b] *Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, Alberta, Canada T6G 2G1*
[c] *Texas Forest Service, John B. Connally Building, 301 Tarrow, Suite 364, College Station, TX 77840-7896, USA*

ARTICLE INFO

Article history:
Received 11 March 2009
Received in revised form 8 June 2009
Accepted 9 June 2009

Keywords:
Boreal mixedwood
Individual tree height prediction
Nonlinear mixed model
Height–diameter function
Impact of stand level variables
Populus tremuloides

ABSTRACT

In this study we developed an individual tree height prediction model for quaking aspen (*Populus tremuloides* Michx.) grown in boreal mixedwood forests in Alberta using the nonlinear mixed model (NLMM) approach. We examined the impacts of density, species composition, and top height on aspen height predictions. Statistically significant stand level variables were incorporated into the base height–diameter model to increase the predictive ability and accuracy of the model at both the population and subject-specific levels. Our analyses showed that top height and density impacted height growth, but species composition did not. More importantly, we found that the inclusion of additional variables into the base model, despite improving model fitting statistics on the modelling data, did not improve the model's predictive ability and accuracy when cross-validated and when tested on an independent testing data set. Under the NLMM framework the base model performed as well as or better than the expanded models that contained other stand level variables. This has important theoretical and practical implications because, other than for biological reasons, more accurate local tree height predictions for aspen can be achieved simply by using the base height–diameter model fitted with the NLMM approach without the inclusion of other variables.

1. Introduction

Since total tree height (*H*) is more difficult and costly to measure than many other variables such as diameter, measurement of only a portion of sample trees for height is common when collecting forestry data. Estimating the "missing" tree heights is thus routinely required in practical management and silvicultural research (Curtis, 1967; Arabatzis and Burkhart, 1992; Huang et al., 1992; Peng et al., 2001; Castedo Dorado et al., 2006). The estimation of tree volume and biomass, the description of stand conditions and their changes over time, as well as the estimation of growth by stand projection methods all heavily depend on the accurate estimation of a complete set of tree heights. Many growth and yield projection systems (e.g., Wykoff et al., 1982; Arney, 1985) also require a complete set of tree heights as basic input.

The most commonly used method of estimating tree heights is to develop height–diameter models. Based on the trees that have both height and diameter measured, a height–diameter model can be developed to express tree height as a function of tree diameter at breast height (DBH). This model can then be used to predict the "missing" heights from measured DBHs.

It has been shown in numerous studies that the predictive ability and accuracy of a base H-DBH model can be improved if additional tree and stand variables (also called covariates) are added into the model, such as the inclusion of tree age, stand age, stand density and competition (e.g., basal area or stems per unit area, crown competition factor, crown ratio, competition index), site quality (e.g., site index, dominant/co-dominant height, top height), and species composition (e.g., tree number proportion, basal area proportion), among others (Curtis, 1967; Huang and Titus, 1994; Zeide and VanderSchaaf, 2002; Temesgen and Gadow, 2004; Adame et al., 2008; Meng et al., 2008; Saunders and Wagner, 2008).

The primary goal of this study was to develop an accurate individual tree height prediction model for aspen (*Populus tremuloides* Michx.) grown in boreal mixedwood forests in Alberta, Canada, using the nonlinear mixed model (NLMM) technique. Previous tree height prediction models available in Alberta were developed based on the nonlinear least squares methods (e.g., Huang, 1999; Huang et al., 2000). They were widely used for making population level predictions but their utility and accuracy at the plot (local) level were often limited. Our initial intent for this study was simple and straightforward: first, to select an

---

[*] Corresponding author. Tel.: +1 780 422 5281; fax: +1 780 427 0084.
E-mail address: shongming.huang@gov.ab.ca (S. Huang).

0378-1127/$ – see front matter. Crown Copyright © 2009 Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.foreco.2009.06.017

appropriate height–diameter function and use it as the base model; then, to evaluate and incorporate other tree and stand level variables into the base model within the NLMM framework to increase the predictive ability and accuracy of the model at both the population and local levels.

The main objectives of this study were threefold. First, to develop, assess and recommend the "best" aspen height prediction model for operational use based on the NLMM technique. The "best" model, besides being flexible and having reasonable biological interpretation of the parameters and the model as a whole, should also provide the most accurate predictions while being the simplest and least expensive in terms of data requirements. Second, to evaluate the effects of stand level attributes on aspen height predictions. We evaluated various attributes but focused on the three commonly used ones in this study: stand density (basal area per ha), top height (i.e., the average height of the 100 largest DBH trees per ha), and species composition (tree number proportion or basal area proportion).

The third objective of this study, arising as a result of the preliminary analyses, was to evaluate the impacts of including or not including statistically significant stand level variables into the base model within the NLMM framework. It is generally thought that the inclusion of additional variables will increase the accuracy of height predictions, where accuracy (δ) = bias^2 + variance of the errors (Cochran, 1977; Arabatzis and Burkhart, 1992). We were interested in confirming if this apparently simple and straightforward conclusion is true within the NLMM framework. If the inclusion of additional variables does not increase the accuracy of height predictions, the time and costs associated with collecting and analyzing the additional variables would be unnecessary for most applications.

2. Material and methods

2.1. Description of data

The data used in this study were collected by the Alberta Forest Service as a part of the provincial forest inventory database. A total of 466 plots were available for this analysis. These plots, which range in size from 200 m^2 to 2000 m^2 with the most common size being 1000 m^2, were randomly located throughout the boreal forests of the province to provide representative information for a variety of densities, heights, species compositions, stand structures, and site qualities. Within each plot, diameters at the breast height of 1.3 m above ground were measured for all trees taller than 1.3 m, and heights were measured for at least 10–20% of the trees in a systematic or random fashion (to cover the range of tree heights). A detailed description of the data collection procedures is provided in Alberta Forest Service (2000). The data were divided into two data sets: one for model fitting and the other for cross-validation. In total, 4371 aspen trees from the first 350 plots (three-quarters of the plots) were used for model fitting, and 2638 aspen trees from the 116 remaining plots (one-quarter of the plots) were used for cross-validation. In addition, an independent data set collected by the Alberta Forest Service, which consists of 2043 aspen trees from 80 plots, was used for model testing (Table 1).

The plot data were summarized to provide the information on stand density (trees per hectare, basal area per hectare) and top height, both for all species combined and by species. To describe the species composition (SC) of aspen grown in boreal mixedwood forests, two measures were used. One is based on tree number proportion (Kelty, 1989), and the other is based on basal area proportion (Huang and Titus, 1994):

$$\mathrm{SC_N} = \frac{N_{\mathrm{aspen}}}{N_{\mathrm{all}}} \tag{1}$$

$$\mathrm{SC_B} = \frac{\mathrm{BA_{aspen}}}{\mathrm{BA_{all}}} \tag{2}$$

where N_{aspen} is the stems/ha for aspen, N_{all} is the total number of stems/ha for all species combined, $\mathrm{BA_{aspen}}$ is the basal area (m^2/ha) for aspen, and $\mathrm{BA_{all}}$ is the total basal area (m^2/ha) for all species combined. Summary statistics of stand density, top height and species composition for aspen from different data sets are provided in Table 1. The data sets are also displayed in Figs. 1 and 2.

2.2. Base models

Based on a detailed examination of the height–diameter data (Fig. 1), a large number of base models were evaluated following well-established procedures (Arabatzis and Burkhart, 1992; Huang et al., 1992; Adame et al., 2008). The models include many presented in detail by Curtis (1967), Huang (1999), Huang et al.

Table 1
Tree level and plot level summary statistics for aspen from different data sets.

Data	Level	Variable	*N*	Mean	Min	Max	SD
Model fitting	Tree	DBH (cm)	4371	18.2121	1.10	61.50	9.03
		H (m)	4371	17.1348	2.00	32.00	5.08
	Plot	BA (m^2/ha)	350	18.2233	0.11	48.64	10.67
		$\mathrm{SC_N}$	350	0.5382	0.01	1.00	0.35
		$\mathrm{SC_B}$	350	0.5755	0.02	1.00	0.34
		Top*H* (m)	350	21.4132	6.33	31.20	4.40
Cross-validation	Tree	DBH	2638	17.9939	1.30	63.00	9.34
		H	2638	17.7981	2.40	30.60	5.40
	Plot	BA (m^2/ha)	116	22.2839	0.15	50.94	13.56
		$\mathrm{SC_N}$	116	0.6225	0.02	1.00	0.35
		$\mathrm{SC_B}$	116	0.6578	0.04	1.00	0.33
		Top*H* (m)	116	22.0479	6.45	29.10	4.22
Model testing	Tree	DBH	2043	30.0608	0.20	61.80	12.34
		H	2043	24.2080	1.40	37.20	7.43
	Plot	BA (m^2/ha)	80	17.2089	2.42	50.23	10.34
		$\mathrm{SC_N}$	80	0.4922	0.01	1.00	0.36
		$\mathrm{SC_B}$	80	0.5985	0.11	1.00	0.30
		Top*H* (m)	80	26.6773	18.95	34.12	3.15

Note: DBH is tree diameter at breast height, *H* is total tree height, BA is basal area per hectare, $\mathrm{SC_N}$ and $\mathrm{SC_B}$ are species compositions defined in Eqs. (1) and (2), Top*H* is top height, and *N* is sample size (i.e., number of trees at tree level, or number of plots at plot level). Min, max and SD are minimum, maximum and standard deviation, respectively.

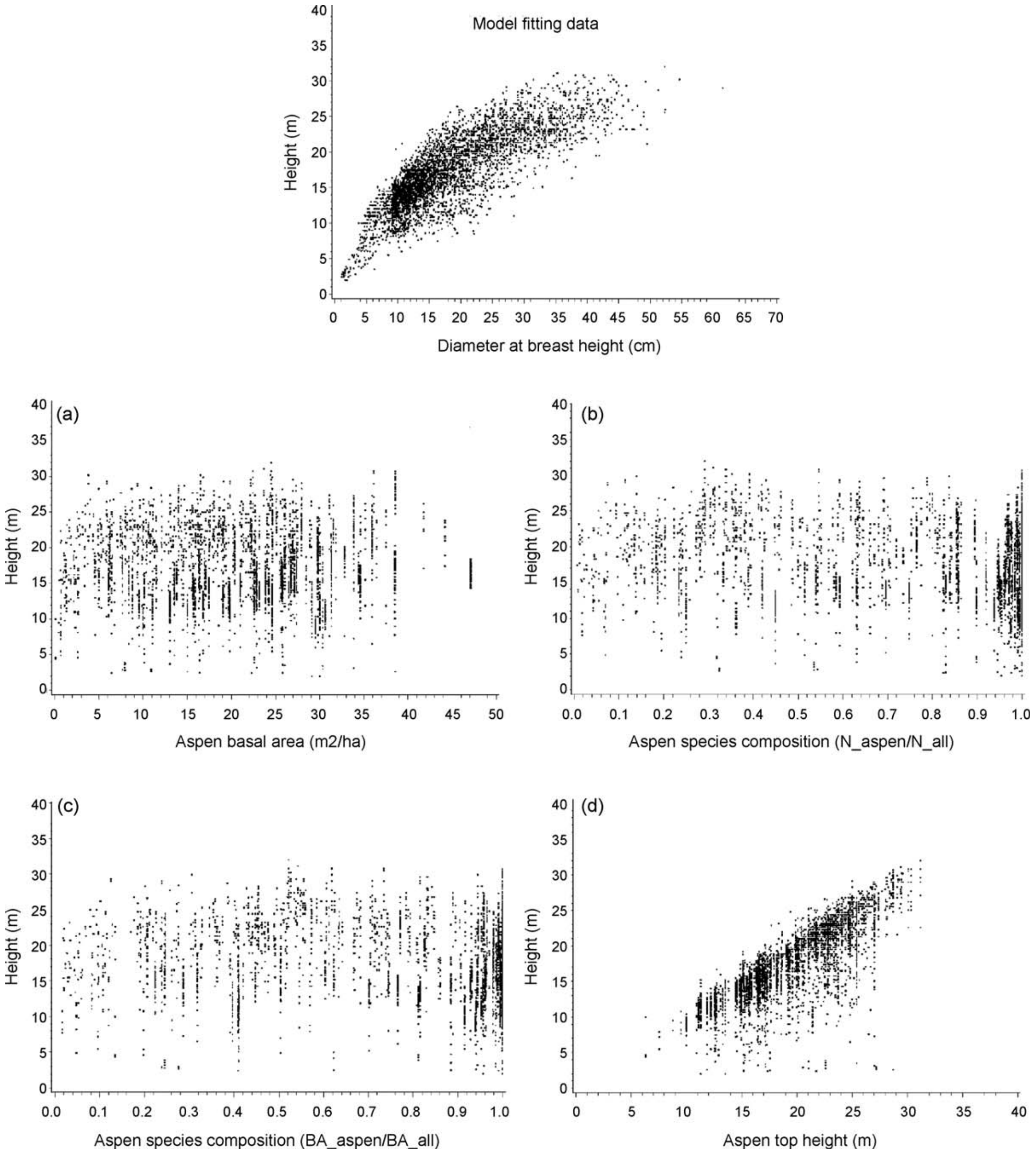

Fig. 1. Plots of model fitting data, where the species composition measures in (b) and (c) are defined in Eqs. (1) and (2), respectively. Summary statistics of the data are provided in Table 1.

(2000), Peng et al. (2001), and Lei et al. (2009). The two base models that showed the "best" performance on the aspen data took the following forms:

$$H = 1.3 + \beta_1[1 - \exp(-\beta_2 \mathrm{DBH})]^{\beta_3} \tag{3}$$

$$H = 1.3 + \frac{\beta_1}{1 + \exp[\beta_2 + \beta_3 \ln(\mathrm{DBH} + 1)]} \tag{4}$$

where H is total tree height (m), DBH is tree diameter (cm) at breast height, and β_1, β_2 and β_3 are model parameters. Model (3) is the Chapman–Richards function and model (4) is a logistic-type function (Ratkowsky, 1990). Both models have been commonly used in various applications and have been ranked favourably among alternative models in other studies (e.g., Huang et al., 1992; Huang, 1999; Peng et al., 2001; Temesgen et al., 2008).

The expressions given in Eqs. (3) and (4) assume that the parameters apply to every tree regardless of plot-specific conditions (i.e., they describe "population-averaged" trend). Since the parameter estimates from individual plots will likely be different due to site-specific differences, following the standard NLMM terminology (Davidian and Giltinan, 1995; Vonesh and Chinchilli, 1997), each parameter in Eqs. (3) and (4) was split into a fixed component, common to all plots within the population, and a random component, specific to each individual plot:

$$H_{ij} = 1.3 + (\beta_1 + b_{1i})[1 - \exp(-(\beta_2 + b_{2i})\mathrm{DBH}_{ij})]^{(\beta_3 + b_{3i})} + \varepsilon_{ij} \tag{5}$$

$$H_{ij} = 1.3 + \frac{(\beta_1 + b_{1i})}{1 + \exp[(\beta_2 + b_{2i}) + (\beta_3 + b_{3i})\ln(\mathrm{DBH}_{ij} + 1)]} + \varepsilon_{ij} \tag{6}$$

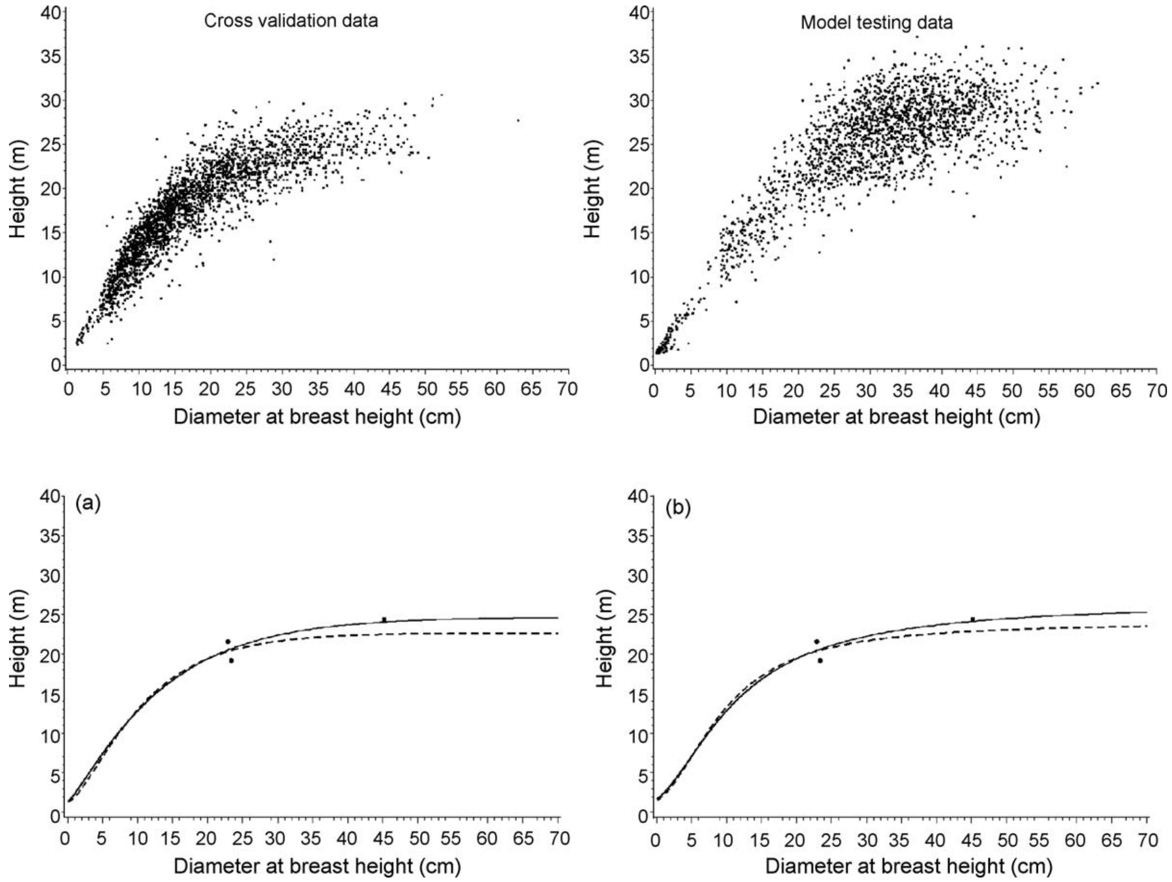

Fig. 2. Cross-validation data (top-left) and independent model testing data (top-right). Height predictions for an example plot of the cross-validation data with three observations are shown in (a) for the Chapman–Richards base model (7) (solid line) and expanded model (12) (dashed line), and in (b) for the logistic base model (8) (solid line) and expanded model (15) (dashed line). Actual data and calculations for the example plot are provided in Table 3.

where H_{ij} and DBH_{ij} are observed height and DBH for the jth tree in the ith plot (subject), $i = 1, 2, \ldots m$, $j = 1, 2, \ldots n_i$, m is the total number of plots in the population, n_i is the number of trees in the ith plot, β_1, β_2 and β_3 are fixed parameters common to all plots, b_{1i}, b_{2i} and b_{3i} are random parameters specific to the ith plot, and ε_{ij} is a normally distributed within-plot error term. In a more compact form, (5) and (6) can be written for the ith plot as

$$\mathbf{H}_i = 1.3 + (\beta_1 + b_{1i})[1 - \exp(-(\beta_2 + b_{2i})\mathbf{DBH}_i)]^{(\beta_3 + b_{3i})} + \varepsilon_i \tag{7}$$

$$\mathbf{H}_i = 1.3 + \frac{(\beta_1 + b_{1i})}{1 + \exp[(\beta_2 + b_{2i}) + (\beta_3 + b_{3i})\ln(\mathbf{DBH}_i + 1)]} + \varepsilon_i \tag{8}$$

where $\mathbf{H}_i$ and $\mathbf{DBH}_i$ are vectors of observed heights and DBHs for the ith plot, and $\varepsilon_i = [\varepsilon_{i_1}, \varepsilon_{i_2}, \ldots, \varepsilon_{i_{n_i}}]'$ is a vector of the within-plot errors.

2.3. Expanded models

To evaluate the potential effects of stand density, top height and species composition on aspen height predictions, basal area per ha, top height, $\mathrm{SC_N}$ and $\mathrm{SC_B}$ as defined in Eqs. (1) and (2) were incorporated into the base models (7) and (8). Different linear and nonlinear combinations of the variables and their transformations were tried and evaluated based on standard NLMM goodness-of-fit measures such as -2 times log-likelihood (-2LL), Akaike information criterion (AIC), Schwarz's Bayesian information criterion (BIC), significance of the parameters, and studentized residual plots (Vonesh and Chinchilli, 1997; Littell et al., 2006). Different specifications of the nature of the parameters as fixed and random, or purely fixed were examined as described in Fang and Bailey (2001), and Calama and Montero (2004). The selected expanded models formulated based on the Chapman–Richards base model (7) were:

$$\mathbf{H}_i = 1.3 + (\beta_1 + \beta_4 \mathrm{BA}_i + b_{1i})[1 - \exp(-(\beta_2 + b_{2i})\mathbf{DBH}_i)]^{(\beta_3 + b_{3i})} + \varepsilon_i \tag{9}$$

$$\mathbf{H}_i = 1.3 + (\beta_1 + \beta_4 \mathrm{SC}_{\mathrm{N}i} + b_{1i})[1 - \exp(-(\beta_2 + b_{2i})\mathbf{DBH}_i)]^{(\beta_3 + b_{3i})} + \varepsilon_i \tag{10}$$

$$\mathbf{H}_i = 1.3 + (\beta_1 + \beta_4 \mathrm{SC_B} + b_{1i})[1 - \exp(-(\beta_2 + b_{2i})\mathbf{DBH}_i)]^{(\beta_3 + b_{3i})} + \varepsilon_i \tag{11}$$

$$\mathbf{H}_i = 1.3 + (\beta_1 + \beta_4 \mathrm{Top}H_i + b_{1i})[1 - \exp(-(\beta_2 + b_{2i})\mathbf{DBH}_i)]^{(\beta_3 + b_{3i})} + \varepsilon_i \tag{12}$$

$$\mathbf{H}_i = 1.3 + (\beta_1 + \beta_4 \mathrm{Top}H_i + \beta_5 \mathrm{BA}_i + b_{1i})[1 - \exp(-(\beta_2 + b_{2i})\mathbf{DBH}_i)]^{(\beta_3 + b_{3i})} + \varepsilon_i \tag{13}$$

$$\mathbf{H}_i = 1.3 + (\beta_1 + \beta_4 \mathrm{Top}H_i + \beta_5 \mathrm{SC}_{\mathrm{N}i} + b_{1i})[1 - \exp(-(\beta_2 + b_{2i})\mathbf{DBH}_i)]^{(\beta_3 + b_{3i})} + \varepsilon_i \tag{14}$$

where BA_i and $\mathrm{Top}H_i$ are aspen density ($\mathrm{m^2/ha}$) and top height (m) for plot i, $\mathrm{SC}_{\mathrm{N}i}$ is the tree number based SC, $\mathrm{SC}_{\mathrm{B}i}$ is the basal area based SC, and β_1–β_5 are fixed parameters common to all plots. Attempts to incorporate the stand level variables in nonlinear forms, such as replacing $(\beta_1 + \beta_4\mathrm{Top}H_i + \beta_5\mathrm{SC}_{\mathrm{N}i})$ in (14) by $(\beta_1\mathrm{Top}H_i^{\beta_4}\mathrm{SC}_{\mathrm{N}i}^{\beta_5})$, resulted in no improvement. Therefore, they were incorporated into the base model in linear forms.

Similar expanded models were also evaluated and formulated for the logistic base model (8). For brevity, only one model which produced the most favourable goodness-of-fit measures is provided:

$$\mathbf{H}_i = 1.3 + \frac{(\beta_1 + \beta_4\mathrm{Top}H_i + b_{1i})}{1 + \exp[(\beta_2 + b_{2i}) + (\beta_3 + b_{3i})\ln(\mathbf{DBH}_i + 1)]} + \varepsilon_i \tag{15}$$

where all variables are as defined above.

We did not consider including tree age (e.g., Curtis, 1967) or stand age (e.g., Eerikäinen, 2003) as a covariate in expanded aspen tree height prediction models for the reasons discussed by Huang and Titus (1994), namely: (1) age is one of the most expensive, time consuming, and difficult variables to measure, particularly for aspen whose rings are very hard to count and more prone to missing or being false than most other species due to decay and fading; (2) for uneven-aged mixed-species stands, age generally accounts for the least amount of additional variation in height–diameter models; and (3) growth and yield models of mixed-species forests with irregular age structures rarely involve the explicit use of tree or stand age as a direct input variable.

2.4. Model estimation

The base and expanded H-DBH models (7)–(15) can all be written in a general form of

$$\mathbf{y}_i = f(\mathbf{x}_i, \boldsymbol{\beta}, \mathbf{b}_i) + \boldsymbol{\varepsilon}_i \tag{16}$$

where $\boldsymbol{y}_i = [y_{i_1}, y_{i_2}, \ldots, y_{in_i}]'$ is a vector of height measurements in plot i, $\mathbf{x}_i$ is a known design matrix of covariates (DBH, and where applicable also Top$H$, BA, $\mathrm{SC_N}$ or $\mathrm{SC_B}$), $\boldsymbol{\beta}$ is a vector of fixed parameters, $\mathbf{b}_i = [b_{1i}, b_{2i}, b_{3i}]'$ is a vector of random parameters, and ε_i is a vector of within-plot errors. Basic assumptions about the $\mathbf{b}_i$ and ε_i are that they are uncorrelated and normally distributed with mean zero and variance–covariance matrices $\mathbf{D}$ and $\mathbf{R}_i$, respectively, that is, $\mathbf{b}_i \sim N(\mathbf{0}, \mathbf{D})$ and $\varepsilon_i \sim N(\mathbf{0}, \mathbf{R}_i)$. Furthermore, since no obvious pattern of unequal error variance was detected from studentized residual plots, which is typical of such models after the inclusion of random parameters (e.g., Castedo Dorado et al., 2006; Adame et al., 2008; Meng et al., 2008), $\mathbf{R}_i$ is assumed to be $\sigma^2\mathbf{I}_{n_i}$, where σ^2 is the error variance and $\mathbf{I}_{n_i}$ is an $n_i \times n_i$ identity matrix, and $\mathbf{D}$ is assumed to be an unstructured covariance matrix constant for every plot within the population:

$$\mathbf{D} = \begin{bmatrix} \sigma_{b_1}^2 & \sigma_{b_1b_2} & \sigma_{b_1b_3} \\ \sigma_{b_1b_2} & \sigma_{b_2}^2 & \sigma_{b_2b_3} \\ \sigma_{b_1b_3} & \sigma_{b_2b_3} & \sigma_{b_3}^2 \end{bmatrix} \tag{17}$$

where $\sigma_{b_1}^2$, $\sigma_{b_2}^2$ and $\sigma_{b_3}^2$ are the variances for random parameters b_1, b_2 and b_3, respectively, and $\sigma_{b_1b_2}$, $\sigma_{b_1b_3}$ and $\sigma_{b_2b_3}$ are the covariances between pairs of random parameters.

Different NLMM methods have been developed to estimate the parameters of model (16). Davidian and Giltinan (1995), Vonesh and Chinchilli (1997), and Pinheiro and Bates (2004) provide comprehensive overviews as well as detailed theoretical developments about the methods and their applications in different fields. In this study we limited ourselves to the more commonly used first-order (FO) method of Beal and Sheiner (1982), which uses a first-order Taylor series expansion of (16) around a $\boldsymbol{\beta}^*$ close to $\boldsymbol{\beta}$ and a $\mathbf{b}_i^*$ set to zero, the expected value of the random parameters (i.e., $\mathbf{b}_i^* = E(\mathbf{b}_i) = \mathbf{0}$), to linearize the nonlinear function (16):

$$\mathbf{y}_i \approx f(\mathbf{x}_i, \boldsymbol{\beta}^*, \mathbf{b}_i^*) + \mathbf{X}_i(\boldsymbol{\beta} - \boldsymbol{\beta}^*) + \mathbf{Z}_i(\mathbf{b}_i - \mathbf{b}_i^*) + \boldsymbol{\varepsilon}_i \tag{18}$$

where the derivative matrices $\mathbf{X}_i$ and $\mathbf{Z}_i$ are defined by

$$\mathbf{X}_i = \left.\frac{\partial f(\mathbf{x}_i, \boldsymbol{\beta}, \mathbf{0})}{\partial \boldsymbol{\beta}'}\right|_{\boldsymbol{\beta}^*, 0} \qquad \mathbf{Z}_i = \left.\frac{\partial f(\mathbf{x}_i, \boldsymbol{\beta}, \mathbf{0})}{\partial \mathbf{b}_i'}\right|_{\boldsymbol{\beta}^*, 0} \tag{19}$$

Rearranging (18) yields (recognizing $\mathbf{b}_i^* = \mathbf{0}$):

$$\mathbf{y}_i - f(\mathbf{x}_i, \boldsymbol{\beta}^*, \mathbf{0}) + \mathbf{X}_i\boldsymbol{\beta}^* = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i + \boldsymbol{\varepsilon}_i \tag{20}$$

Define the left-hand side of (20) as the so-called pseudo-response function $\mathbf{y}_i^*$

$$\mathbf{y}_i^* = \mathbf{y}_i - f(\mathbf{x}_i, \boldsymbol{\beta}^*, \mathbf{0}) + \mathbf{X}_i\boldsymbol{\beta}^* \tag{21}$$

Therefore, (20) can be written as the standard linear mixed model:

$$\mathbf{y}_i^* = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i + \boldsymbol{\varepsilon}_i \tag{22}$$

Following linear mixed model theory, the generalized least squares estimator $\hat{\boldsymbol{\beta}}$ of the fixed parameters $\boldsymbol{\beta}$ in (22) can be obtained through an iterative procedure, as described in Fitzmaurice et al. (2004) and Vonesh and Chinchilli (1997). The random parameters predictor $\hat{\boldsymbol{b}}_i$ of the $\mathbf{b}_i$ is (Davidian and Giltinan, 1995):

$$\hat{\mathbf{b}}_i = \hat{\mathbf{D}}\mathbf{Z}_i'(\mathbf{Z}_i\hat{\mathbf{D}}\mathbf{Z}_i' + \hat{\mathbf{R}}_i)^{-1}[\mathbf{y}_i - f(\mathbf{x}_i, \hat{\boldsymbol{\beta}}, \mathbf{0})] \tag{23}$$

where $\hat{\mathbf{D}}$ is an estimate of $\mathbf{D}$ in (17), $\mathbf{Z}_i$ is defined in (19), and $\hat{\mathbf{R}}_i$ is an estimate of $\mathbf{R}_i$.

All model fittings were carried out using the SAS macro %NLINMIX, with the keyword expand = zero for the FO method (SAS Institute Inc, 2004; Littell et al., 2006).

2.5. Predictions using the cross-validation data and independent model testing data

The appropriateness of a fitted model needs to be tested through cross-validation or using independent model testing data (Picard and Cook, 1984; Yang et al., 2004). In order to evaluate the performance of the mixed H-DBH models fitted in this study, both the cross-validation data and the independent model testing data were used. At the plot level (or subject-specific level) where the main interest of NLMMs lies, the random parameters $\hat{\mathbf{b}}_i$ must first be calculated using Eq. (23). Once the $\hat{\mathbf{b}}_i$ are available, the pseudo-response function $\mathbf{y}_i^*$ in Eq. (22) can be predicted by replacing $\boldsymbol{\beta}$ and $\mathbf{b}_i$ with $\hat{\boldsymbol{\beta}}$ and $\hat{\mathbf{b}}_i$, respectively: $\hat{\mathbf{y}}_i^* = \mathbf{X}_i\hat{\boldsymbol{\beta}} + \mathbf{Z}_i\hat{\mathbf{b}}_i$. Substituting $\mathbf{y}_i^*$ in Eq. (21) by $\hat{\mathbf{y}}_i^*$ and rearranging (21) for $\mathbf{y}_i$ (and recognizing $\boldsymbol{\beta}^* = \hat{\boldsymbol{\beta}}$), we obtain the predicted $\mathbf{y}_i$ for plot i:

$$\hat{\mathbf{y}}_i = \mathbf{X}_i\hat{\boldsymbol{\beta}} + \mathbf{Z}_i\hat{\mathbf{b}}_i + f(\mathbf{x}_i, \hat{\boldsymbol{\beta}}, \mathbf{0}) - \mathbf{X}_i\hat{\boldsymbol{\beta}} = f(\mathbf{x}_i, \hat{\boldsymbol{\beta}}, \mathbf{0}) + \mathbf{Z}_i\hat{\mathbf{b}}_i \tag{24}$$

The errors (residuals) associated with the predictions are $e_{ij} = y_{ij} - \hat{y}_{ij}$, where y_{ij} and $\hat{y}_{ij}$ are the observed and predicted values for tree j in plot i. For a population of N observations from m plots (subjects), where $N = \sum_{i=1}^{m} \mathrm{n}_i$, the overall mean bias $(\bar{e})$ is the average of the residuals. An overall accuracy measure combining the mean bias and the standard deviation (SD) of the residuals, is calculated as follows (Cochran, 1977; Arabatzis and Burkhart, 1992):

$$\delta = \bar{e}^2 + \mathrm{SD}^2 \tag{25}$$

In addition to the errors calculated using both fixed and random parameters, we also calculated the errors using only the fixed parameters, for in practice, many researchers (e.g., Fang and Bailey, 2001; Calama and Montero, 2004) have used the fixed parameters estimated as a part of a NLMM to represent the "typical" population-averaged response curve via $\hat{\mathbf{y}}_{i\_\text{fix}} = f(\mathbf{x}_i, \hat{\boldsymbol{\beta}}, \mathbf{0})$ for the "typical" population-averaged predictions based on the fixed parameters $\hat{\boldsymbol{\beta}}$ only (with the random parameters $\hat{\mathbf{b}}_i$ set to zero), and $\mathbf{e}_{i\_\text{fix}} = \mathbf{y}_i - \hat{\mathbf{y}}_{i\_\text{fix}}$ for the associated residuals.

Since the predominant interest of a NLMM is to provide subject-specific predictions at a plot level using both fixed and random parameters, we also calculated the mean bias (or simply bias) for each individual subject. This is more important than the calculation of the overall bias averaged over all observations across all subjects, because the overall bias could sometimes be misleading due to the fact that the positive (under-prediction) and negative (over-prediction) errors from individual subjects could be cancelled, even though the under- or over-prediction for individual subjects could be large. In the worse case scenario, a NLMM could fit individual subjects poorly, yet the overall mean bias across all subjects could be zero or near zero. The mean bias $(\bar{e}_i)$ and the mean absolute bias $|\bar{e}_i|$ calculated for plot i with n_i observations are:

$$\bar{e}_i = \frac{\sum_{j=1}^{n_i} e_{ij}}{n_i} = \frac{\sum_{j=1}^{n_i} (y_{ij} - \hat{y}_{ij})}{n_i} \tag{26}$$

$$|\bar{e}_i| = \frac{\sum_{j=1}^{n_i} |e_{ij}|}{n_i} = \frac{\sum_{j=1}^{n_i} |y_{ij} - \hat{y}_{ij}|}{n_i} \tag{27}$$

Two other traditional goodness-of-fit statistics were also calculated by subject. One is the percent bias (Bias%), which puts the mean bias on a relative scale for easier interpretation, and the other is the root mean square error (RMSE):

$$\text{Bias\%}_i = \frac{\bar{y}_i - \bar{\hat{y}}_i}{\bar{y}_i} \times 100\% = \frac{\bar{e}_i}{\bar{y}_i} \times 100\% \tag{28}$$

$$\text{RMSE}_i = \sqrt{\frac{1}{n_i} \sum_{j=1}^{n_i} (y_{ij} - \hat{y}_{ij})^2} \tag{29}$$

where $\bar{y}_i$ and $\bar{\hat{y}}_i$ are subject-specific averages of the observed (y_{ij}) and predicted $(\hat{y}_{ij})$ values.

Distributions of the calculated $\bar{e}_i$, $|\bar{e}_i|$, Bias%$_i$ and RMSE$_i$ values from all plots in the cross-validation data and independent model testing data were examined and compared to determine the goodness-of-prediction of the base and expanded height–diameter models.

2.6. Impact of the number of prior observations per subject on predictions

At subject-specific level, the goodness-of-prediction of a NLMM is dependent on the availability and the number of prior observations per subject. There will be no subject-specific predictions if no previous H-DBH observation is available. Only the population-averaged marginal predictions can be made using the fixed parameters if no prior observation is available.

To assess the relationship between the goodness-of-prediction and the number of prior observations per subject, the cross-validation data and the independent model testing data were used. Within each plot, 1, 2, 3,. . ., 9, or all available prior observations (termed 9+) were randomly selected and used to predict the random parameters. The prediction errors were obtained in each case. This process was repeated 10 times and the averages of the prediction errors and relevant statistics were calculated. For a comparison, prediction errors from 0 prior observation (i.e., population-averaged marginal predictions based on the fixed parameters $\hat{\boldsymbol{\beta}}$ only) were also obtained in each case.

3. Results and discussion

3.1. Choice of the expanded models

Results of the estimated parameters and relevant fit statistics for the base models (7) and (8) and expanded models (12) and (15) are listed in Table 2.

For the expanded models formulated based on the Chapman–Richards function, when stand density (basal area per ha) was included in the H-DBH model (as in Eq. (9)), the estimated coefficient associated with BA is positive and significant ($p = 0.0021$). The positive estimate implies that increasing stand density has a positive effect on aspen height growth, and denser stands tend to result in taller aspen trees for the same DBH (provided other conditions are the same). It is generally accepted that for trees grown in even-aged pure species stands, the average height of the dominant/co-dominant trees or top height trees is relatively unaffected by a wide range of stand densities. This is the foundation of the site index concept (Clutter et al., 1983). However, the average height of all trees in the stand could be affected by stand density (Zeide and VanderSchaaf, 2002; Temesgen et al., 2008; Adame et al., 2008), and differences in stand structure (Saunders and Wagner, 2008), particularly for trees grown in

Table 2
Parameter estimates and fit statistics for different height–diameter models for aspen.

Parameter	Chapman–Richards		Logistic-type	
	Base model (7)	Expand model (12)	Base model (8)	Expand model (15)
β_1	22.1453	2.8392	24.0450	3.1154
β_2	0.09691	0.1231	4.3080	4.7083
β_3	1.3235	1.5302	−1.7723	−2.0692
β_4		0.8476		0.8827
$\sigma_{b_1}^2$	17.0377	0.2306	22.5203	1.6736
$\sigma_{b_1 b_2}$	−0.0939	−0.0192	0.1623	−0.9799
$\sigma_{b_2}^2$	0.00178	0.00216	0.8341	2.3565
$\sigma_{b_1 b_3}$	−0.1903	−0.0267	0.6901	0.6269
$\sigma_{b_2 b_3}$	0.01621	0.02134	−0.3427	−0.8877
$\sigma_{b_3}^2$	0.2488	0.5028	0.1826	0.3898
σ^2	1.6640	1.6096	1.6600	1.5946
P	10	11	10	11
−2LL	16171.1	15302.6	16670.9	15281.8
AIC	16191.1	15324.6	16690.9	15303.8
BIC	16229.7	15367.1	16731.1	15346.3

Note: $\sigma_{b_1}^2$, $\sigma_{b_2}^2$ and $\sigma_{b_3}^2$ are the variances for random parameters b_1, b_2 and b_3, respectively, $\sigma_{b_1 b_2}$, $\sigma_{b_1 b_3}$ and $\sigma_{b_2 b_3}$ are the covariances between pairs of random parameters, σ^2 is the error (residual) variance, P is the total number of parameters, −2LL is −2 times log-likelihood, AIC is Akaike information criterion, and BIC is the Schwarz's Bayesian information criterion.

mixedwood stands (Huang and Titus, 1994). Apparently, this was the case for aspen in boreal mixedwood stands in Alberta.

The estimated coefficients associated with species composition (SC_N or SC_B) were both insignificant ($p = 0.3450$ for SC_N and $p = 0.7856$ for SC_B), suggesting that the proportion of aspen in the stand has little to do with aspen height growth. This is likely a result of the competitive advantages of aspen over other species in boreal forests, as the shade-intolerant and fast-growing aspen typically establishes its dominance on the site quickly by occupying the upper layer of the canopy with little or no influence from the shade-tolerant and slow-growing species such as white spruce (*Picea glauca* (Moench) Voss), black spruce (*Picea mariana* (Mill.) B.S.P.) and balsam fir (*Abies balsamea* (L.) Mill) often found underneath aspen.

When top height was included in the Chapman–Richards H-DBH function (Eq. (12)), the estimated coefficient associated with top height is positive and highly significant ($p < 0.0001$). This is expected because, if top height is considered alone, it corresponds to the development of the stand, so the bigger the top height, the bigger the individual heights become because the stand is older. In addition, once the top height was included in the models, the addition of stand density (β_5 in Eq. (13)) or species composition (β_5 in Eq. (14)) becomes insignificant ($p > 0.05$). This suggests that top height is the most significant contributor among the three stand level variables, and it is statistically unnecessary to include stand density and/or species composition into the base model once top height is included. Model (12) with top height alone is therefore the "best" expanded Chapman–Richards model for aspen.

Similar analyses were conducted for models expanded based on the logistic base model (8). It was also found that when density was included in (8) without the inclusion of other stand level variables, it was significant. However, once top height was included in the models, density and/or species composition became insignificant. Therefore, the selected expanded model (15) for aspen based on the logistic-type function also included top height only.

Compared to their respective base models (7) and (8), the expanded models (12) and (15) have smaller AIC and BIC values (Table 2). All parameters in (12) and (15) are significant ($p < 0.05$), and the studentized residual plots (Fig. 3) are satisfactory. These are strong indications that the expanded models are better than the base models. Therefore, based exclusively on model fitting statistics, one of the two expanded models should be chosen over the base models to improve local tree height predictions for aspen grown in boreal mixedwood forests.

3.2. Predictions on cross-validation and independent model testing data sets

Based on the estimated parameters and the variance–covariance components (Table 2), height predictions from the base models (7) and (8) and the expanded models (12) and (15) were made using the cross-validation data and the independent model testing data. To illustrate the computations involved when using the Chapman–Richards base model (7), the first plot from the cross-validation data was chosen. This plot has three height–diameter observations, which are given in Table 3 and shown in Fig. 2.

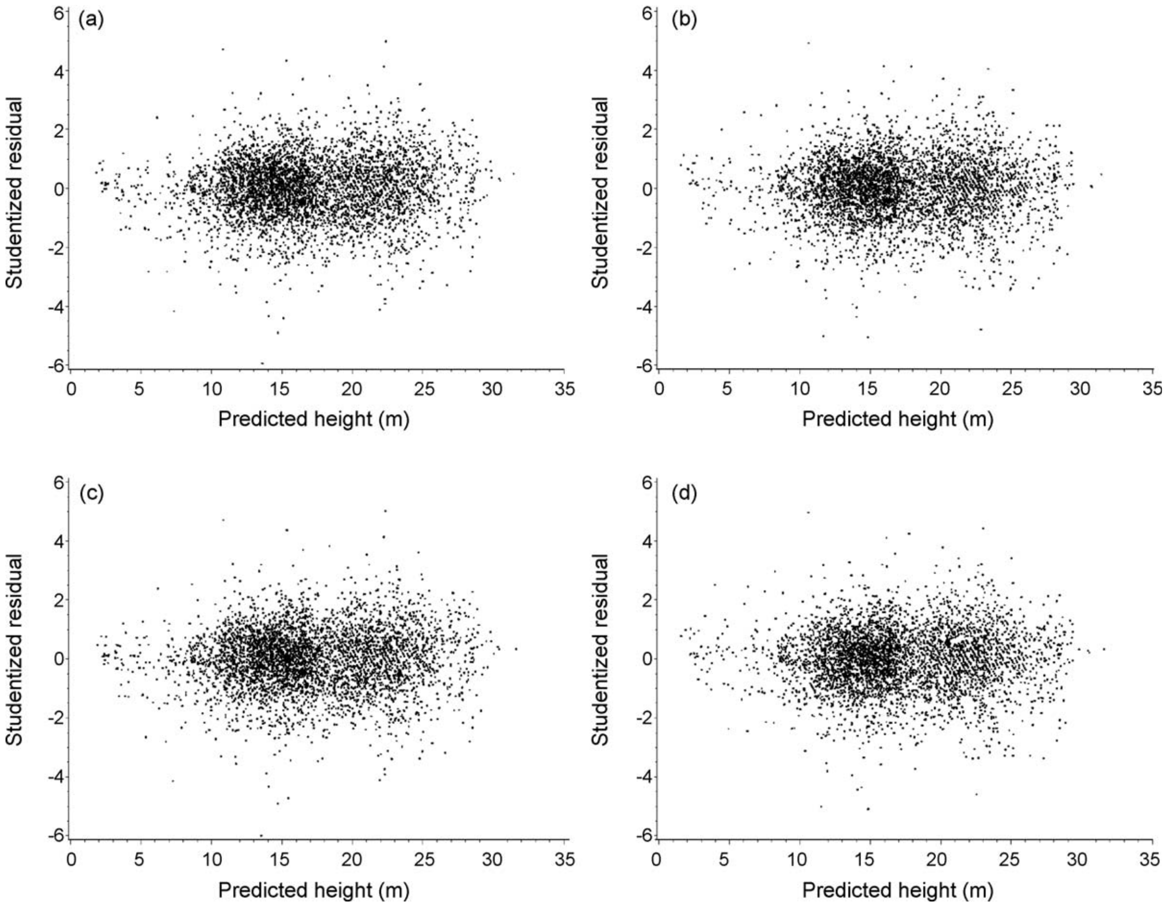

Fig. 3. Studentized residual plots from the nonlinear mixed Chapman–Richards base model (7) (a), Chapman–Richards expanded model (12) (b), logistic base model (8) (c), and logistic expanded model (15) (d).

Table 3
Height predictions for an example plot from the base and expanded models.

Model	Observed		Predicted							Bias%
	DBH	H	H_fix	e_fix	der_b_1	der_b_2	der_b_3	H_ss	e_ss	
(7)	22.9	21.6	20.32	1.2829	0.8587	70.2865	−2.1882	20.50	1.0981	−0.02
	23.4	19.2	20.46	−1.2624	0.8653	68.5514	−2.0947	20.67	−1.4724	
	45.2	24.4	23.08	1.3209	0.9835	16.5204	−0.2744	24.04	0.3626	
	Mean	21.73	21.29	0.4472				21.74	−0.0039	
(12)	22.9	21.6	20.65	0.9497	0.9102	43.0246	−1.1904	20.33	1.2706	2.93
	23.4	19.2	20.76	−1.5625	0.9154	41.4226	−1.1238	20.46	−1.2597	
	45.2	24.4	22.44	1.9642	0.9941	5.6248	−0.0812	22.5	1.8979	
	Mean	21.73	21.28	0.4505				21.1	0.6363	
(8)	22.9	21.6	20.26	1.3359	0.7887	−4.0073	−12.7185	20.45	1.1502	0.02
	23.4	19.2	20.41	−1.2096	0.7947	−3.9224	−12.5303	20.61	−1.4106	
	45.2	24.4	23.50	0.9041	0.9231	−1.7069	−6.5426	24.13	0.2703	
	Mean	21.73	21.39	0.3435				21.73	0.0033	
(15)	22.9	21.6	20.59	1.0068	0.8652	−2.6009	−8.2550	20.36	1.2435	2.26
	23.4	19.2	20.70	−1.5029	0.8701	−2.5203	−8.0511	20.48	−1.281	
	45.2	24.4	22.75	1.6549	0.9617	−0.8216	−3.1493	22.89	1.5089	
	Mean	21.73	21.35	0.3863				21.24	0.4904	

Note: (7) and (8) are the base Chapman–Richards and logistic models, respectively, (12) and (15) are their respective expanded models, DBH and H are tree diameter (cm) and height (m), H_fix and e_fix are height prediction and associated residual based on the fixed parameters only, der_b_1, der_b_2 and der_b_3 are derivatives with respect to the random parameters b_1, b_2 and b_3, respectively, H_ss and e_ss are subject-specific height prediction and associated residual based on the fixed and random parameters, and Bias% = 100 × mean of e_ss/mean of H (from Eq. (28)).

The population-averaged marginal height predictions based on the fixed parameters only are obtained directly from (7), with the random parameters set to zero. That is,

$$\hat{\mathbf{H}}_{i\_\text{fix}} = 1.3 + \hat{\beta}_1[1 - \exp(-\hat{\beta}_2 \mathbf{DBH}_i]^{\hat{\beta}_3}$$

where $\hat{\beta}_1 = 22.1453$, $\hat{\beta}_2 = 0.09691$ and $\hat{\beta}_3 = 1.3235$ (from Table 2). The predicted $\hat{\mathbf{H}}_{i\_\text{fix}}$ are listed in Table 3. The derivatives of (7) with respect to the three random parameters are:

$$\text{der\_}b_1 = \frac{\partial f(\mathbf{x}_i, \hat{\boldsymbol{\beta}}, \mathbf{0})}{\partial b_1} = [1 - \exp(-\hat{\beta}_2 \mathbf{DBH}_i)]^{\hat{\beta}_3}$$

$$\text{der\_}b_2 = \frac{\partial f(\mathbf{x}_i, \hat{\boldsymbol{\beta}}, \mathbf{0})}{\partial b_2} = \hat{\beta}_1 \mathbf{DBH}_i \hat{\beta}_3 \exp(-\hat{\beta}_2 \mathbf{DBH}_i)[1 - \exp(-\hat{\beta}_2 \mathbf{DBH}_i)]^{(\hat{\beta}_3 - 1)}$$

$$\text{der\_}b_3 = \frac{\partial f(\mathbf{x}_i, \hat{\boldsymbol{\beta}}, \mathbf{0})}{\partial b_3} = \hat{\beta}_1 \ln[1 - \exp(-\hat{\beta}_2 \mathbf{DBH}_i)][1 - \exp(-\hat{\beta}_2 \mathbf{DBH}_i)]^{\hat{\beta}_3}$$

The calculated derivatives at the given DBHs are listed in Table 3. They constitute the $\mathbf{Z}_i$ matrix defined in (19),

$$\mathbf{Z}_i = \begin{bmatrix} 0.8587 & 70.2865 & -2.1882 \\ 0.8653 & 68.5514 & -2.0947 \\ 0.9835 & 16.5204 & -0.2744 \end{bmatrix}.$$ From Table 2 and Eq. (17), we have

$$\hat{\mathbf{R}}_i = \begin{bmatrix} 1.6640 & 0 & 0 \\ 0 & 1.6640 & 0 \\ 0 & 0 & 1.6640 \end{bmatrix}$$

$$\hat{\mathbf{D}} = \begin{bmatrix} 17.0377 & -0.0939 & -0.1903 \\ -0.0939 & 0.00178 & 0.01621 \\ -0.1903 & 0.01621 & 0.2488 \end{bmatrix}$$

Hence, the random parameters predicted from Eq. (23) are:

$$\hat{\mathbf{b}}_i = \hat{\mathbf{D}}\mathbf{Z}_i'(\mathbf{Z}_i\hat{\mathbf{D}}\mathbf{Z}_i' + \hat{\mathbf{R}}_i)^{-1}(\mathbf{H}_i - \hat{\mathbf{H}}_{i\_\text{fix}})$$
$$= [1.1978, -0.01477, -0.08894]'$$

Once the $\hat{\mathbf{b}}_i$ are available, the predicted heights are obtained by $\hat{\mathbf{H}}_i = \hat{\mathbf{H}}_{i\_\text{fix}} + \mathbf{Z}_i\hat{\mathbf{b}}_i$ (from (24)). They are listed in Table 3, along with the associated residuals.

Height predictions from the other models were obtained in a similar manner. Results of the predictions and associated biases from all four models for the example plot are provided in Table 3. Note that the mean biases from the expanded models (12) and (15) are larger than their respective counterparts from the base models (7) and (8). This can also be seen in Fig. 2.

Following the above procedure, random parameters were calculated and height predictions were made by plot for all 116 plots of the cross-validation data and 80 plots of the model testing data. Predictions were also made based on the fixed parameters only. Summary statistics of the prediction biases for different models from all plots are listed in Table 4.

An examination of Table 4 reveals two distinct trends. First, the overall biases for the subject-specific predictions are much lower than those for the population-averaged predictions. This is expected because the population-averaged marginal predictions were obtained using the fixed parameters only (i.e., $\hat{\mathbf{H}}_{i\_\text{fix}} = 1.3 + \hat{\beta}_1[1 - \exp(-\hat{\beta}_2 \mathbf{DBH_i}]^{\hat{\beta}_3}$) ignoring any subject-specific trend in the data, whereas each subject (plot) in the subject-specific predictions has its own localized curve obtained using both fixed and random parameters.

As emphasized by Davidian and Giltinan (2003) and Fitzmaurice et al. (2004, pp. 360–364), and demonstrated by Hu et al. (1998), Young et al. (2007) and Huang (2008), the fixed parameters estimated as a part of a NLMM do not fully characterize the mean responses of the population, especially when the amount of the inter-subject variability is large. As a result, the "population-averaged" marginal estimates usually provide a biased partial representation of the true population mean responses, not an unbiased one as some forest modellers may have assumed (e.g., Fang and Bailey, 2001). Corrections, separate fits, or numerical integrations are often necessary to obtain unbiased population predictions.

The second trend observed from Table 4 is that, after calibrating the model for each plot using both fixed and random parameters, which is the main reason for developing a NLMM, the absolute sizes of the mean biases and the δ values from the base models are consistently smaller than those from their respective expanded models on both the cross-validation and independent model testing data sets, suggesting more accurate predictions from the base models on both data sets. As an example, the base Chapman–Richards model (7) produced a mean bias of −0.0052, with an SD of 1.60 (δ = 2.56). However, the expanded Chapman–Richards model (12) produced a mean bias of 0.0106 and a larger SD of 1.66 (δ = 2.76). The same phenomenon was also observed for the base and expanded logistic models (8) and (15) (Table 4). In both cases,

Table 4
Summary statistics for the overall height prediction errors from different data sets.

Data	Type	Model	N	Mean ($\bar{e}$)	Min	Max	SD	Bias%	δ
Cross-validation	SS	(7)	2638	−0.0052	−8.59	10.15	1.60	−0.03	2.56
		(12)	2638	0.0106	−8.68	13.74	1.66	0.06	2.76
		(8)	2638	−0.0036	−8.60	10.11	1.60	−0.02	2.56
		(15)	2638	0.0161	−8.69	12.26	1.64	0.09	2.69
	PA	(7)	2638	0.9275	−9.67	10.43	2.27	5.21	6.01
		(12)	2638	0.3860	−10.57	13.35	2.11	2.17	4.60
		(8)	2638	0.9780	−9.66	10.46	2.24	5.49	5.97
		(15)	2638	0.3902	−10.63	13.29	2.11	2.19	4.60
Model testing	SS	(7)	2043	0.0189	−13.81	7.82	1.91	0.08	3.65
		(12)	2043	0.2667	−13.99	8.01	1.99	1.10	4.03
		(8)	2043	0.0187	−13.89	7.72	1.92	0.08	3.69
		(15)	2043	0.1220	−14.07	7.74	1.95	0.50	3.82
	PA	(7)	2043	4.0205	−7.52	14.60	3.57	16.61	28.91
		(12)	2043	−0.3286	−17.41	7.79	2.51	−1.36	6.41
		(8)	2043	3.9061	−7.46	14.43	3.51	16.14	27.58
		(15)	2043	−0.3949	−17.25	7.69	2.46	−1.63	6.21

Note: SS refers to the subject-specific predictions based on the fixed and random parameters, PA refers to the population-averaged predictions based on the fixed parameters only, (7) and (12) are the base and expanded Chapman–Richards models, (8) and (15) are the base and expanded logistic models, N is the total number of observations (trees), min, max and SD are minimum, maximum and standard deviation, respectively, Bias% $= 100 \times \bar{e}/$mean height, and δ is the overall accuracy measure ($\delta = \bar{e} + \text{SD}^2$).

the base models are more accurate with lower standard deviations and smaller δ values than those from the expanded models. This is unexpected because the expanded models, which included an additional variable top height, should have provided more accurate predictions with lower standard deviations and smaller δ values.

3.3. Distributions of biases from base and expanded models

To help determine the possible reasons for larger biases from the expanded models, we examined the distributions of subject-specific biases from the base and expanded models. Since the overall mean biases averaged over all observations from all subjects (as in Table 4) could be misleading because the positive and negative errors from individual subjects could cancel one another out, the mean bias, mean absolute bias, Bias% and RMSE were calculated plot-by-plot using Eqs. (26)–(29), respectively. Summary statistics of the calculated results from all 116 plots of the cross-validation data are listed in Table 5. The results (not shown here) from the model testing data are similar. The plots where the Bias% exceeded ±2.5% of the observed mean were

Table 5
A summary of height prediction statistics from 116 plots of the cross-validation data.

Model	Data	Variable	N	Mean	Min	Max	SD	δ
(7)	All	$\bar{e}_i$	116	−0.0012	−0.390	0.296	0.098	0.010
		$\lvert\bar{e}_i\rvert$	116	0.9296	0.075	2.177	0.429	
		Bias$\%_i$	116	−0.1323	−6.187	1.330	0.936	
		RMSE$_i$	116	1.1372	0.075	2.839	0.552	
	$\lvert$Bias%$\rvert > 2.5$	$\bar{e}_i$	2	−0.3691	−0.390	−0.349	0.029	0.137
		$\lvert\bar{e}_i\rvert$	2	0.7464	0.635	0.858	0.157	
		Bias$\%_i$	2	−6.1859	−6.187	−6.185	0.001	
		RMSE$_i$	2	0.9634	0.850	1.077	0.161	
(12)	All	$\bar{e}_i$	116	0.0178	−1.022	1.436	0.403	0.163
		$\lvert\bar{e}_i\rvert$	116	0.9976	0.153	2.263	0.461	
		Bias$\%_i$	116	0.3702	−5.048	11.617	2.544	
		RMSE$_i$	116	1.2404	0.154	3.041	0.605	
	$\lvert$Bias%$\rvert > 2.5$	$\bar{e}_i$	25	−0.0299	−1.022	1.436	0.796	0.635
		$\lvert\bar{e}_i\rvert$	25	1.2459	0.443	2.182	0.447	
		Bias$\%_i$	25	1.0613	−5.048	11.617	5.212	
		RMSE$_i$	25	1.5580	0.478	2.922	0.633	
(8)	All	$\bar{e}_i$	116	0.0024	−0.338	0.290	0.091	0.008
		$\lvert\bar{e}_i\rvert$	116	0.9330	0.057	2.221	0.433	
		Bias$\%_i$	116	−0.0947	−5.371	1.305	0.806	
		RMSE$_i$	116	1.1410	0.057	2.862	0.554	
	$\lvert$Bias%$\rvert > 2.5$	$\bar{e}_i$	2	−0.3054	−0.338	−0.272	0.047	0.096
		$\lvert\bar{e}_i\rvert$	2	0.7066	0.568	0.845	0.196	
		Bias$\%_i$	2	−5.1036	−5.371	−4.837	0.378	
		RMSE$_i$	2	0.9110	0.798	1.024	0.160	
(15)	All	$\bar{e}_i$	116	0.0141	−0.919	1.071	0.355	0.126
		$\lvert\bar{e}_i\rvert$	116	0.9865	0.133	2.217	0.452	
		Bias$\%_i$	116	0.3255	−4.539	10.333	2.207	
		RMSE$_i$	116	1.2222	0.147	2.968	0.582	
	$\lvert$Bias%$\rvert > 2.5$	$\bar{e}_i$	21	0.0260	−0.919	1.071	0.716	0.513
		$\lvert\bar{e}_i\rvert$	21	1.1918	0.368	2.200	0.462	
		Bias$\%_i$	21	1.2816	−4.539	10.333	4.718	
		RMSE$_i$	21	1.5098	0.426	2.603	0.603	

Note: $\bar{e}_i$, $\lvert\bar{e}_i\rvert$, Bias$\%_i$ and RMSE$_i$ are defined in Eqs. (26)–(29), respectively, N is the number of subjects (plots), min, max and SD are minimum, maximum and standard deviation, respectively, and δ is the overall accuracy measure ($\delta = \bar{e}_i^2 + \text{SD}^2$).

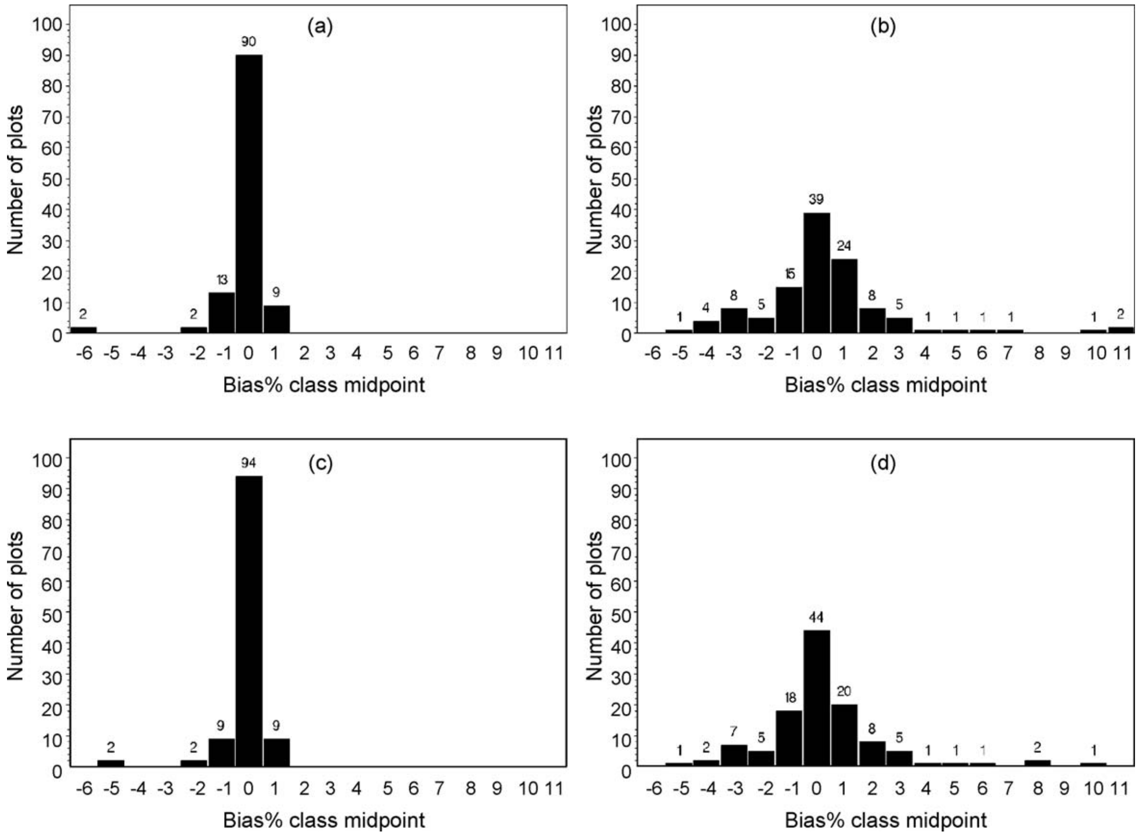

Fig. 4. Frequency distributions of Bias% from the base Chapman–Richards model (7) (a), expanded Chapman–Richards model (12) (b), base logistic model (8) (c), and expanded logistic model (15) (d) on the cross-validation data. The number on top of the bar represents the number of plots fall in that class.

identified and listed separately in Table 5 (the ±2.5% threshold was chosen by dividing the one-sided 5% significance level commonly used in statistical inferences into two-sides). The frequency distributions of the biases, absolute biases, Bias%s and RMSEs from all 116 plots were also obtained. For the sake of brevity, we only show the frequency distributions of the Bias%s from the base and expanded models (Fig. 4), because the frequency distributions of the other statistics showed similar patterns.

Judging from the results shown in Table 5 and Fig. 4, the contrasts between the base and expanded models are striking. The expanded models produced larger biases on more plots than the base models. For the Chapman–Richards models (7) and (12), among the 116 plots of the cross-validation data, only two plots produced biases exceeding ±2.5% of the observed mean when the base model (7) was used (Table 5). However, when the expanded model (12) was used, 25 plots produced biases exceeding ±2.5% of the observed mean.

The results for the logistic models are similar: only two plots had biases exceeding ±2.5% of the observed mean for the base model (8), whereas 21 plots had biases exceeding ±2.5% of the observed mean for the expanded model (15). The frequencies of poor predictions (i.e., with the $|\text{Bias\%}| > 2.5$) from the expanded models are much greater than those from the base models. This is clearly shown in Fig. 4, where the biases are centered around zero but there is more variability among the biases from the expanded models compared to the base models, suggesting larger prediction variations and lower accuracy for the expanded models. The standard deviations of the biases from the expanded models (0.403 and 0.355 for (12) and

Table 6
Prediction statistics on the model testing data obtained using different numbers of prior observations per plot.

No. of prior observations	Base model (7)				Expanded model (12)			
	$\bar{e}_i$	$\lvert\bar{e}_i\rvert$	Bias%	RMSE	$\bar{e}$	$\lvert\bar{e}\rvert$	Bias%	RMSE
0	4.0205	4.4290	16.6080	5.3790	−0.3286	1.8410	−1.3574	2.5291
1	0.9141	2.2108	3.7758	3.0339	−0.1437	1.7518	−0.5936	2.4126
2	0.4676	1.9058	1.9314	2.6424	0.0347	1.7270	0.1433	2.3634
3	0.2645	1.7351	1.0926	2.3717	−0.0092	1.6816	−0.0380	2.3074
4	0.2108	1.6628	0.8710	2.2689	0.0496	1.6407	0.2048	2.2433
5	0.1196	1.6121	0.4939	2.2099	0.0901	1.6124	0.3723	2.2104
6	0.1134	1.5717	0.4685	2.1469	0.1611	1.6108	0.6656	2.2024
7	0.0581	1.5412	0.2399	2.1024	0.1352	1.5812	0.5583	2.1477
8	0.0639	1.5212	0.2641	2.0748	0.1471	1.5744	0.6076	2.1447
9	0.0081	1.5014	0.0335	2.0498	0.2014	1.5586	0.8317	2.1147
9+	0.0189	1.4090	0.0781	1.9144	0.2667	1.4981	1.1015	2.0067

Note: $\bar{e}$, $|\bar{e}|$, Bias% and RMSE are mean bias, mean absolute bias, percent bias and root mean square error of predictions obtained on the independent model testing data. No. of prior observations refers to the number of observations used per plot to predict the random parameters, "0" indicates population-averaged marginal predictions based on the fixed parameters only, and "9+" indicates all available observations from each plot were used to predict the random parameters.

(15), respectively, see Table 5) were about four times larger than their respective counterparts from the base models (0.098 and 0.091 for (7) and (8), respectively). The δ values from the expanded models (0.163 and 0.126 for (12) and (15), respectively) were about 16 times larger than their respective counterparts from the base models (0.010 and 0.008 for (7) and (8), respectively).

3.4. Impact of the number of prior observations per subject on prediction accuracy

Table 6 lists the prediction statistics obtained on the independent model testing data for the base and expanded Chapman–Richards models. A graphic depiction of the statistics is shown in Fig. 5. Prediction statistics obtained on the cross-validation data, and for the base and expanded logistic models showed similar trends.

It is rather obvious from Table 6 and Fig. 5 that the prediction accuracy of the aspen H-DBH models is enhanced as the number of prior observations per plot increases. The most drastic change occurs from no. 0 to 1 prior observation. This is similar to what was observed by Temesgen et al. (2008) in the study of tree height prediction strategies. When no prior observation is available, the NLMMs reduce to marginal models with fixed parameters only, which performed poor on the testing data, particularly for the base model (7). When using at least one prior observation to predict the random parameters, the prediction statistics improved quite drastically. For instance, for the base model (7), the mean absolute bias $|\bar{e}|$ decreased from 4.4290 to 2.2108 and the RMSE decreased from 5.3790 to 3.0339. The reductions among other numbers of prior observations become smaller as the number of prior observations increases.

The results shown in Table 6 also suggest that, when 0–5 prior observations were used, the expanded model produced better prediction statistics suggesting its superiority over the base model. However, when there were 6 or more prior observations, which were typically available in forest inventories in Alberta, the base model produced smaller values of $\bar{e}$, $|\bar{e}|$, Bias% and RMSE. For instance, when all available prior observations were used, the $\bar{e}$, $|\bar{e}|$, Bias% and RMSE values for the base model are 0.0189, 1.4090, 0.0781 and 1.9144, respectively, whereas for the expanded model, they are all larger, at 0.2667, 1.4981, 1.1015 and 2.0067, respectively. This suggests that the base model provided more accurate predictions than the expanded model when 6 or more prior observations are available.

The poorer predictive performance of the expanded models is somewhat surprising. It was expected that the predictive performance of a base height–diameter model would be improved by including additional tree and stand level variables regardless of the number of prior observations. However, this was not observed for aspen grown in boreal mixedwood forests. We found that the addition of top height improved the fit of the base height–diameter models on the modelling data, and based on Table 4 alone, it also improved the accuracy of the base models on the cross-validation data and model testing data for the (biased) population-averaged

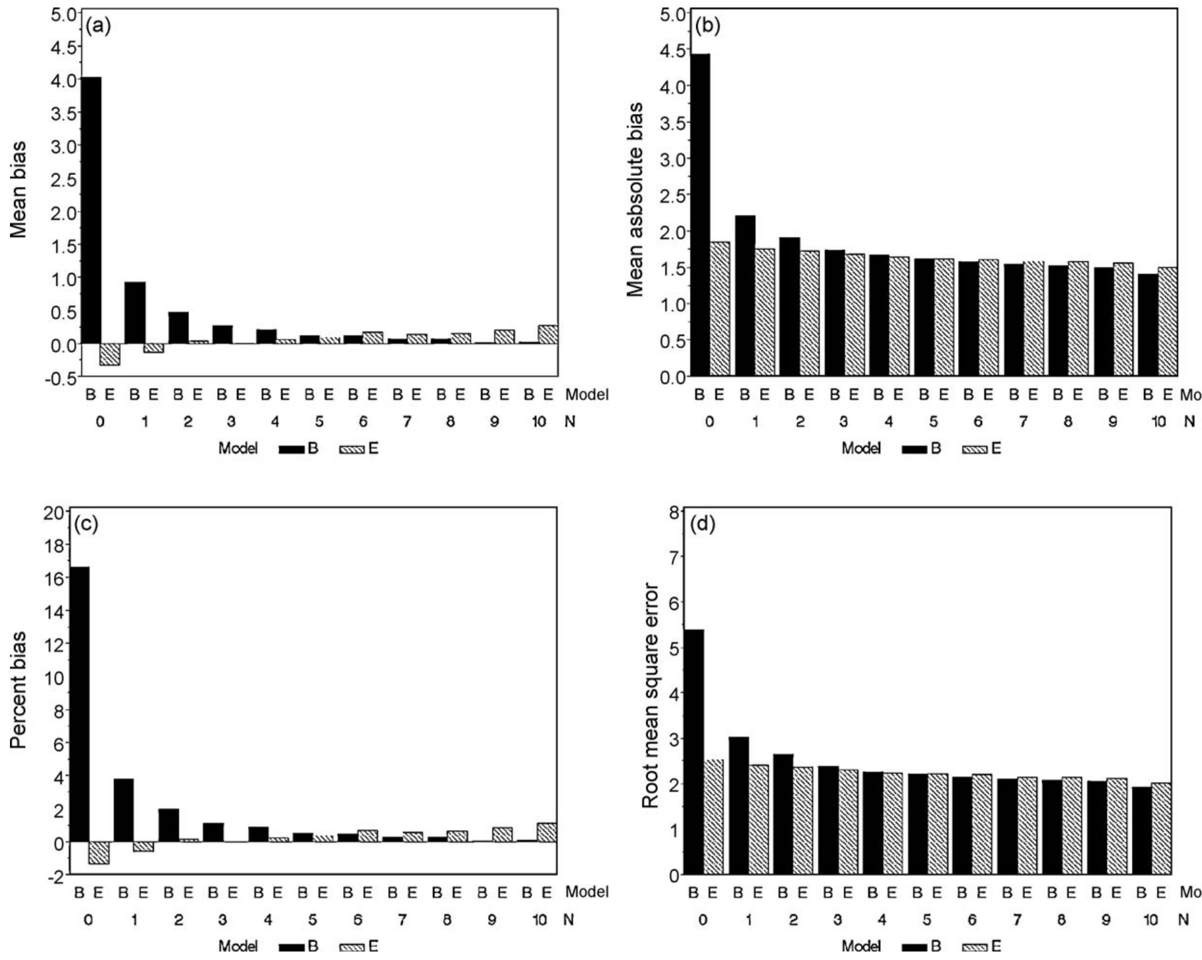

Fig. 5. Relationship between height prediction statistics and the number of prior observations per plot obtained on the independent model testing data, where "B" and "E" refer to the base and expanded Chapman–Richards models (7) and (12), respectively, *N* is the number of prior observations per plot. Actual values associated with the graphs are listed in Table 6, where "9+" corresponds to the "10" on the *x*-axis of the graphs.

marginal predictions using the fixed parameters only, as the biases, standard deviations and the overall accuracy measure δ from the expanded models were smaller than their counterparts from the base models. However, for subject-specific predictions, which are the main focus of a NLMM, no improvement from the expanded models was achieved over the base models. In fact, the expanded models performed consistently worse than the base models (Tables 4 and 5; Fig. 4). Only when there are fewer than 6 prior observations per subject, the expanded models outperformed the based models on the model testing data (Table 6; Fig. 5). Similar results (not shown here) were obtained when the expanded models were expanded further to include other variables such as quadratic mean diameter, mean height, and stand density in terms of stems/ha.

At this stage we are not exactly certain about what caused the base models to perform better in predictions than the expanded models. To avoid a random occurrence of the outcome, we split the data in various other forms instead of the reported 75–25% split for modelling and cross-validation. We also evaluated similar models on a different tree species (black spruce). The trends were identical to those we described here: under the NLMM framework, a similar or a better prediction ability of a model can be achieved by the base model alone rather than the expanded models with added covariates, even though the expanded models appear to have better fitting statistics on modelling data.

In previous NLMM analyses, the evaluation and the incorporation of additional variables into the base models played an important part in such analyses (e.g., Castedo Dorado et al., 2006; Trincado et al., 2007; Temesgen et al., 2008; Adame et al., 2008; Meng et al., 2008). However, to our knowledge, none of the analyses has compared the predictive results between the base and expanded models in a way as we did here (e.g., through the examination and comparison of the distributions of the biases from individual plots on cross-validation data and independent model testing data). It has generally been taken for granted that the inclusion of additional stand variables into the base models would automatically result in better predictions (helping to justify the increased costs associated with measuring the additional variables).

We believe that the inclusion of appropriate subject-specific random parameters in the base height–diameter models already allowed the plot level variations related to many known and unknown factors such as topography, soil type, nutrient status, genetics, climate, silvicultural regime, environment, intra- and inter-specific competitions, to be accounted for without actually requiring that they be identified or measured. This is a noted advantage of the NLMM technique. In many cases it eliminates the need to include additional plot level variables such as top height, stand density and species composition, which also implicitly represent the same biogeoclimatic, environmental and management effects that the random parameters already represent. In the context of over-fitting a model, the finding of greater bias in the models with more variables could also be attributed to the increased variability of estimated parameters, which would likely lead to greater bias in random parameter models.

For the purpose of evaluating the biological significance of other variables on aspen height growth within the NLMM framework, it is still useful to examine and incorporate, when warrant, these variables into the base H–D models to determine their importance (Temesgen et al., 2008; Meng et al., 2008; Saunders and Wagner, 2008). However, the results should be interpreted with great caution as the random parameters will likely convolute and mask their importance. For site-specific predictions at the plot level, we found that the addition of other variables beyond the base H-DBH models is not needed when 6 or more prior observations are available. Apparently the random parameters in the base models can take into account all or most of the other known and unknown factors for aspen grown in boreal mixedwood forests, enabling the plot level variations to be accounted for without the need to include other factors. This finding has important practical implications as more accurate tree height predictions for aspen can be achieved simply by using the mixed base H-DBH models without requiring measurements of other variables. Additional studies to further evaluate this finding on larger data sets, with added variables, and for other tree species and other types of models where a base model and expanded models are developed, will be very useful.

4. Conclusions

Based on the NLMM method, we developed individual tree height prediction models for aspen grown in boreal mixedwood forests in Alberta. We found that top height was the most significant predictor among the three stand level variables, and that species composition had no impact on aspen height growth. More importantly, we found that accurate local tree height predictions for aspen can be achieved using either the Chapman–Richards base model (7) or the logistic base model (8) without the need to include any other additional stand level variable. Contrary to intuition, we observed that the inclusion of other variables into the base height–diameter models actually produced consistently less accurate height predictions at the subject-specific level where the main interest of NLMMs lies.

Although for various other reasons it may still be useful to assess the significance of other tree and stand level variables on aspen height growth within the NLMM framework, for the purpose of prediction, there is no gain by including other variables into the base height–diameter models if a set of prior observations of 6 or more is available. The results of this study can have important practical implications because they demonstrate that the efforts, time and costs associated with collecting additional variables may not be necessary. Accurate subject-specific individual tree height predictions for aspen grown in boreal mixedwood forests can be better achieved using simple height–diameter models fitted by the NLMM technique.

Acknowledgements

This research was supported by the Alberta Government and Forest Resource Improvement Association of Alberta (the FRIAA-GYPSY project). Our special thanks to the 11 project partners, and to Dr. F. He, Mr. D. Morgan and Dr. K. McClain for reviewing an earlier draft. The constructive suggestions from the reviewers are greatly appreciated.

References

Adame, P., del Rio, M., Canellas, I., 2008. A mixed nonlinear height–diameter model for Pyrenean oak (*Ouercus pyrenaica* Willd.). For. Ecol. Manage. 256, 88–98.

Alberta Forest Service, 2000. Permanent Sample Plots Field Procedures Manual. Forest Manage. Branch, Alberta Forest Service, Edmonton, Alberta, Pub. FMOPC 83-03, 106 pp.

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

Arney, J.D., 1985. A modelling strategy for the growth projection of managed stands. Can. J. For. Res. 15, 511–518.

Beal, S.L., Sheiner, L.B., 1982. Estimating population kinetics. Crit. Rev. Biomed. Eng. 8, 195–222.

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

Castedo Dorado, F., Diéguez-Aranda, U., Barrio Anta, M., Rodríguez, M.S., Gadow, K.V., 2006. A generalized height–diameter model including random components for radiata pine plantations in northwestern Spain. For. Ecol. Manage. 229, 202–213.

Clutter, J.L., Fortson, J.C., Pienaar, L.V., Brister, G.H., Bailey, R.L., 1983. Timber Management—A Quantitative Approach. Wiley, New York, p. 333.

Cochran, W.G., 1977. Sampling Techniques, 3rd ed. Wiley, New York, p. 428.

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

Davidian, M., Giltinan, D.M., 1995. Nonlinear Models for Repeated Measurement Data. Chapman & Hall, New York, p. 360.

Davidian, M., Giltinan, D.M., 2003. Nonlinear models for repeated measurement data: an overview and update. J. Agric. Biol. Environ. Stat. 8, 387–419.

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

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

Fitzmaurice, G.M., Laird, N.M., Ware, J.H., 2004. Applied Longitudinal Analysis. Wiley, New York, p. 506.

Hu, F.B., Goldberg, J., Hedeker, D., Flay, B.R., Pentz, M.A., 1998. Comparison of population-averaged and subject-specific approaches for analyzing repeated binary outcomes. Am. J. Epidemiol. 147, 694–703.

Huang, S., 1999. Ecoregion-based individual tree height–diameter models for lodgepole pine in Alberta. West. J. Appl. For. 14, 186–193.

Huang, S., 2008. A generalized procedure for bias-free predictions at population and local levels. A Discussion Paper. Forest Management Branch, Alberta Ministry of Sustainable Resource Development, Edmonton, Alberta. Pub-Feb-22-2008, 61 pp.

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

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

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

Kelty, M.J., 1989. Productivity of New England hemlock/hardwood stands as affected by species composition and canopy structure. For. Ecol. Manage. 28, 237–257.

Lei, X., Peng, C., Wang, H., Zhou, X., 2009. Individual height–diameter models for young black spruce (*Picea mariana*) and jack pine (*Pinus banksiana*) plantations in New Brunswick, Canada. Forest. Chron. 85, 43–56.

Littell, R.C., Milliken, G.A., Stroup, W.W., Wolfinger, R.D., Schabenberger, O., 2006. SAS System for Mixed Models, 2nd ed. SAS Institute Inc., Cary, NC, p. 814.

Meng, X.S., Huang, S., Lieffers, V.J., Nunifu, T., Yang, Y., 2008. Wind speed and crown class influence the height–diameter relationship of lodgepole pine: nonlinear mixed effects modeling. For. Ecol. Manage. 256, 570–577.

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

Picard, R.R., Cook, R.D., 1984. Cross-validation of regression models. J. Am. Stat. Assoc. 79, 575–583.

Pinheiro, J.C., Bates, D.M., 2004. Mixed Effects Models in S and S-PLUS. Springer, New York, p. 528.

Ratkowsky, D.A., 1990. Handbook of Nonlinear Regression Models. Dekker, New York, p. 241.

SAS Institute Inc., 2004. SAS/STAT 9.1 User's Guide. SAS Institute Inc., Cary, NC, 5121 pp.

Saunders, M.R., Wagner, R.G., 2008. Height–diameter models with random coefficients and site variables for tree species of central Maine. Ann. For. Sci. 65, 203.

Temesgen, H., Gadow, K.V., 2004. Generalized height–diameter models—an application for major tree species in complex stands of interior British Colombia. Eur. J. For. Res. 123, 45–51.

Temesgen, H., Monleon, V.J., Hann, D.W., 2008. Analysis and comparison of nonlinear tree height prediction strategies for Douglas-fir forests. Can. J. For. Res. 38, 553–565.

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

Vonesh, E.F., Chinchilli, V.M., 1997. Linear and Nonlinear Models for the Analysis of Repeated Measurements. Marcel Dekker, New York, p. 560.

Wykoff, W.R., Crookston, N.L., Stage, A.R., 1982. User's guide to the stand prognosis model. In: USDA For. Serv. Gen. Tech. Rep. INT-133. p. 122.

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

Young, M.L., Preisser, J.S., Qaqish, B.F., Wolfson, M., 2007. Comparison of subject-specific and population averaged models for count data from cluster-unit intervention trials. Stat. Methods Med. Res. 16, 167–184.

Zeide, B., VanderSchaaf, C.L., 2002. The effect of density on the height–diameter relationship. In: Proceedings of the 11th Biennial Southern Silvicultural Research Conference, USDA For. Serv. Gen. Tech. Rep. SRS-48, Asheville, NC, pp. 463–466.