Population and plot-specific individual tree height-diameter models for major Alberta tree species

Prepared by:
Shongming Huang, Yuqing Yang, Darren Aitkin

August 1, 2013

Technical Report Pub. No.: T/600(2013)

ISBN No. 978-1-4601-0753-9 (Printed Edition)
ISBN No. 978-1-4601-0754-6 (On-line Edition)

For additional copies of this publication, please contact:

Forest Management Branch
Alberta Ministry of Environment and Sustainable Resource Development
8th Floor, Great West Life Building
9920–108 Street
Edmonton, Alberta, Canada T5K 2M4

Main reception: Tel: (780) 427-8474 Fax: (780) 427-0085

This publication will be made available on the Alberta Government website. Please check the latest information on the Alberta Government website at:
http://esrd.alberta.ca/lands-forests/forest-management/default.aspx

Edmonton, Alberta, Canada
August 2013

Executive Summary and Acknowledgement

Based on the nonlinear least squares method and nonlinear mixed-effects modeling technique, provincial and subregion-specific height-diameter models were developed for major Alberta tree species. These models can be used to predict the missing heights of individual trees at both population and plot-specific levels with varying accuracies. Different procedures for obtaining the most reasonable predictions under different circumstances are presented and evaluated. Examples of model application from real data are provided.

We appreciate the help and reviews of an early draft by Karl Peck, Daryl Price and Dave Morgan, who provided many constructive comments and suggestions.

Table of Contents

1. Introduction

Previous tree height-diameter models available in Alberta were developed based on the nonlinear least squares (NLS) method and stem analysis data (e.g., Huang et al. 1994, Huang et al. 2000). They were widely used for making population level predictions, where "population" may refer to the entire province or a specific natural region, subregion or forest management unit. However, the utility and accuracy of these population-based models at a plot-specific level were often limited.

This study uses both the NLS method and the nonlinear mixed-effects modeling (NMM) technique to develop new height-diameter models based on the expanded data (i.e., new measurements on existing plots and/or new plots) from stem analysis and Permanent Sample Plots (PSPs). The main objective of the study is to increase the predictive accuracy of the height-diameter models at a plot-specific level, while maintaining or improving their predictive accuracy at the population level as well.

To facilitate the understanding and use of the fitted models, detailed application examples were provided based on real data. These examples demonstrate model application procedures that readers can use in practice to make the "best" prediction at population and plot-specific levels.

1.1 Species, Species Code and Species Grouping

Whenever possible and reasonable, models developed in this study were species-specific. Models were also developed for deciduous species combined and coniferous species combined. Due to their limited sample sizes, whitebark pine, limber pine, Engelmann spruce, alpine fir, western larch and alpine larch were grouped into relevant species as defined in Table 1. A total of 10 species are present after the grouping: three deciduous and seven coniferous. All models were developed for each of these 10 species (last column of Table 1). Lodgepole pine, white spruce, aspen and black spruce are the four leading tree species in Alberta.

Table 1. List of Alberta tree species, species code and species grouping.

Deciduous/ coniferous	Species	Scientific name	Species code	Grouped species code
Deciduous	Aspen	*Populus tremuloides* Michx.	AW	AW
	Balsam poplar	*Populus balsamifera* L.	PB	PB
	White birch	*Betula papyrifera* Marsh.	BW	BW
Coniferous	Lodgepole pine	*Pinus contorta* var. *latifolia* Engelm.	PL	PL
	Whitebark pine	*Pinus albicaulis* Engelm.	PW	PL
	Limber pine	*Pinus flexilis* E. James	PF	PL
	Jack pine	*Pinus banksiana* Lamb.	PJ	PJ
	White spruce	*Picea glauca* (Moench) Voss	SW	SW
	Engelmann spruce	*Picea engelmannii* Parry ex Engelm.	SE	SW
	Black spruce	*Picea mariana* (Mill.) B.S.P.	SB	SB
	Balsam fir	*Abies balsamea* (L.) Mill.	FB	FB
	Alpine fir	*Abies lasiocarpa* (Hook.) Nutt.	FA	FB
	Douglas-fir	*Pseudotsuga menziesii* (Mirb.) Franco	FD	FD
	Tamarack larch	*Larix laricina* (Du Roi) K. Koch	LT	LT
	Western larch	*Larix occidentalis* Nutt.	LW	LT
	Alpine larch	*Larix lyallii* Parlatore	LA	LT

1.2 Natural Regions and Subregions of Alberta

Alberta is divided into six natural regions and 21 natural subregions listed in Table 2 and shown in Figure 1. They are defined by the Natural Regions Committee (2006) based on the biogeoclimatic characteristics of the regions in Alberta. Natural regions and subregions have also been referred to as ecoregions.

Table2. Natural regions and subregions of Alberta.

Natural region	Natural subregion	Numeric code	Character code
Boreal Forest	Central Mixedwood	1	CM
	Dry Mixedwood	2	DM
	Northern Mixedwood (Wetland Mixedwood)	3	NM
	Boreal Subarctic	4	BSA
	Peace-Athabasca Delta (Peace River Lowlands)	5	PAD
	Lower Boreal Highlands (Boreal Highlands)	6	LBH
	Upper Boreal Highlands (Boreal Highlands)	21	UBH
	Athabasca Plain	12	AP
Rocky Mountain	Alpine	7	ALP
	Subalpine	8	SA
	Montane	9	MT
Foothills	Upper Foothills	10	UF
	Lower Foothills	11	LF
Canadian Shield	Kazan Upland	13	KU
Parkland	Foothills Parkland	14	FP
	Peace River Parkland	15	PRP
	Central Parkland	16	CP
Grassland	Dry Mixedgrass	17	DMG
	Foothills Fescue	18	FF
	Northern Fescue	19	NF
	Mixedgrass	20	MG

The vast majority of the data used in this study were collected in four main forest production areas: lower foothills subregion, upper foothills subregion, and central and dry mixedwood subregions.

For four species (one deciduous and three coniferous) with relatively small sample sizes or concentrated species distribution ranges, only provincial models were developed based on the data combined from all natural subregions:

–BW, FD, LT, PJ

For six other species (two deciduous and four coniferous) with relatively large sample sizes and wide species distribution ranges, the models were developed provincially and by natural subregion or group of natural subregions:

–AW, PB, PL, SW, SB, FB

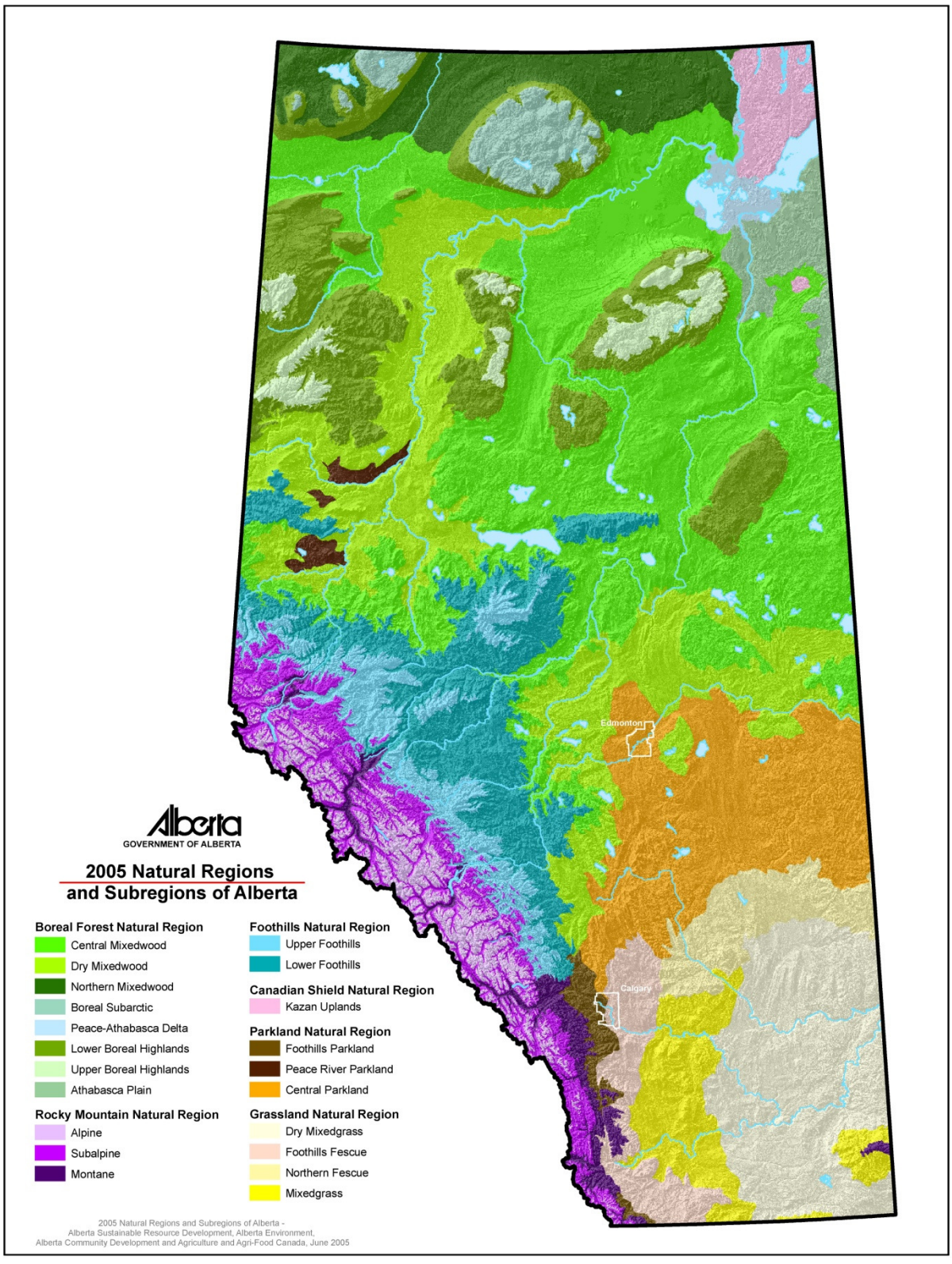

Figure 1. Natural regions and subregions of Alberta. Designated numeric and character codes for each natural subregion are listed in Table 2.

2. Data

Data from the Alberta government stem analysis database and PSP database were combined to fit the height-diameter models for major Alberta tree species. A summary of the data is presented in Table 3.

Table 3. A summary of the data used to fit height-diameter models.

Species	Subregion	*N*	*m*	DBH (cm)				H (m)			
				Mean	Min	Max	SD	Mean	Min	Max	SD
AW	7-10	490	192	24.07	1.60	57.40	8.75	18.94	2.00	30.70	5.33
	11	8323	1330	26.25	1.10	67.10	11.39	21.50	1.50	34.70	5.87
	1-6, 12-21	7802	1358	21.43	0.20	73.10	10.68	19.46	1.40	37.20	6.03
	Provincial	16615	2880	23.92	0.20	73.10	11.24	20.47	1.40	37.20	6.02
PB	7-11	2412	537	20.74	0.70	67.30	12.95	16.44	1.60	33.80	7.84
	1-6, 12-21	1501	415	21.71	1.10	80.80	12.61	17.98	1.70	36.60	6.90
	Provincial	3913	952	21.11	0.70	80.80	12.83	17.03	1.60	36.60	7.53
BW	Provincial	2088	608	6.95	0.10	36.60	6.55	7.69	1.40	27.20	5.25
Deciduous	Provincial	22616	3431	21.87	0.10	80.80	12.21	18.69	1.40	37.20	7.28
PL	7-9	3721	221	15.99	0.60	53.50	6.04	13.80	1.50	26.25	3.63
	10	19458	1411	20.11	0.80	59.20	6.92	18.21	1.50	33.80	4.37
	11	24814	1830	22.74	0.50	64.60	7.62	20.90	1.40	36.80	4.65
	1-6, 12-21	2964	401	25.08	5.10	55.40	7.85	21.02	6.10	34.10	4.50
	Provincial	50957	3863	21.38	0.50	64.60	7.56	19.36	1.40	36.80	4.90
SW	7-9	1766	146	17.43	0.70	63.40	10.19	12.77	1.40	30.50	5.53
	10	9063	864	21.50	0.50	78.50	11.67	16.52	1.40	36.00	6.96
	11	14150	1604	24.35	0.20	74.00	13.44	19.17	1.40	40.20	8.74
	1-6, 12-21	20314	1684	25.99	0.20	77.70	11.24	22.03	1.40	43.30	6.88
	Provincial	45293	4298	24.24	0.20	78.50	12.21	19.67	1.40	43.30	7.90
SB	7-10	5168	703	12.56	0.40	45.00	5.50	11.09	1.40	25.20	4.23
	11	8597	935	13.26	0.40	55.30	5.74	12.65	1.40	32.00	4.48
	1-6, 12-21	1321	245	15.39	0.80	41.10	7.19	14.04	1.60	30.70	5.41
	Provincial	15086	1883	13.21	0.40	55.30	5.85	12.24	1.40	32.00	4.58
FB	7-9	1369	93	11.30	1.10	53.00	5.49	9.73	1.50	24.50	4.29
	10	4337	440	17.47	0.30	57.60	10.12	13.64	1.40	31.40	7.16
	11	4748	507	10.00	0.20	51.30	8.29	9.13	1.40	31.50	6.84
	1-6, 12-21	2399	251	12.96	0.30	44.90	7.61	11.47	1.40	32.00	6.38
	Provincial	12853	1291	13.21	0.20	57.60	9.18	11.15	1.40	32.00	6.93
FD	Provincial	841	45	18.64	1.30	60.90	8.77	13.67	1.80	26.60	4.13
LT	Provincial	1378	150	9.00	1.10	37.70	5.67	8.51	1.60	27.90	5.33
PJ	Provincial	3681	217	15.55	1.00	45.00	6.89	14.49	1.60	28.80	4.58
Coniferous	Provincial	130089	6625	20.31	0.20	78.50	10.37	17.54	1.40	43.30	7.12

Note:

Species and subregions are defined in Tables 1 and 2, respectively;
Provincial = all subregions combined;
Deciduous = all deciduous species combined;
Coniferous = all coniferous species combined;
N = total number of observations (trees);
m = number of plots;
DBH = tree diameter (in centimeters, cm) at the breast height of 1.30 meters above ground;
H = total tree height (in meters, m);
Min, max and SD = minimum, maximum and standard deviation, respectively.

The stem analysis data (13,174 trees from 2,489 plots) were collected from felled trees in temporary sample plots. For PSPs, each measurement was considered a temporary sample plot. Due to several factors, including the changes in measurement protocol since 1981 (Alberta Sustainable Resource Development 2005), the large sample sizes, and the need to reserve some data for possible model validation, for the six species with large sample sizes and wide distribution ranges, only the PSPs established prior to 1981 plus the stem analysis data were used for model development in this study:

- AW, PB, PL, SW, SB, FB

For four other species with relatively small sample sizes or concentrated species distribution ranges, all data from PSPs and stem analysis were used:

- BW, FD, LT, PJ

In addition, as a standard data screening procedure in developing tree height-diameter models, trees identified with the following conditions were excluded from modeling:

- Dead top or dieback
- Multiple leaders
- Poor form/stem form defect
- Broken top
- Standing dead
- Dead top dieback with new leader
- Severe leaning
- Fork
- Pronounced crook
- Broken stem
- Dead or down

3. Height-Diameter Models and Goodness-of-Fit Statistics

3.1 Base Models

Comparison of alternative base model forms suggests that the following base models are appropriate for describing the population level tree height-diameter relationship for deciduous and coniferous species in Alberta:

[1] $$H = 1.30 + b_1[1 - \exp(-b_2 DBH)]^{b_3}$$ (Deciduous species)

[2] $$H = 1.30 + \frac{b_1}{1 + \exp[b_2 + b_3 \ln(DBH)]}$$ (Coniferous species)

where H is total tree height (m), DBH is tree diameter (cm) at the breast height of 1.30 m above ground, b_1, b_2 and b_3 are model parameters applicable to the population, exp denotes the exponential function, and ln denotes the natural logarithm to the base e (e≈2.718281828).

3.2 Mixed Models

The mixed models take the following forms:

[3] $$H = 1.30 + (b_1 + u_1)[1 - \exp(-(b_2 + u_2)DBH)]^{b_3}$$ (Deciduous species)

[4] $$H = 1.30 + \frac{(b_1 + u_1)}{1 + \exp[(b_2 + u_2) + b_3 \ln(DBH)]}$$ (Coniferous species)

where b_1, b_2 and b_3 are fixed parameters applicable to every plot in the population, and u_1 and u_2 are random parameters unique for each plot in the population. Attempts to incorporate a third random parameter u_3 to b_3 produced better fit statistics but failed to give stable predictions in some cases. This is illustrated in "Additional Notes" for interested readers (Section 8.1). For BW, only one random parameter u_1 was used in the mixed model [3] due to data and prediction issues. This is also illustrated in "Additional Notes".

3.3 Parameter Estimates and Goodness-of-Fit Statistics

Parameter estimates for base and mixed models are listed in Appendix 1 (Tables A1 to A7), along with relevant residual and spaghetti plots (Figures A1 to A16). Summary goodness-of-fit statistics associated with different types of predictions are listed in Appendix 2. The parameter estimates for base models were obtained from the ordinary NLS method. The parameter estimates for mixed models were obtained from the first-order method of the NMM technique.

The fitted models can be used in different ways. The most common examples are demonstrated in Sections 5 and 6, after some related background material is presented in Section 4.

The summary goodness-of-fit statistics listed in Appendix 2 pertain to different types of predictions discussed in Section 4. Readers who are familiar with how to use the fitted models to make predictions at population and plot-specific levels can skip the background material and application examples, and go directly to "Recommendations" (Section 7). For interested modelers, several additional notes about the model development are presented in Section 8.

4. Prediction Types and Goodness-of-Fit Measures

4.1 Population and Subject-Specific Models

In this study, the word "population" refers to the entire province, or a specific natural subregion or a group of natural subregions. Readers can use the word to denote the aggregated data from a company, a region, or a specific area consists of *m* experimental units. For height-diameter models, the experimental units, also referred to as "subjects" in mixed model parlance, are sample plots. The elements within each sample plot are trees. Height and diameter observations are made on the trees within each plot, where "diameter" in this study always refers to the diameter at the breast height of 1.30 m above ground (DBH).

Regression models can generally be classified as population-based models and subject-specific models. Traditional regression models estimated from least squares methods are typically population-based models. They are intended to describe the population averages, and as such, they are also referred to as "population-averaged" models. As an example, a population-based height-diameter model estimated from the NLS method for a population consisting of *m* plots can be written as:

[5] $$H = 1.30 + b_1[1 - \exp(-b_2 DBH)]^{b_3}$$

where H is total tree height (m), DBH is tree diameter (cm) at breast height, and b_1, b_2 and b_3 are estimated coefficients for the population.

The estimated coefficients in [5] (b_1, b_2 and b_3) obtained from the NLS method are derived as though all data in the population had come from one big plot. They apply to the entire population. There is no mechanism in [5] to differentiate the plots that the model applies to. Given the same DBH, the same H will be predicted regardless of which plot the DBH is observed in.

Population-based models developed from the least squares method predict the population averages from the observed *x*-variable(s). As the name implies, they are "population-based", not subject-specific. One common problem with population-based models is that, due to the intrinsic variation among the subjects within a population and the polymorphic nature of biological growth on different sites, the trends exhibited by the data from individual subjects within a population could be quite different from the trend exhibited by the population averages. This is illustrated in Figure 2, where the data from individual subjects display differing trends than that of the population averages. As a result, it is quite possible that a population-based model may fit or predict the data well on average for the entire population, but it could perform poorly for the individual subjects in the population. Sometimes population averages could be meaningless at a subject-specific level.

Subject-specific models, on the other hand, are primarily designed to provide better fits and predictions for individual subjects within a population. They can be developed using different procedures (e.g., repeated subject-specific NLS fits, indicator variable and varying coefficient approaches). They can also be developed based on the NMM technique, which is the focus of this study as all models developed in this study are nonlinear models. As an example, a subject-specific nonlinear mixed height-diameter model corresponding to [5] can be written as:

[6] $$H_{ij} = 1.30 + (b_1 + u_{1i})[1 - \exp(-(b_2 + u_{2i})DBH_{ij})]^{(b_3 + u_{3i})}$$

where H_{ij} and DBH_{ij} are the observed height and diameter for the *j*th tree in the *i*th plot, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n_i$, *m* is the number of plots in the population, n_i is the number of trees in the *i*th plot, b_1, b_2 and b_3 are fixed

parameters common to every plot in the population, and u_{1i}, u_{2i} and u_{3i} are random parameters unique for each plot in the population. Since the random parameters are unique for each plot in the population, it is necessary to use subscripts to differentiate the plots within the population.

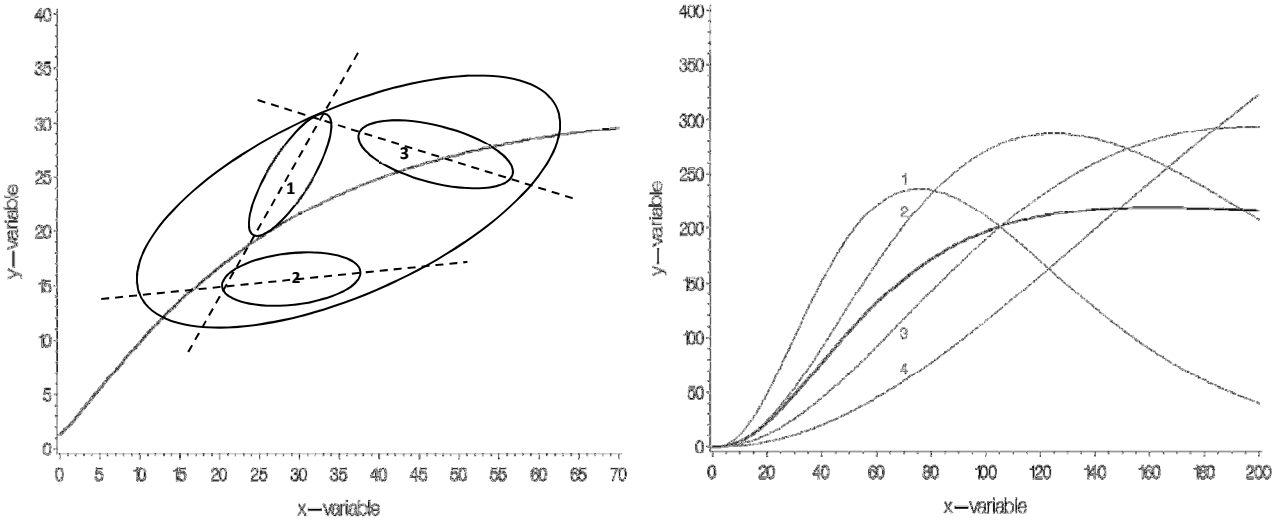

Figure 2. Illustrations of population-based (solid line) and subject-specific (dashed lines) models, where 1, 2, 3 and 4 represent subject-specific models for subjects 1, 2, 3 and 4, respectively.

The key difference between subject-specific models and population-based models is the inclusion of random parameters in subject-specific models. Random parameters are primarily used to:

1). Account for the idiosyncrasies of individual subjects within a population;

2). Account for the remnant impacts of the *x*-variables already included in the model – this can be important when the true model specification is unknown;

3). Account for the impacts of other known and unknown variables left-out by the model without actually requiring these variables to be identified or measured – this can be a good or a bad trait;

4). Alleviate or eliminate entirely the correlation and heteroskedasticity issues commonly occurred in forest modeling from repeatedly measured cross-sectional data.

In essence, a subject-specific model has a unique set of coefficients for each subject in the population. For the subject-specific height-diameter model [6] developed from *m* plots, there are *m* unique sets of coefficients for *m* plots in the population:

$$\begin{aligned}
&\text{Plot 1: } H_{1j} = 1.30 + b_{11}[1-\exp(-b_{21}DBH_{1j})]^{b_{31}} && (b_{11}=b_1+u_{11}, b_{21}=b_2+u_{21}, b_{31}=b_3+u_{31})\\
&\text{Plot 2: } H_{2j} = 1.30 + b_{12}[1-\exp(-b_{22}DBH_{2j})]^{b_{32}} && (b_{12}=b_1+u_{12}, b_{22}=b_2+u_{22}, b_{32}=b_3+u_{32})\\
&\quad\vdots \qquad\qquad \vdots \qquad\qquad \vdots && \qquad\vdots \qquad\qquad \vdots \qquad\qquad \vdots\\
&\text{Plot } m\text{: } H_{mj} = 1.30 + b_{1m}[1-\exp(-b_{2m}DBH_{mj})]^{b_{3m}} && (b_{1m}=b_1+u_{1m}, b_{2m}=b_2+u_{2m}, b_{3m}=b_3+u_{3m})
\end{aligned}$$

Because a unique set of coefficients is developed for each subject in the population, rather than assigning the same set of population-based coefficients obtained for the entire population to each subject in the population,

a subject-specific model is much more flexible and powerful than a population-based model. It can mimic the data trends exhibited by individual subjects more closely. A subject-specific model typically provides more accurate predictions on a subject-specific level than a population-based model.

Of course, subject-specific models also have some drawbacks, including high dependence on local data and more complexity in computation. Sometimes, depending on the model specification and the NMM method used, as well as the specific data involved, subject-specific models could have too much flexibility in curve shape. There are also confusions and inconsistences in the forest literature about the "appropriate" use of the nonlinear mixed model technique in estimating a model and using the estimated model to make predictions.

Both population-based models estimated from the NLS method and subject-specific models estimated from the NMM method were developed in this study. Throughout this study, population-based models were referred to as "base models" and subject-specific models were referred to as "mixed models".

For mixed models, both the first-order (FO) method and the first-order conditional expectation (FOCE) method can be used to estimate parameters. The equations involved in both methods are listed in Table 4, which are taken from Huang et al. (2009a) (with minor corrections). Details about the two methods are also summarized in Meng and Huang (2009), Huang et al. (2009a, 2009b), and Yang and Huang (2011a). Interested readers may wish to read relevant chapters in Davidian and Giltinan (1995), Vonesh and Chinchilli (1997), and Pinheiro and Bates (2004), who provide comprehensive overviews as well as detailed theoretical developments about the NMM methods and their applications in different fields.

The FO method was implemented in this study because the FOCE method did not achieve convergence in most cases. The FOCE method also requires numerical iteration when predicting random parameters, which made the computation more difficult (e.g., see Appendix 1 in Meng and Huang (2009) and Appendix A in Huang et al. (2009a)). Had convergence been achieved in most cases, the FOCE method would likely have been chosen because it has some advantages in maintaining the original shape of model specification, whereas the FO method could alter the original shape of model specification in prediction.

It is worthwhile to emphasize that the equations listed in Table 4 for the FO and FOCE methods are different. Besides different ways of predicting random parameters, each method has its unique equations for, e.g.:

- Making subject-specific predictions; and
- Calculating residuals or prediction errors.

Mixing the equations from the FO and FOCE methods would be mathematically incorrect, even though statistically the consequence of mixing could vary (and be masked) depending on the data and model involved. This explains why sometimes an unsuspecting user could mix the equations and still get a "correct" answer from the wrong equations.

Indeed, when estimating a model and using the estimated model to make predictions, it is essential to ensure that the model prediction procedure to be implemented on new data is consistent with or equivalent to the model prediction procedure embedded in the model estimation process. Otherwise, the results could be inconsistent or simply wrong, and the model fitting statistics would not be as meaningful as they supposed to be. This principle of consistency or equivalence between model estimation and model prediction procedures applies to any type of modeling and model prediction, regardless of whether the models are simple linear models, nonlinear models, or mixed models. Before a model can be used to make prediction elsewhere, it is important to ensure that on the model fitting data, the prediction results are equivalent to the estimation results (e.g., prediction errors are equivalent to residuals).

Table 4. A summary of first-order (FO) and first-order conditional expectation (FOCE) methods.

Method = FO	Method = FOCE
Taylor series expansion	
$\mathbf{y}_i \approx f(\mathbf{x}_i,\mathbf{b}^*,\mathbf{0})+\mathbf{X}_i(\mathbf{b}-\mathbf{b}^*)+\mathbf{Z}_i\mathbf{u}_i+\boldsymbol{\varepsilon}_i$	$\mathbf{y}_i \approx f(\mathbf{x}_i,\mathbf{b}^*,\mathbf{u}_i^*)+\mathbf{X}_i(\mathbf{b}-\mathbf{b}^*)+\mathbf{Z}_i(\mathbf{u}_i-\mathbf{u}_i^*)+\boldsymbol{\varepsilon}_i$
Pseudo-response function	
$\mathbf{y}_i^* = \mathbf{y}_i - f(\mathbf{x}_i,\mathbf{b}^*,\mathbf{0})+\mathbf{X}_i\mathbf{b}^*$	$\mathbf{y}_i^* = \mathbf{y}_i - f(\mathbf{x}_i,\mathbf{b}^*,\mathbf{u}_i^*)+\mathbf{X}_i\mathbf{b}^*+\mathbf{Z}_i\mathbf{u}_i^*$
Linearized model	
$\mathbf{y}_i^* = \mathbf{X}_i\mathbf{b}+\mathbf{Z}_i\mathbf{u}_i+\boldsymbol{\varepsilon}_i$	$\mathbf{y}_i^* = \mathbf{X}_i\mathbf{b}+\mathbf{Z}_i\mathbf{u}_i+\boldsymbol{\varepsilon}_i$
Design matrices	
$\mathbf{X}_i = \left.\frac{\partial f(\mathbf{x}_i,\mathbf{b},\mathbf{u}_i)}{\partial \mathbf{b}'}\right\|_{\hat{\mathbf{b}},\mathbf{0}}$	$\mathbf{X}_i = \left.\frac{\partial f(\mathbf{x}_i,\mathbf{b},\mathbf{u}_i)}{\partial \mathbf{b}'}\right\|_{\hat{\mathbf{b}},\mathbf{u}_i^*}$
$\mathbf{Z}_i = \left.\frac{\partial f(\mathbf{x}_i,\mathbf{b},\mathbf{u}_i)}{\partial \mathbf{u}_i'}\right\|_{\hat{\mathbf{b}},\mathbf{0}}$	$\mathbf{Z}_i = \left.\frac{\partial f(\mathbf{x}_i,\mathbf{b},\mathbf{u}_i)}{\partial \mathbf{u}_i'}\right\|_{\hat{\mathbf{b}},\mathbf{u}_i^*}$
Random parameters predictor	
$\hat{\mathbf{u}}_i = \hat{\mathbf{D}}\mathbf{Z}_i'(\mathbf{Z}_i\hat{\mathbf{D}}\mathbf{Z}_i'+\hat{\mathbf{R}}_i)^{-1}[\mathbf{y}_i - f(\mathbf{x}_i,\hat{\mathbf{b}},\mathbf{0})]$	$\hat{\mathbf{u}}_i = \hat{\mathbf{D}}\mathbf{Z}_i'(\mathbf{Z}_i\hat{\mathbf{D}}\mathbf{Z}_i'+\hat{\mathbf{R}}_i)^{-1}[\mathbf{y}_i - f(\mathbf{x}_i,\hat{\mathbf{b}},\hat{\mathbf{u}}_i)+\mathbf{Z}_i\hat{\mathbf{u}}_i]$
Marginal prediction from fixed parameters	
$\hat{\mathbf{y}}_{i\_fix} = f(\mathbf{x}_i,\hat{\mathbf{b}},\mathbf{0})$	$\hat{\mathbf{y}}_{i\_fix} = f(\mathbf{x}_i,\hat{\mathbf{b}},\mathbf{0})$
Subject-specific prediction with independent and identically distributed error structure ($\hat{\mathbf{R}}_i = \sigma^2\mathbf{I}$)	
$\hat{\mathbf{y}}_i = f(\mathbf{x}_i,\hat{\mathbf{b}},\mathbf{0})+\mathbf{Z}_i\hat{\mathbf{u}}_i$	$\hat{\mathbf{y}}_i = f(\mathbf{x}_i,\hat{\mathbf{b}},\hat{\mathbf{u}}_i)$
Residuals or prediction errors	
$\hat{\mathbf{e}}_i = \mathbf{y}_i - f(\mathbf{x}_i,\hat{\mathbf{b}},\mathbf{0})-\mathbf{Z}_i\hat{\mathbf{u}}_i$	$\hat{\mathbf{e}}_i = \mathbf{y}_i - f(\mathbf{x}_i,\hat{\mathbf{b}},\hat{\mathbf{u}}_i)$
Subject-specific forecast with generalized error structure ($\hat{\mathbf{R}}_i = \sigma^2\boldsymbol{\Psi}$)	
$\hat{\mathbf{y}}_{0i} = f(\mathbf{x}_{0i},\hat{\mathbf{b}},\mathbf{0})+\mathbf{Z}_{0i}\hat{\mathbf{u}}_i+\mathbf{V}'\boldsymbol{\Psi}^{-1}\hat{\mathbf{e}}_i$	$\hat{\mathbf{y}}_{0i} = f(\mathbf{x}_{0i},\hat{\mathbf{b}},\hat{\mathbf{u}}_i)+\mathbf{V}'\boldsymbol{\Psi}^{-1}\hat{\mathbf{e}}_i$

Note:

$\mathbf{y}_i$ and $\hat{\mathbf{y}}_i$ are observed and predicted values for subject *i*;

$\mathbf{x}_i$ is a matrix of covariate(s);

$\mathbf{X}_i$ and $\mathbf{Z}_i$ are partial derivatives with respect to fixed parameters $\mathbf{b}$ and random parameters $\mathbf{u}_i$, respectively;

$\hat{\mathbf{b}}$ and $\hat{\mathbf{u}}_i$ are predictors of $\mathbf{b}$ and $\mathbf{u}_i$, respectively;

$\hat{\mathbf{D}}$ and $\hat{\mathbf{R}}_i$ are estimated variance-covariance matrices for $\mathbf{u}_i$ and $\boldsymbol{\varepsilon}_i$, respectively;

$\hat{\mathbf{y}}_{0i}$, $\mathbf{x}_{0i}$ and $\mathbf{Z}_{0i}$ denote the variables associated with future observations;

$\mathbf{V}$ contains the correlations between the elements of past and future errors;

$\boldsymbol{\Psi}$ is the correlation matrix of past errors.

4.2 Predictions on Model Fitting and Model Application Data

The word "prediction" in forestry often refers to the prediction on new data not used in model fitting, after a model has already been fitted. In this study, predictions were made on model application data not used in model fitting, as well as on model fitting data.

Figure 3 shows the predictions on model fitting and model application data. The model application data could come from different regions, companies, data collection protocols, or a combination of these and other factors (model validation data can be considered a specific case of model application data). While the goodness-of-fit statistics obtained on the modeling data are important indicators of model performance, the true applicability and goodness-of-fit of a model are better judged on model application data not used in model fitting.

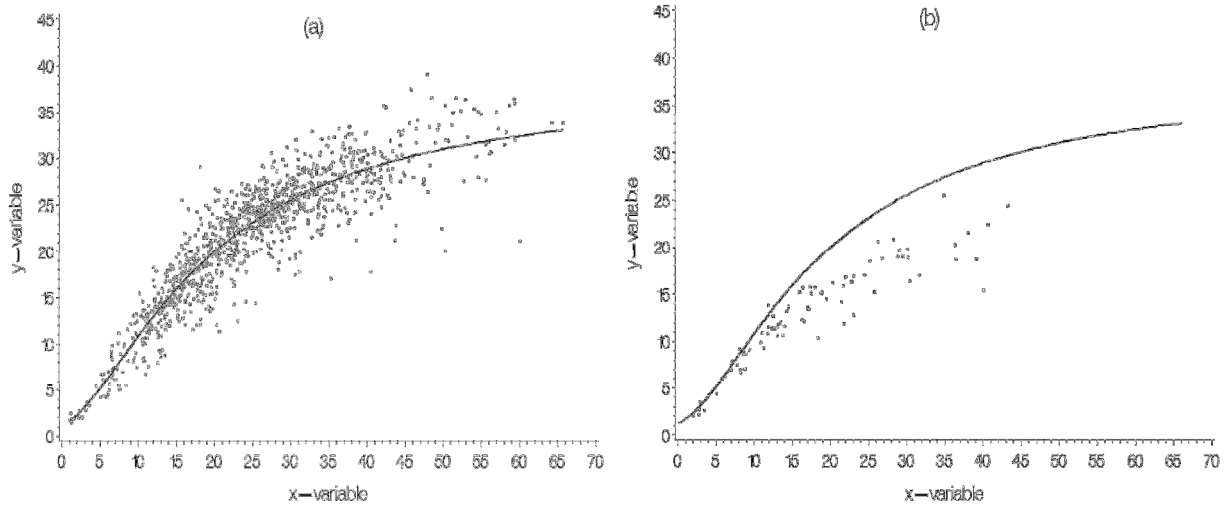

Figure 3. An illustration of predictions on model fitting data (a) and model application data (b). The solid line represents the model fitted on model fitting data.

On the model fitting data, if the model estimation procedure is consistent with or equivalent to the model prediction procedure, which it should be, the residuals and model fitting statistics are identical to the prediction errors and prediction statistics. For a population model estimated by the classical least squares method, as shown in Figure 3(a), the mean bias ($\bar{e}$) is "equivalent" to zero for an appropriately fitted model:

[7] $$\bar{e} = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n_i} e_{ij}}{\sum_{i=1}^{m} n_i} = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n_i}(y_{ij} - \hat{y}_{ij})}{N}$$

where $\bar{e}$ is the mean bias, y_{ij} and $\hat{y}_{ij}$ are the *j*th observed and predicted *y*-values for the *i*th subject, *i* = 1, 2, …, *m*, *j* = 1, 2, …, n_i, *m* is the number of subjects in the population, n_i is the number of observations in the *i*th subject, and *N* is the total number of observations in the population. Since $\bar{e}$ is equivalent to zero, the percent mean bias defined in [8] is also equivalent to zero on the model fitting data:

[8] $$\bar{e}\% = \frac{\bar{e}}{\bar{y}} \times 100 \qquad (\bar{y} = \frac{1}{N}\sum_{i=1}^{m}\sum_{j=1}^{n_i} y_{ij})$$

where $\bar{y}$ is the average of the observed values.

Several important notes regarding the mean bias and percent mean bias defined in [7]-[8] are described here:

1. For linear models fitted by the ordinary least squares method, the mean bias ($\bar{e}$) and the percent mean bias ($\bar{e}\%$) are guaranteed to be zeros on the model fitting data. This is one of the most noteworthy characteristics of the least squares principle. For linear models fitted by the linear mixed model technique, the $\bar{e}$ and $\bar{e}\%$ are also guaranteed to be zeros on the model fitting data.

2. For nonlinear models fitted by the ordinary nonlinear least squares method or the nonlinear mixed model technique, both $\bar{e}$ and $\bar{e}\%$ are not guaranteed to be zeros on the model fitting data. But they should be "very close" or "asymptotically approximately equivalent" to zeros when the sample size is "large enough" and when the models are appropriately specified and fitted.

   Here, the word "asymptotically" means that the $\bar{e}$ approaches zero when the sample size becoming increasingly large. The word "approximately" means that even if the asymptotic condition is met, the zero equivalence is still "approximately" true, because many of the nonlinear computations involved in nonlinear estimation and prediction need to be approximated, e.g., by linearization and/or by dropping some "negligible" terms such as quadratics, cubics, and cross-products.

   Of course, in practice, we seldom argue about the precise semantics of these words. We generally consider the $\bar{e}$ to be "close enough" or "statistically equivalent" to zero for an appropriately specified and fitted nonlinear model as long as the sample size is "reasonably large" or "large enough". But how "close" and what is "reasonably large" depend on a number of non-trivial factors, amongst others, the approximation method used, the complexity of the relationship to be built and the number of model parameters and variables involved. No definitive answer is given in the literature. A threshold of 30 reasonably distributed observations has frequently been considered "large enough" in many statistical applications for some relatively simple relationships.

3. In this study, we define the closeness to zero on the model fitting data as "$\bar{e}\%$ less than half-a-percent". That is, the percent mean bias defined in [8] must be less than 0.5% (1/200) on either side of the observed mean to be considered "close" or "equivalent" to zero. Otherwise, the model is considered to have a non-zero mean bias or simply biased. We also define a sample size of "100 reasonably spread observations from 30 subjects per *x*-variable" to be the minimum sample size for "large enough". For instance, for a model involving three *x*-variables, 300 reasonably spread observations from 90 plots are considered to be the minimum "large enough" sample size for estimating a model, which typically has less than 10 parameters.

4. On independent model application data not used in model fitting, following the common practice in Alberta, if the percent mean bias is within 10% of the observed mean on either side, the model is generally considered acceptable, provided that it makes biological sense and that the graphics showing the prediction errors do not indicate any gross abnormality. Of course, it is always possible to adjust the predictions on model application data such that the mean bias and percent mean bias of the adjusted predictions equal to zero exactly. But the adjusted "bias-free" predictions do not necessarily lead to the "best" or "most accurate" predictions, as the "best" or "most accurate" predictions are not determined by the mean bias alone.

5. When a fitted model is used to make predictions on model application data not used in model fitting, the $\bar{e}$ does not equal to zero (except by chance). In addition, the model application data may not follow the general trend exhibited by the fitted model. These are illustrated in Figure 3(b). They can occur if a provincial model is applied to a specific region or a sub-population from a company.

4.3 Population-Based Predictions from Base Model

4.3.1 Unadjusted Population Prediction

Unadjusted population prediction refers to the direct use of a base model estimated from the NLS method for the prediction of population averages from the known *x*-values. As an example, assuming the base model (solid line) in Figure 4 depicts the height-diameter relationship for white spruce in Alberta, tree height H (*y*-variable) can be predicted directly from the base model from the observed tree DBH (*x*-variable), regardless of the plot in which the DBH is observed from. For the example population in Figure 4, the predictions obtained directly from the base model deviate from the actual data considerably.

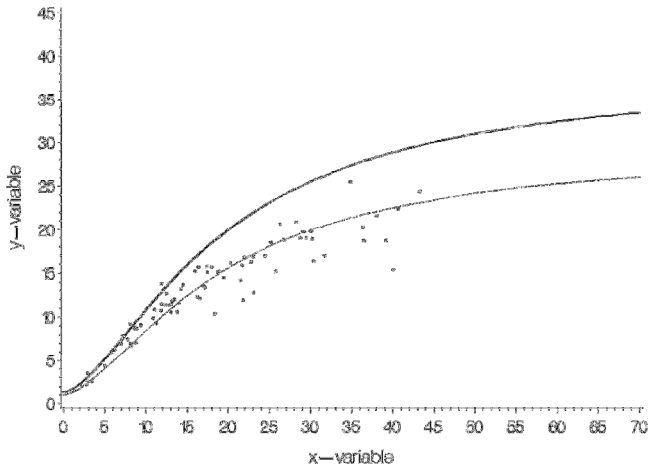

Figure 4. An illustration of unadjusted (solid line) and adjusted (dashed line) population predictions from a base model applied to a model application data set. The adjusted predictions were obtained from the proportional adjustment method.

4.3.2 Adjusted Population Prediction

The unadjusted population prediction obtained directly from a fitted base model can be poor. Some better predictions may be obtained if we could adjust the model. The proportional adjustment method is the simplest and most effective method in many cases for obtaining adjusted predictions (e.g., Figure 4). This method is implemented through the calculation of a ratio, called the proportional adjustment ratio (PAR) in this study, from all data in the model application population:

$$[9] \qquad \text{PAR} = \frac{\bar{y}}{\bar{\hat{y}}} = \frac{\frac{1}{N}\sum_{i=1}^{m}\sum_{j=1}^{n_i} y_{ij}}{\frac{1}{N}\sum_{i=1}^{m}\sum_{j=1}^{n_i} \hat{y}_{ij}} = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n_i} y_{ij}}{\sum_{i=1}^{m}\sum_{j=1}^{n_i} \hat{y}_{ij}}$$

where PAR is the proportional adjustment ratio for the population in which the base model is applied, y_{ij} is the *j*th observed *y*-value for the *i*th subject in the population and $\hat{y}_{ij}$ is its prediction from the base model, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n_i$, m is the number of subjects in the population, n_i is the number of observations in the *i*th subject, and N is the total number of observations in the population. Re-arrange [9] produces:

$$[10] \qquad \sum_{i=1}^{m}\sum_{j=1}^{n_i} y_{ij} - \text{PAR} \cdot \sum_{i=1}^{m}\sum_{j=1}^{n_i} \hat{y}_{ij} = 0$$

which implies that the summation of the observed values equals to the PAR times the summation of the predicted values. If we define the adjusted predictions as:

[11] $$\hat{y}_{ij\_adj} = \text{PAR} \cdot \hat{y}_{ij}$$

We have:

[12] $$\overline{\hat{y}}_{ij\_adj} = \frac{1}{N}\sum_{i=1}^{m}\sum_{j=1}^{n_i}\hat{y}_{ij\_adj} = \frac{\text{PAR}}{N}\sum_{i=1}^{m}\sum_{j=1}^{n_i}\hat{y}_{ij} = \text{PAR} \cdot \overline{\hat{y}} = \overline{y}$$

Therefore:

[13] $$\overline{y} - \text{PAR} \cdot \overline{\hat{y}} = 0 \ \ (\text{or } \overline{y} - \overline{\hat{y}}_{ij\_adj} = 0)$$

which means that the mean of the observed values equals to the mean of the proportionally adjusted predictions from the base model. What is really important about the expressions given in [9]-[13] is that, they all suggest that the mean bias ($\overline{e}_{ij\_adj}$) of the adjusted predictions for the model application population is zero:

[14] $$\overline{e}_{ij\_adj} = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n_i} e_{ij\_adj}}{N} = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n_i}(y_{ij} - \hat{y}_{ij\_adj})}{N} = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n_i} y_{ij} - \text{PAR}\cdot\sum_{i=1}^{m}\sum_{j=1}^{n_i}\hat{y}_{ij}}{N} = 0$$

As an example, the adjusted predictions (dashed line) in Figure 4 obtained from the proportional adjustment method are guaranteed to have a zero mean bias.

The proportional adjustment method relies on the assumption which requires that the new data a model is to be applied to are (more or less) proportional to the model across the *x*-range. If this assumption does not hold, a mean bias of zero is still guaranteed for adjusted predictions, but it may not mean a good fit to the data.

4.4 Subject-Specific Predictions

Regardless of the adjustment, population-based predictions ignore the idiosyncrasies of individual subjects within a population. To account for the intrinsic variation and to reflect the polymorphic growth of individual subjects within a population, different procedures can be used to make subject-specific predictions. The most common ones are described here.

4.4.1 Proportional Adjustment of Base Model

Calculation for adjusted plot-specific prediction from a base model is similar to that for adjusted population prediction, except that it is carried out plot-by-plot rather than for the entire population. For instance, a plot-specific PAR_i for plot *i* is calculated by:

[15] $$\text{PAR}_i = \frac{\overline{y}_i}{\overline{\hat{y}}_i} = \frac{\sum_{j=1}^{n_i} y_{ij}}{\sum_{j=1}^{n_i}\hat{y}_{ij}}$$

where PAR_i is the proportional adjustment ratio for plot *i* in the population, and all other variables are as defined before. Following equations [10] to [14], for any specific plot *i* in the population, we have:

[16] $$\sum_{j=1}^{n_i} y_{ij} - \text{PAR}_i \cdot \sum_{j=1}^{n_i} \hat{y}_{ij} = 0$$

[17] $$\hat{y}_{ij\_adj} = \text{PAR}_i \cdot \hat{y}_{ij}$$

[18] $$\bar{\hat{y}}_{ij\_adj} = \frac{1}{n_i}\sum_{j=1}^{n_i} \hat{y}_{ij\_adj} = \frac{\text{PAR}_i}{n_i}\sum_{j=1}^{n_i} \hat{y}_{ij} = \text{PAR}_i \cdot \bar{\hat{y}}_i = \bar{y}_i$$

[19] $$\bar{y}_i - \text{PAR}_i \cdot \bar{\hat{y}}_i = 0 \ \ (\text{or } \bar{y}_i - \bar{\hat{y}}_{ij\_adj} = 0)$$

[20] $$\bar{e}_{ij\_adj} = \frac{\sum_{j=1}^{n_i} e_{ij\_adj}}{n_i} = \frac{\sum_{j=1}^{n_i} (y_{ij} - \hat{y}_{ij\_adj})}{n_i} = \frac{\sum_{j=1}^{n_i} y_{ij} - \text{PAR}_i \cdot \sum_{j=1}^{n_i} \hat{y}_{ij}}{n_i} = 0$$

where $\bar{e}_{ij\_adj}$ is the mean bias of the adjusted prediction errors for plot *i* in the population.

Adjusted predictions obtained for each plot from the plot-specific proportional adjustment method are guaranteed to have a zero mean bias. Since the mean bias for each plot in the population is zero, the mean bias of the adjusted predictions for the entire population is also guaranteed to be zero.

While the proportional adjustment of a base model is the simplest and most effective method in many cases for deriving subject-specific predictions, the method relies heavily on the assumption which requires that the subject-specific data to be (more or less) proportional to the population averages represented by the base model across the *x*-range. If this assumption is violated or difficult to know, a mean bias of zero is still guaranteed for adjusted predictions, but it may not mean a good fit for a subject. This is illustrated in Figure 5, where the proportionality between the population averages and subject-specific data does not exist (a), or is difficult to know due to the narrow range of the data (b).

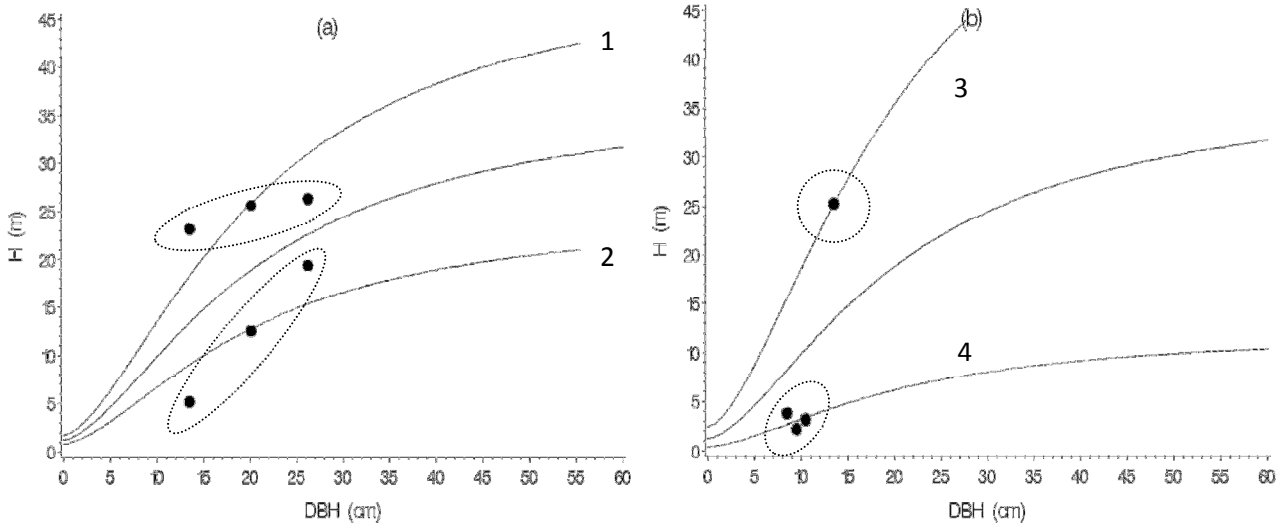

Figure 5. An illustration of subject-specific predictions (dashed lines) obtained from the proportional adjustment method, where 1, 2, 3, and 4 represent four subjects, H and DBH are tree height and diameter, and the solid lines represent the population averages from a base model. The graphs show that the proportionality between the population averages and subject-specific data does not exist (a), or is difficult to know due to the narrow range of the data (b).

4.4.2 Mixed Model

The equations associated with subject-specific predictions from nonlinear mixed model methods are listed in Table 4. As noted earlier (Section 4.1), mixed models contain both fixed and random parameters, where the fixed parameters refer to population-level parameters common to all subjects in the population and the random parameters refer to subject-specific parameters unique for each subject in the population. Because of the inclusion of subject-specific random parameters, a mixed model essentially provides a localized fit to each and every subject in the population. It can closely track the data trends of individual subjects in the population.

In addition, using the mixed model technique, a mixed model fit for any subject in the population can be supplemented and enhanced during the estimation and prediction processes by the information about the other subjects in the population. This effectively allows a subject to "borrow strength" across all subjects in the population to get a better local fit for the subject.

4.4.3 Proportional Adjustment of Mixed Model

Although for most practical purposes the mixed model technique can generally be considered unbiased, the unbiasedness property, and indeed, many other properties associated with the NMM technique, hold only in an "asymptotically approximated" sense. When the sample sizes of the subjects are not "large enough", or when model specification is problematic, mixed model can produce biased predictions.

The potential bias of the NMM leads naturally to the third method of subject-specific prediction, termed the combination method in this study. The combination method combines the mixed model method with the subject-specific proportional adjustment method.

For the combination method, the biased subject-specific predictions from a mixed model are scaled up or down by a constant proportion determined by the subject-specific proportional adjustment ratio PAR_i. The PAR_i for subject *i* is calculated according to [15], with the predictions derived directly from a NMM method (e.g., Table 4). The predictions derived directly from the NMM method are called unadjusted mixed model predictions. The scaled predictions obtained via $\hat{y}_{ij\_adj} = \mathrm{PAR}_i \cdot \hat{y}_{ij}$ are called adjusted mixed model predictions. Figure 6 shows the unadjusted and adjusted predictions for four subjects with small sample sizes.

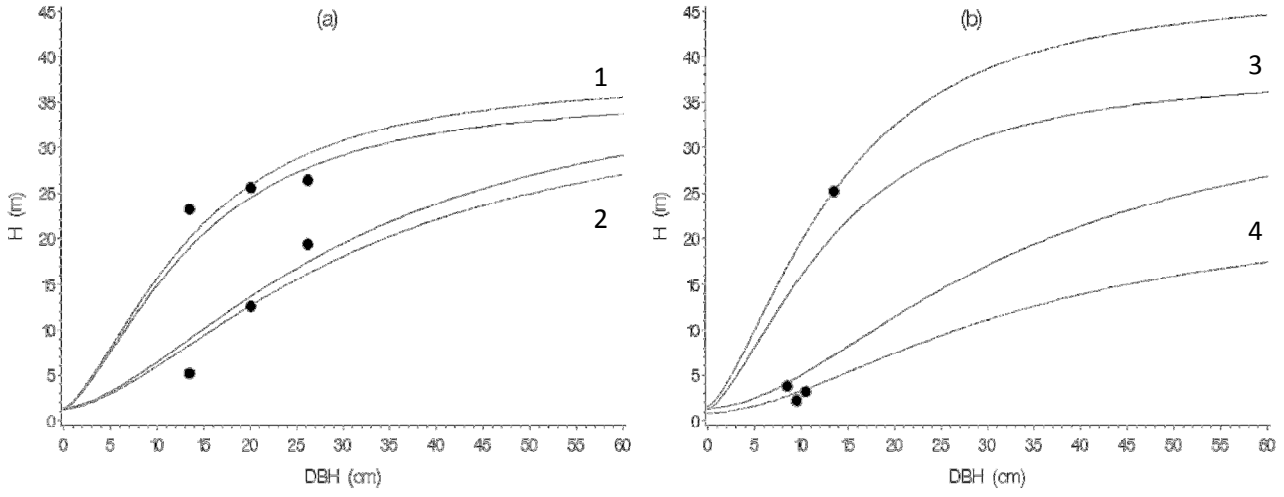

Figure 6. Unadjusted (solid lines) and adjusted (dashed lines) predictions for four subjects, where the unadjusted predictions were obtained directly from the mixed model method, and the adjusted predictions were obtained from the combination method.

In Figure 6, the unadjusted predictions were obtained directly from the mixed model method (in this case, the FO method), and the adjusted predictions were obtained from the combination method (actual data and computations for the examples shown in Figure 6 are available to interested readers). Note that for any subject with one observation only (e.g., subject 3 in Figure 6(b)), the adjusted predictions always pass through that observation, whereas the unadjusted predictions do not pass through that observation.

Following the logic embedded in equations [16]-[20], it can be inferred that the adjusted mixed model predictions obtained from the combination method are guaranteed to have a zero mean bias. This is true on both the subject-specific level and the population level.

The essence of the combination method is to utilize the power of the mixed model method to track the trends of subject-specific data in a population, while simultaneously combining this power with the proportional adjustment method to alleviate the sample size and nonlinear approximation and asymptotic issues (Huang 2008). This "deadly" (i.e., towards the bias) combination of the mixed model method and proportional adjustment method allows for a mixed model to fit any subject-specific data as close as possible. With the combination method it is always possible, provided that model specification does not become a limiting factor, to achieve a "bias-free" fit that closely mimics the data of any subject in a population, even though the "bias-free" fit still does not necessarily mean the "best" fit.

The performance of the combination method is affected by the ability of the mixed model to track subject-specific data trends. If the mixed model tracks the trends well, which should be the case for an appropriately specified and fitted mixed model, the combination method will likely produce the "best" results. If the mixed model does not track the trends well, the predictions from the combination method may not be the best even though the mean bias is still guaranteed to be zero. Knowledge, skill and experience in selecting the right mixed model specification play a more important role than fit statistics in this regard. The selected model specification should have enough but not too much flexibility for the relevant data and relationship to be built.

4.5 Prediction Types Assessed for Height-Diameter Models

For the height-diameter models developed in this study, four prediction types are assessed:

1. B-U: denotes "base model, unadjusted", i.e., the predictions are obtained directly from the base model fitted from the ordinary NLS method.

2. B-A: denotes "base model, adjusted", i.e., the predictions are obtained first from the base model, then adjusted through the proportional adjustment method.

3. M-U: denotes "mixed model, unadjusted", i.e., the predictions are obtained directly from the first-order method of the NMM technique.

4. M-A: denotes "mixed model, adjusted", i.e., the predictions are obtained first from the first-order method of the NMM technique, then adjusted through the proportional adjustment method.

For population-based predictions, only B-U and B-A types of predictions are relevant. They are illustrated in Figure 7 (a, b). The PAR for population-based predictions is calculated based on all data in the population.

For subject-specific predictions, all four types of predictions could be used. The PAR_i used in B-A and M-A types of predictions is calculated based on the data from plot *i* only. Figure 7 (c, d, e, f) illustrates the four types of subject-specific predictions.

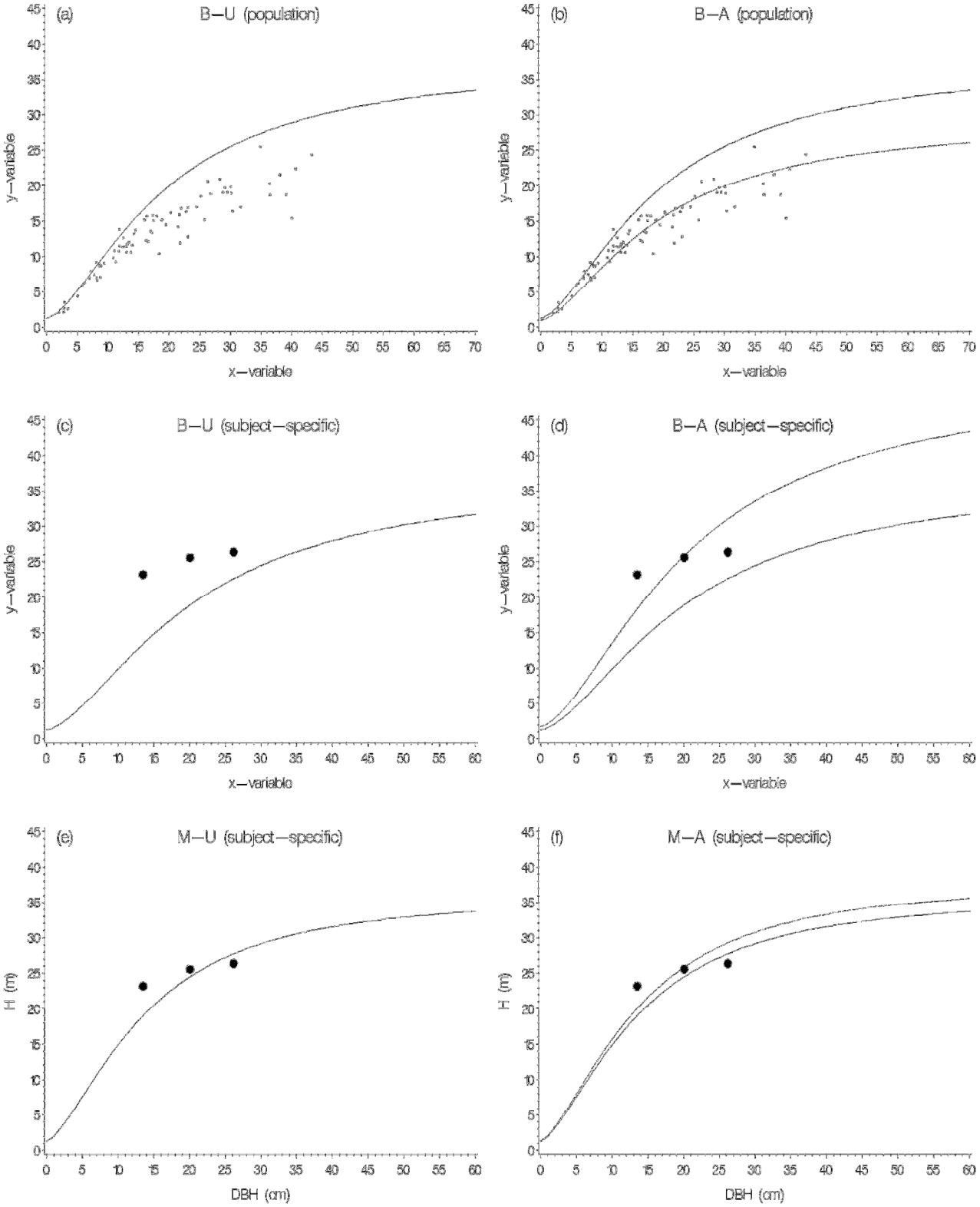

Figure 7. An illustration of prediction types: B-U (base model, unadjusted), B-A (base model, adjusted), M-U (mixed model, unadjusted), and M-A (mixed model, adjusted). Solid lines represent unadjusted predictions. Dashed lines represent adjusted predictions. For a comparison, relevant solid lines are also plotted on the right-hand side graphs.

4.6 Goodness-of-Fit Measures

In order to judge the goodness-of-fit (or goodness-of-prediction) of different models and prediction types, a variety of goodness-of-fit measures listed in Table 5 were calculated. For mixed models on model fitting data, Akaike information criterion (AIC) and Schwarz's Bayesian information criterion (BIC) were also calculated as a part of the standard model fitting output.

Table 5. Goodness-of-fit measures calculated in this study.

Goodness-of-fit measure	Computation formula
1. Mean bias (or mean error)	$\bar{e} = \frac{1}{N}\sum_{i=1}^{m}\sum_{j=1}^{n_i}(y_{ij} - \hat{y}_{ij}) = \frac{1}{N}\sum_{i=1}^{m}\sum_{j=1}^{n_i} e_{ij}$
2. Percent mean bias	$\bar{e}\% = \frac{\bar{e}}{\bar{y}} \times 100$
3. Standard deviation	$SD = \sqrt{\frac{1}{N-1}\sum_{i=1}^{m}\sum_{j=1}^{n_i}(e_{ij} - \bar{e})^2}$
4. Mean absolute deviation	$MAD = \frac{1}{N}\sum_{i=1}^{m}\sum_{j=1}^{n_i}\left\|y_{ij} - \hat{y}_{ij}\right\|$
5. Mean square error (on model fitting data)	$MSE = \frac{1}{N\text{-}p}\sum_{i=1}^{m}\sum_{j=1}^{n_i}(y_{ij} - \hat{y}_{ij})^2$
6. Mean square error (on model application data)	$MSE = \frac{1}{N}\sum_{i=1}^{m}\sum_{j=1}^{n_i}(y_{ij} - \hat{y}_{ij})^2$
7. Coefficient of determination	$R^2 = 1 - \sum_{i=1}^{m}\sum_{j=1}^{n_i}(y_{ij} - \hat{y}_{ij})^2 \Big/ \sum_{i=1}^{m}\sum_{j=1}^{n_i}(y_{ij} - \bar{y})^2$
8. Concordance correlation coefficient	$CC = 1 - \frac{\sum_{i=1}^{m}\sum_{j=1}^{n_i}(y_{ij} - \hat{y}_{ij})^2}{\sum_{i=1}^{m}\sum_{j=1}^{n_i}(y_{ij} - \bar{y})^2 + \sum_{i=1}^{m}\sum_{j=1}^{n_i}(\hat{y}_{ij} - \bar{\hat{y}})^2 + N(\bar{y} - \bar{\hat{y}})^2}$
9. Mean percent error	$MPE = \frac{1}{N}\sum_{i=1}^{m}\sum_{j=1}^{n_i}(\frac{y_{ij} - \hat{y}_{ij}}{y_{ij}}) \times 100$
10. Mean absolute percent error	$MAPE = \frac{1}{N}\sum_{i=1}^{m}\sum_{j=1}^{n_i}\left\|\frac{y_{ij} - \hat{y}_{ij}}{y_{ij}}\right\| \times 100$
11. Number of absolute percent errors >10%	$e_{10} = \frac{number\ \ of\ \left\|PE_{ij}\right\| > 10}{N}$, $PE_{ij} = (\frac{y_{ij} - \hat{y}_{ij}}{y_{ij}}) \times 100$
12. Overall accuracy	$\delta = \bar{e}^2 + SD^2$
13. Akaike information criterion	$AIC = -2\ln(L) + 2P$
14. Schwarz's Bayesian information	$BIC = -2\ln(L) + P\ln(m)$
Grand means and total number	$\bar{y} = \frac{1}{N}\sum_{i=1}^{m}\sum_{j=1}^{n_i} y_{ij}$, $\bar{\hat{y}} = \frac{1}{N}\sum_{i=1}^{m}\sum_{j=1}^{n_i}\hat{y}_{ij}$, $N = \sum_{i=1}^{m} n_i$

In Table 5, y_{ij} and $\hat{y}_{ij}$ are the *j*th observed and predicted values for the *i*th subject, *i* = 1, 2, …, *m*, *j* = 1, 2, …, n_i, *m* is the number of subjects in the data, n_i is the number of observations for the *i*th subject, $\bar{y}$ and $\bar{\hat{y}}$ are the grand means of observed and predicted values, *N* is the total number of observations, *L* is the maximized value of the likelihood function for the estimated model, *p* (lowercase) is the number of fixed parameters, *P* (uppercase) is the total number of effective parameters in mixed model estimation (includes fixed parameters, variance-covariance components of the random parameters, plus the residual variance component), and AIC and BIC are information criteria used for mixed model on model fitting data only.

Many of the measures listed in Table 5 have been used in various studies for determining the goodness-of-fit of a model. They have also been used for model selection. The "number of absolute percent errors greater than 10%" (e_{10}) is a new measure used to represent the distribution/spread of the prediction errors. It describes the proportion of predictions whose errors are greater than ±10% of the observed values. A larger e_{10} corresponds to a poorer model.

Two similar measures were also calculated. One is the number of absolute percent errors greater than 5% (e_5), and the other is the number of absolute percent errors greater than 20% (e_{20}):

$$e_5 = \frac{\text{number of } |PE_{ij}| > 5}{N} \qquad e_{20} = \frac{\text{number of } |PE_{ij}| > 20}{N} \qquad \left(PE_{ij} = \left(\frac{y_{ij} - \hat{y}_{ij}}{y_{ij}}\right) \times 100\right)$$

where PE_{ij} is the percent error for the *j*th observation in the *i*th plot, and *N* is the total number of observations (see Table 5). Results of the calculated e_5 and e_{20} statistics are available but not listed in this study.

Below are some important notes with regard to goodness-of-fit measures:

1. Each goodness-of-fit measure listed in Table 5 has its own pros and cons, and each usually reflects one aspect of a fitted model. In general, the overall accuracy measure δ ($\delta = \bar{e}^2 + SD^2$), which combines the mean bias ($\bar{e}^2$) and the variance of the prediction errors (SD^2), is considered a good overall indicator of accuracy. However, to get a more complete picture about the fit of a model, other goodness-of-fit measures should also be looked at, together with relevant graphics. Picking only a single measure to represent the goodness-of-fit of a model may not be enough if other measures particularly graphics are ignored.

2. It is worthwhile to emphasize that:

   Accuracy = Bias + Precision

   which means that the goodness of a model cannot be judged by either "bias" or "precision" alone. An "unbiased" model could give poor predictions if the precision is low. The same is true if the precision is high but the model is biased. One shall not over-emphasize nor be duped by "unbiased models" or "unbiased something", for they could still be very bad. Our goal is to minimize bias, while maximizing precision (i.e., by minimizing the variance or standard deviation).

3. The goodness-of-fit measures listed in Table 5 apply to the population as a whole. To calculate the goodness-of-fit measures for any subject *i* in the population (*i* = 1, 2, …, *m*), the formulas given in Table 5 are still correct, except that the calculation will be done on the n_i observations from subject *i*, not the *N* observations from all subjects. For instance, subject-specific mean bias, standard deviation and overall accuracy for subject *i* are calculated as follows:

$$\bar{e}_i = \frac{1}{n_i}\sum_{j=1}^{n_i}(y_{ij} - \hat{y}_{ij}) = \frac{1}{n_i}\sum_{j=1}^{n_i} e_{ij} \qquad SD_i = \sqrt{\frac{1}{n_i - 1}\sum_{j=1}^{n_i}(e_{ij} - \bar{e}_i)^2} \qquad \delta_i = \bar{e}_i^{\,2} + SD_i^{\,2}$$

Other subject-specific goodness-of-fit measures can be calculated in a similar manner. They can be used to assess the fit of a model on individual subjects within a population. The distributions of these subject-specific measures from all subjects in the population can be graphed, compared and analyzed further to detect any possible outlier and abnormality.

4. While subject-specific goodness-of-fit measures are valuable in assessing subject-specific goodness-of-fit of a model, they could be overwhelming considering if there are hundreds or even thousands of subjects in a population (e.g., Table 3). Hence, some summary goodness-of-fit statistics for all subjects combined would be very useful. In this study, these summary statistics were calculated by the formulas given in Table 5, with the $\hat{y}_{ij}$ obtained from subject-specific predictions. The summary statistics were <u>not averaged</u> from the goodness-of-fit statistics obtained from individual subjects.

5. Some readers may wish to calculate the summary statistics by taking the averages of the goodness-of-fit statistics obtained from individual subjects. <u>This is not recommended in general</u>. At a minimum, if such summary statistics are to be calculated at all, they shall be weighted averages of the goodness-of-fit statistics from individual subjects, weighted by the number of observations in each subject. For instance, for a population with eight subjects:

$$\bar{e} = \frac{\bar{e}_1 \times n_1 + \bar{e}_2 \times n_2 + \cdots + \bar{e}_8 \times n_8}{n_1 + n_2 + \cdots + n_8} \qquad \bar{e}\% = \frac{\bar{e}_1\% \times n_1 + \bar{e}_2\% \times n_2 + \cdots + \bar{e}_8\% \times n_8}{n_1 + n_2 + \cdots + n_8}$$

$$SD = \frac{SD_1 \times n_1 + SD_2 \times n_2 + \cdots + SD_8 \times n_8}{n_1 + n_2 + \cdots + n_8} \qquad R^2 = \frac{R_1^2 \times n_1 + R_2^2 \times n_2 + \cdots + R_8^2 \times n_8}{n_1 + n_2 + \cdots + n_8}$$

where $\bar{e}_i$, $\bar{e}_i\%$, SD_i, R_i^2 and n_i are subject-specific mean bias, percent mean bias, standard deviation, coefficient of determination and sample size, respectively (i =1, 2,…, 8).

6. In addition to the goodness-of-fit measures listed in Table 5, some other goodness-of-fit measures based on prediction, tolerance, and confidence intervals could also be constructed once the prediction errors were obtained. For instance, for a subject with n_i observations, the prediction interval for a future observation, the tolerance interval to contain at least proportion γ of the subject, and the confidence intervals for subject mean bias ($\bar{e}_i$) and standard deviation (SD_i), can be calculated respectively as follows (α=0.05 throughout this study):

$$\bar{e}_i \pm t_{(1-\alpha/2),(n_i-1)} SD_i \sqrt{1 + \frac{1}{n_i}} \qquad \bar{e}_i \pm z_{\frac{1+\gamma}{2}} \left(1 + \frac{1}{2n_i}\right) SD_i \sqrt{\frac{n_i - 1}{\chi^2_{\alpha,(n_i-1)}}}$$

$$\bar{e}_i \pm t_{(1-\alpha/2),(n_i-1)} \frac{SD_i}{\sqrt{n_i}} \qquad SD_i \sqrt{\frac{n_i - 1}{\chi^2_{(1-\alpha/2),(n_i-1)}}} \rightarrow SD_i \sqrt{\frac{n_i - 1}{\chi^2_{(\alpha/2),(n_i-1)}}}$$

where t, χ^2 and z are Student's t, χ^2 and standard normal distributions, respectively. The calculations assume that the prediction errors are normally distributed. If the normality assumption is not met, approximated distribution-free intervals could be constructed using a 10% trimmed mean and a jackknifed standard deviation (e.g., Huang et al. 2009a).

Since hypothesis testing and interval estimation were not commonly used as goodness-of-fit measures in most practical circumstances, we only listed the equations here for those who might wish to consider them as potential goodness-of-fit measures (selected results of the interval-based goodness-of-fit statistics for the models developed in this study are available to interested readers).

Mean squared error (MSE) or its square root (root mean squared error, RMSE) plays a similar role to that of the overall accuracy measure δ. It incorporates both the variance and the bias of the errors. The calculation for MSE was slightly different in different studies. In this study, the MSE on model fitting data is calculated consistently following Measure #5 in Table 5, for both base and mixed models:

$$\text{MSE} = \frac{1}{N\text{-}p} \sum_{i=1}^{m} \sum_{j=1}^{n_i} (y_{ij} - \hat{y}_{ij})^2$$

where *p* is the number of fixed parameters.

The MSE on model application data is calculated consistently following Measure #6 in Table 5, for both base and mixed models:

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^{m} \sum_{j=1}^{n_i} (y_{ij} - \hat{y}_{ij})^2$$

In regression literature, the denominator in MSE on model fitting data denotes the degrees of freedom, i.e., the total number of observations reduced by the number of "model parameters" had to be estimated in obtaining the $\hat{y}_{ij}$. For a base model, the number of model parameters had to be estimated in obtaining the $\hat{y}_{ij}$ is the number of fixed parameters *p* (lowercase *p*) in the model. However, for a mixed model, the number of model parameters had to be estimated in obtaining the $\hat{y}_{ij}$ is *P* (uppercase *P*), where *P* is the total number of effective parameters in mixed model estimation. The effective parameters in mixed model estimation include fixed parameters, variance-covariance components of the random parameters, plus the residual variance component. Hence, for a mixed model, the MSE could also be calculated as follows:

$$\text{MSE} = \frac{1}{N\text{-}P} \sum_{i=1}^{m} \sum_{j=1}^{n_i} (y_{ij} - \hat{y}_{ij})^2$$

The MSEs calculated in this manner are not presented in this report but are available to interested readers.

5. Application Examples on a Small Data Set

White spruce height-diameter data from a small population of eight plots were used to demonstrate the application of the fitted models. The actual data are listed in Table 6 (and shown later in Figures 8 and 9). We purposely chose this small population with eight plots only such that the procedures and computations demonstrated here could be duplicated step-by-step by interested readers. The procedures could also serve as an integral part of model validation (Huang 2002). Two larger populations were used later to further demonstrate model application in Section 6.

Table 6. White spruce height-diameter data and predictions from the base model [2].

Plot	Tree	DBH	H	Population-based					Plot-specific					
				$\hat{H}$	PAR	$\hat{H}_{adj}$	e	e_{adj}	$\hat{H}$	$\overline{H}$	$\overline{\hat{H}}$	PAR	$\hat{H}_{adj}$	e_{adj}
1	1	32.5	24.70	25.49	0.9857	25.13	-0.79	-0.43	25.49	21.53	21.42	1.0055	25.63	-0.93
1	2	11.9	14.90	11.82	0.9857	11.65	3.08	3.25	11.82	21.53	21.42	1.0055	11.89	3.01
1	3	36.6	25.00	26.93	0.9857	26.55	-1.93	-1.55	26.93	21.53	21.42	1.0055	27.08	-2.08
2	1	30.0	23.22	24.46	0.9857	24.12	-1.24	-0.90	24.46	17.99	22.32	0.8060	19.72	3.50
2	2	22.0	12.76	20.17	0.9857	19.89	-7.41	-7.13	20.17	17.99	22.32	0.8060	16.26	-3.50
3	1	29.8	22.72	24.38	0.9857	24.03	-1.66	-1.31	24.38	21.75	25.32	0.8588	20.93	1.79
3	2	34.6	20.77	26.27	0.9857	25.89	-5.50	-5.12	26.27	21.75	25.32	0.8588	22.56	-1.79
4	1	38.7	16.90	27.57	0.9857	27.18	-10.67	-10.28	27.57	16.90	27.57	0.6130	16.90	0.00
5	1	25.7	17.10	22.37	0.9857	22.05	-5.27	-4.95	22.37	17.10	19.08	0.8961	20.05	-2.95
5	2	11.7	11.60	11.62	0.9857	11.46	-0.02	0.14	11.62	17.10	19.08	0.8961	10.41	1.19
5	3	27.4	22.60	23.25	0.9857	22.92	-0.65	-0.32	23.25	17.10	19.08	0.8961	20.84	1.76
6	1	13.5	19.20	13.39	0.9857	13.20	5.81	6.00	13.39	21.18	19.03	1.1133	14.90	4.30
6	2	19.0	19.20	18.07	0.9857	17.82	1.13	1.38	18.07	21.18	19.03	1.1133	20.12	-0.92
6	3	26.2	24.40	22.64	0.9857	22.32	1.76	2.08	22.64	21.18	19.03	1.1133	25.20	-0.80
6	4	20.1	22.60	18.88	0.9857	18.61	3.72	3.99	18.88	21.18	19.03	1.1133	21.02	1.58
6	5	18.5	19.80	17.69	0.9857	17.44	2.11	2.36	17.69	21.18	19.03	1.1133	19.70	0.10
6	6	27.9	21.90	23.50	0.9857	23.16	-1.60	-1.26	23.50	21.18	19.03	1.1133	26.16	-4.26
7	1	30.4	29.10	24.64	0.9857	24.29	4.46	4.81	24.64	24.15	21.26	1.1361	27.99	1.11
7	2	19.2	18.50	18.22	0.9857	17.96	0.28	0.54	18.22	24.15	21.26	1.1361	20.70	-2.20
7	3	20.8	21.80	19.37	0.9857	19.10	2.43	2.70	19.37	24.15	21.26	1.1361	22.01	-0.21
7	4	26.5	27.20	22.80	0.9857	22.47	4.40	4.73	22.80	24.15	21.26	1.1361	25.90	1.30
8	1	26.2	25.91	22.64	0.9857	22.32	3.27	3.59	22.64	20.34	20.44	0.9947	22.52	3.39
8	2	20.8	17.68	19.37	0.9857	19.10	-1.69	-1.42	19.37	20.34	20.44	0.9947	19.27	-1.59
8	3	31.2	27.74	24.97	0.9857	24.62	2.77	3.12	24.97	20.34	20.44	0.9947	24.84	2.90
8	4	22.1	16.46	20.24	0.9857	19.95	-3.78	-3.49	20.24	20.34	20.44	0.9947	20.13	-3.67
8	5	15.5	15.24	15.22	0.9857	15.00	0.02	0.24	15.22	20.34	20.44	0.9947	15.14	0.10
8	6	24.4	24.99	21.65	0.9857	21.34	3.34	3.65	21.65	20.34	20.44	0.9947	21.53	3.46
8	7	20.3	14.33	19.02	0.9857	18.75	-4.69	-4.42	19.02	20.34	20.44	0.9947	18.92	-4.59
Grand mean			20.65	20.95			-0.30	0	20.95				20.65	0

Note: DBH is tree diameter (cm), H is tree height (m), $\hat{H}$ is the unadjusted predicted height (m), $\hat{H}_{adj}$ is the adjusted predicted height (m), e is the unadjusted prediction error (H - $\hat{H}$), e_{adj} is the adjusted prediction error (H - $\hat{H}_{adj}$), $\overline{H}$ is the mean of H, $\overline{\hat{H}}$ is the mean of $\hat{H}$, and PAR is the proportional adjustment ratio.

5.1 Population-Based Predictions from the Base Model

Using the white spruce base model [2] and the estimated provincial coefficients as the example (b_1=35.7854, b_2=4.8482 and b_3=-1.6040, from Table A4 of Appendix 1), tree heights were predicted directly from [2]. For instance, for Tree 1 in Plot 1:

$$\hat{H} = 1.30 + \frac{35.7854}{1 + \exp[4.8482 + (-1.6040) \cdot \ln(32.5)]} = 25.49$$

$$e = H - \hat{H} = 24.70 - 25.49 = -0.79$$

Predicted heights and prediction errors for all eight plots are listed in Table 6 (5th column and 8th column). The average of the predicted heights is $\overline{\hat{H}} = 20.95$ (last row, Table 6). The average of the observed heights is $\overline{H} = 20.65.$ The mean bias of the predictions is $\overline{e} = -0.30$ (last row, Table 6).

The above height predictions involve the direct use of the fitted base model [2]. They are unadjusted. To derive adjusted predictions, the PAR value for the population is first calculated using equation [9], which gives:

$$\text{PAR} = \frac{\overline{H}}{\overline{\hat{H}}} = \frac{20.65}{20.95} = 0.9857$$

Consequently, using the first observation in Table 6 as an example:

$$\hat{H}_{adj} = \text{PAR} \times \hat{H} = 0.9857 \times 25.49 = 25.13$$

$$e_{adj} = 24.70 - 25.13 = -0.43$$

Adjusted predictions for all other observations were calculated in a similar manner. They are listed in Table 6.

The goodness-of-fit statistics associated with the unadjusted and adjusted population-based predictions are listed in Table 7. As expected, the mean bias ($\overline{e}$) of the adjusted predictions from the proportional adjustment method is guaranteed to be zero. But the differences between the unadjusted and adjusted predictions are small (e.g., δ=15.4934 for B-U and δ=15.2342 for B-A). Both types of predictions appear to be reasonable on average.

Table 7. Goodness-of-fit statistics for population-based predictions.

Type	Goodness-of-fit measure										
	$\overline{e}$	$\overline{e}$%	SD	MAD	MSE	R^2	CC	MPE	MAPE	e_{10}	δ
B-U	-0.2975	-1.4405	3.9249	3.0537	14.9432	0.2919	0.6197	-3.80	16.04	53.57	15.4934
B-A	0	0	3.9031	3.0423	14.6901	0.3039	0.6203	-2.32	15.90	57.14	15.2342

Note: Type refers to prediction type. "B-U" denotes base model, unadjusted. "B-A" denotes base model, adjusted from the proportional adjustment method. The base model is [2] (with provincial coefficients). The data used in computations are listed in Table 6 (under "population-based"). The goodness-of-fit measures are defined in Table 5.

The unadjusted and adjusted population-based predictions are also shown in Figure 8. It is evident that their difference is small. Apparently for this population, the base model could be used to make population-based predictions with or without adjustment.

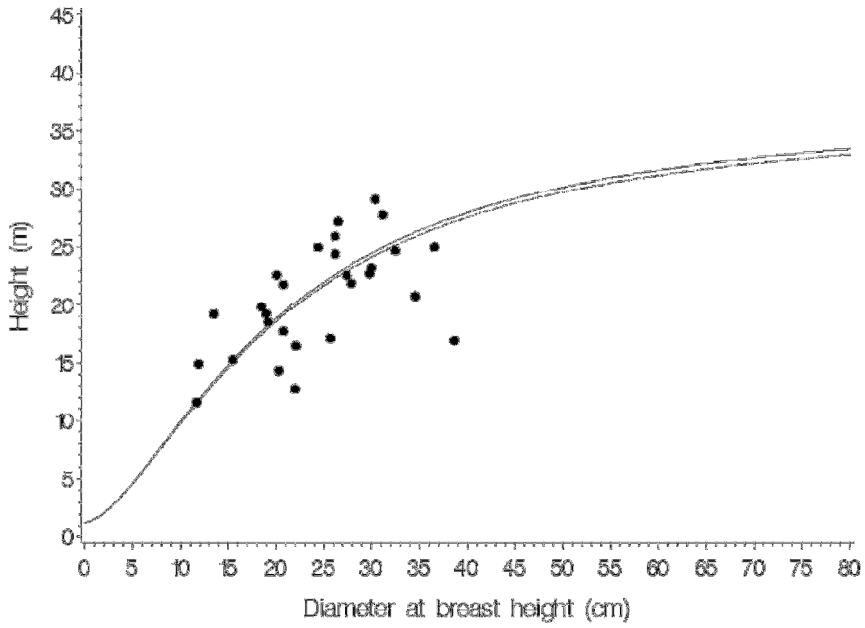

Figure 8. White spruce height-diameter data from an example population of eight plots. The solid line is the unadjusted population prediction obtained directly from the base model [2]. The dashed line is the adjusted population prediction obtained from the proportional adjustment method. Table 6 lists the actual data and computations (under "population-based").

5.2 Plot-Specific Predictions from the Base Model

Calculation for plot-specific predictions from the base model [2] is done by plot. Using plot 1 in Table 6 as an example, the averages of the observed and predicted heights for this plot are:

$\overline{H}$ = 21.53 (21.5333) $\quad$ $\overline{\hat{H}}$ = 21.42 (21.4164)

Therefore, the proportional adjustment ratio for plot 1 is:

$$PAF = \frac{\overline{H}}{\overline{\hat{H}}} = \frac{21.5333}{21.4164} = 1.0055$$

Hence, the adjusted predictions for plot 1 are 1.0055 times the unadjusted predictions obtained directly from the base model [2]. For instance, for the first observation in Table 6, the adjusted prediction and prediction error are:

$$\hat{H}_{adj} = PAR \times \hat{H} = 1.0055 \times 25.49 = 25.63$$

$$e_{adj} = 24.70 - 25.63 = -0.93$$

Adjusted predictions and prediction errors for other observations are calculated in a similar manner. Results of the plot-specific calculations are listed in Table 6 (under the heading of "plot-specific"). Figure 9 shows the unadjusted and adjusted predictions across the DBH range for all eight plots. Note that the unadjusted predictions are identical for each plot. They are all obtained directly from the base model [2].

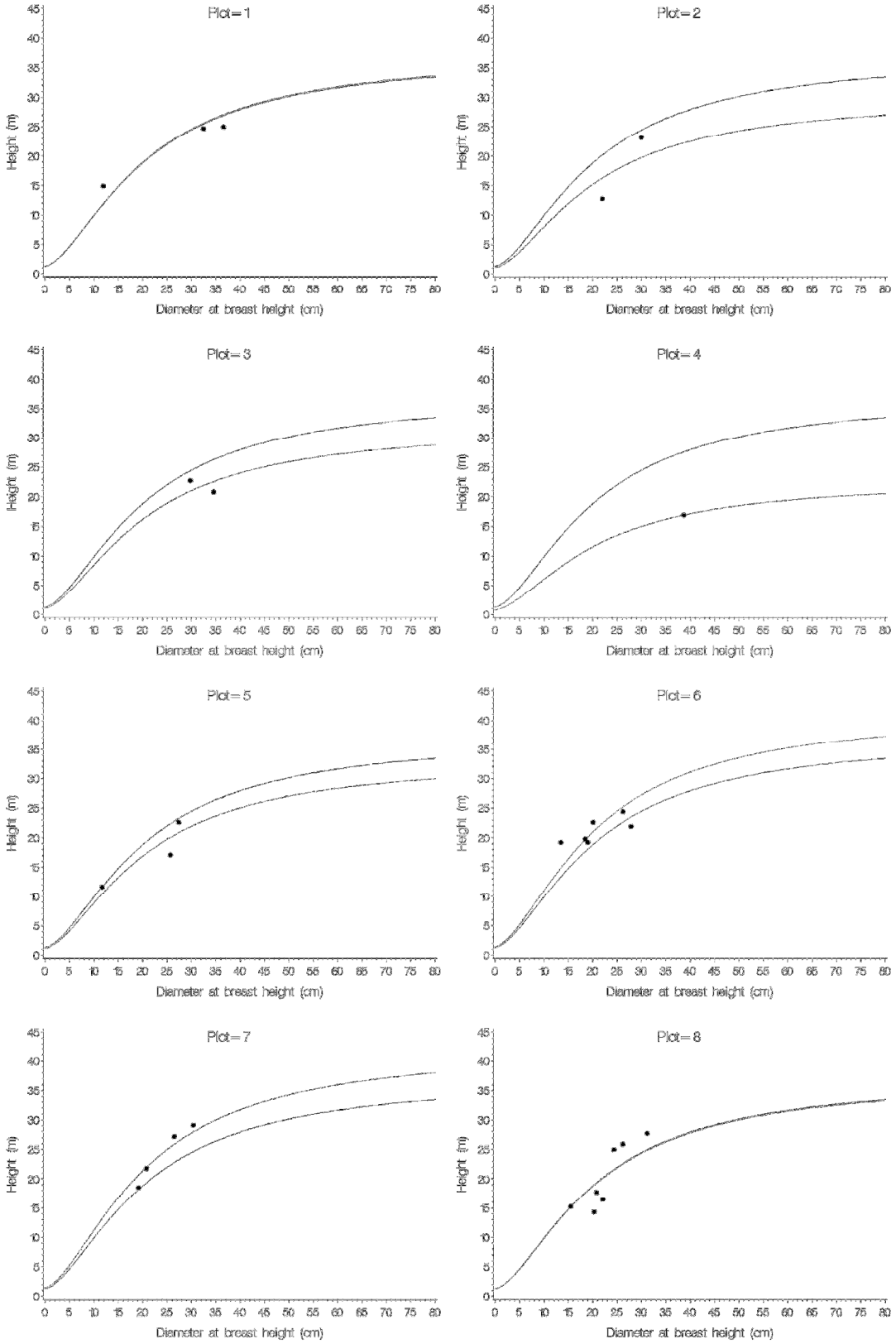

Figure 9. Unadjusted (solid lines) and adjusted (dashed lines) plot-specific predictions from the white spruce base model [2]. The unadjusted predictions are identical for each plot. They are obtained directly from the base model. Table 6 lists the actual data and computations (under "plot-specific").

Table 8 lists the plot-specific goodness-of-fit statistics for the unadjusted and adjusted predictions. For plot 4 with one observation only, the SD and R^2 are not applicable because the denominator is zero.

Table 8. Goodness-of-fit statistics for plot-specific predictions from the base model [2].

Type	Goodness-of-fit measure											
	Plot	$\bar{e}$	$\bar{e}$%	SD	MAD	MSE	R^2	CC	MPE	MAPE	e_{10}	δ
B-U	1	0.1169	0.5431	2.6267	1.9348	4.6133	0.7905	0.9326	3.24	10.53	33.33	6.9131
	2	-4.3296	-24.0668	4.3624	4.3296	28.2609	-0.0332	0.4426	-31.73	31.73	50.00	37.7762
	3	-3.5765	-16.4472	2.7143	3.5765	16.4747	-16.3304	-0.1258	-16.88	16.88	50.00	20.1584
	4	-10.6718	-63.1465	N/A	10.6718	113.8863	N/A	0	-63.15	63.15	100.00	113.8863
	5	-1.9822	-11.5919	2.8664	1.9822	9.4066	0.5336	0.8193	-11.30	11.30	33.33	12.1453
	6	2.1554	10.1752	2.4959	2.6883	9.8373	-1.6061	0.5002	10.53	12.96	50.00	10.8755
	7	2.8932	11.9803	1.9834	2.8932	11.3214	0.3644	0.6551	11.04	11.04	75.00	12.3049
	8	-0.1081	-0.5315	3.3814	2.7950	9.8120	0.6416	0.7242	-4.16	14.48	57.14	11.4454
B-A	1	0	0	2.6718	2.0087	4.7591	0.7838	0.9309	2.71	10.77	33.33	7.1386
	2	0	0	4.9509	3.5008	12.2559	0.5519	0.5961	-6.18	21.26	100.00	24.5118
	3	0	0	2.5257	1.7859	3.1895	-2.3552	-0.9833	-0.37	8.23	0	6.3791
	4	0	0	N/A	0	0	N/A	0	0	0	0	0
	5	0	0	2.5692	1.9653	4.4006	0.7818	0.8967	0.26	11.75	66.67	6.6009
	6	0	0	2.8507	1.9947	6.7722	-0.7941	0.6181	0.40	9.58	33.33	8.1267
	7	0	0	1.6144	1.2050	1.9547	0.8903	0.9259	-1.07	5.36	25.00	2.6062
	8	0	0	3.3900	2.8145	9.8502	0.6403	0.7224	-3.61	14.48	71.43	11.4920
B-U	All	-0.2975	-1.4405	3.9249	3.0537	14.9432	0.2919	0.6197	-3.80	16.04	53.57	15.4934
B-A	All	0	0	2.5515	2.1066	6.2776	0.7025	0.8407	-1.12	10.96	46.43	6.5101

Note: Type refers to prediction type. "B-U" denotes base model, unadjusted. "B-A" denotes base model, adjusted from the proportional adjustment method. "N/A" denotes "not applicable" (due to a zero denominator). Actual data used in computations are listed in Table 6 (under "plot-specific"). The goodness-of-fit measures are defined in Table 5.

The plot-specific goodness-of-fit statistics listed in Table 8 suggest that, for the unadjusted predictions (B-U), the absolute values of the percent mean bias $\bar{e}$% exceeded 10% for six of the eight plots. This is an indication that the model if unadjusted would produce large biases for most plots in the population, even though it was shown earlier (in Table 7 and Figure 8) that the model appeared to have predicted well on average for the population. The differences between the unadjusted and adjusted predictions are obvious in Figure 9.

The plot-specific goodness-of-fit statistics (Table 8) indicate that:

1. The differences between the unadjusted and adjusted predictions can be large for many plots. For example, for plot 3, δ=20.1584 for unadjusted predictions, while δ=6.3791 for adjusted predictions.

2. The adjusted predictions are not guaranteed to be better (i.e., more accurate) than the unadjusted predictions, and a zero mean bias does not imply better predictions. For instance, for plot 1, $\bar{e}$ = 0.1169 and δ = 6.9131 for unadjusted predictions, while $\bar{e}$ = 0 and δ = 7.1386 for adjusted predictions. The unadjusted predictions are more accurate in terms of the overall δ value.

Table 8 also lists the summary goodness-of-fit statistics for all eight plots combined (bottom two rows, with the plot number designated as "All"). They were calculated using the formulas given in Table 5, based on all 28 observations listed in Table 6. To illustrate the calculation of the summary statistics, for instance, for the unadjusted predictions obtained directly from the base model [2], the grand means of the observed heights, predicted heights and prediction errors from all *N*=28 observations are obtained as follows:

$$\bar{y} = \frac{1}{N}\sum_{i=1}^{m}\sum_{j=1}^{n_i} y_{ij} = 20.6543 \qquad \bar{\hat{y}} = \frac{1}{N}\sum_{i=1}^{m}\sum_{j=1}^{n_i} \hat{y}_{ij} = 20.9518 \qquad \bar{e} = \frac{1}{N}\sum_{i=1}^{m}\sum_{j=1}^{n_i} (y_{ij} - \hat{y}_{ij}) = -0.2975$$

where y represents the dependent variable H (so the expressions are more generic), m=8, i=1, 2, …, 8, n_1=3, n_2=2, n_3=2, n_4=1, n_5=3, n_6=6, n_7=4 and n_8=7 (the calculated results listed in Table 6 only include two decimal places). Hence, e.g., the percent mean bias, standard deviation and R^2 for the entire population are:

$$\bar{e}\% = (\bar{e}/\bar{y}) \times 100 = (-0.2975/20.6543) \times 100 = -1.4405\%$$

$$SD = \sqrt{\frac{1}{28-1}\sum_{i=1}^{8}\sum_{j=1}^{n_i}[e_{ij} - (-0.2975)]^2} = 3.9249$$

$$R^2 = 1 - \sum_{i=1}^{8}\sum_{j=1}^{n_i}(y_{ij} - \hat{y}_{ij})^2 \Big/ \sum_{i=1}^{8}\sum_{j=1}^{n_i}(y_{ij} - 20.6543)^2 = 0.2919$$

Other summary goodness-of-fit statistics can be calculated in a similar manner. They are listed in Table 8 identified by the prediction type of "B-U" and Plot=All, where "All" refers to all plots combined. Note that the summary goodness-of-fit statistics for unadjusted predictions from all plots combined (Type=B-U and Plot=All in Table 8) are identical to those for unadjusted population-based predictions (Type=B-U in Table 7). This is because the two types of predictions have the same unadjusted height predictions ($\hat{H}$'s in Table 6) obtained directly from the base model [2].

For the adjusted plot-specific predictions, calculations for the summary goodness-of-fit statistics are similar except that the adjusted predictions and errors ($\hat{H}_{adj}$ and e_{adj} in Table 6) are used in place of the unadjusted predictions and errors. For instance:

$$\bar{\hat{y}}_{adj} = \frac{1}{28}\sum_{i=1}^{8}\sum_{j=1}^{n_i}\hat{y}_{ij\_adj} = 20.6543$$

$$\bar{e} = \frac{1}{28}\sum_{i=1}^{8}\sum_{j=1}^{n_i} e_{ij\_adj} = 0 \qquad \bar{e}\% = 0$$

$$SD = \sqrt{\frac{1}{28-1}\sum_{i=1}^{8}\sum_{j=1}^{n_i}(e_{ij\_adj} - 0)^2} = 2.5515$$

$$R^2 = 1 - \sum_{i=1}^{8}\sum_{j=1}^{n_i}(y_{ij} - \hat{y}_{ij\_adj})^2 \Big/ \sum_{i=1}^{8}\sum_{j=1}^{n_i}(y_{ij} - 20.6543)^2 = 0.7025$$

Other summary goodness-of-fit statistics can be calculated in a similar manner. Results are listed in the bottom row in Table 8 (Type=B-A, Plot=All).

The summary goodness-of-fit statistics listed in Table 8 suggest that when all plots are combined, the adjusted predictions (δ=6.5101) are more than twice as accurate as the unadjusted predictions (δ=15.4934). A more detailed look of the plot-by-plot goodness-of-fit statistics reveals that this is mainly caused by the substantial accuracy gains for plots 2, 3, 4, 5, 6 and 7 after adjustment. Even for plots 1 and 8 where the accuracy is reduced after adjustment, the differences between unadjusted (δ=6.9131 for plot 1 and δ=11.4454 for plot 8) and adjusted (δ=7.1386 for plot 1 and δ=11.4920 for plot 8) predictions are small. These can also be seen in Figure 9.

5.3 Plot-Specific Predictions from the Mixed Model

The same white spruce height-diameter data from the small population of eight plots (Table 9) were used to demonstrate the application of the fitted mixed model. Step-by-step computations relevant to the variables listed in Table 9 are described next.

Table 9. Original data and height predictions from the white spruce mixed model [4].

Plot	Tree	DBH	H	$\hat{H}_{fix}$	der_u_1	der_u_2	$\hat{H}$	e	$\overline{H}$	$\overline{\hat{H}}$	PAR	$\hat{H}_{adj}$	e_{adj}
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
1	1	32.5	24.70	24.20	0.6416	-8.2069	24.64	0.07	21.53	21.17	1.0170	25.05	-0.35
1	2	11.9	14.90	11.79	0.2939	-7.4060	13.14	1.76	21.53	21.17	1.0170	13.36	1.54
1	3	36.6	25.00	25.58	0.6803	-7.7629	25.74	-0.74	21.53	21.17	1.0170	26.18	-1.18
2	1	30.0	23.22	23.23	0.6145	-8.4550	20.82	2.40	17.99	18.70	0.9621	20.04	3.18
2	2	22.0	12.76	19.29	0.5039	-8.9223	16.57	-3.81	17.99	18.70	0.9621	15.94	-3.18
3	1	29.8	22.72	23.15	0.6122	-8.4737	21.25	1.47	21.75	22.03	0.9870	20.98	1.74
3	2	34.6	20.77	24.94	0.6623	-7.9831	22.81	-2.04	21.75	22.03	0.9870	22.51	-1.74
4	1	38.7	16.90	26.20	0.6976	-7.5288	18.97	-2.07	16.90	18.97	0.8910	16.90	0.00
5	1	25.7	17.10	21.29	0.5601	-8.7940	19.97	-2.87	17.10	17.21	0.9938	19.85	-2.75
5	2	11.7	11.60	11.61	0.2888	-7.3302	10.91	0.70	17.10	17.21	0.9938	10.84	0.76
5	3	27.4	22.60	22.10	0.5829	-8.6777	20.74	1.86	17.10	17.21	0.9938	20.61	1.99
6	1	13.5	19.20	13.19	0.3332	-7.9303	15.99	3.21	21.18	20.85	1.0161	16.25	2.95
6	2	19.0	19.20	17.39	0.4509	-8.8366	20.18	-0.98	21.18	20.85	1.0161	20.50	-1.30
6	3	26.2	24.40	21.54	0.5670	-8.7627	23.79	0.61	21.18	20.85	1.0161	24.17	0.23
6	4	20.1	22.60	18.12	0.4712	-8.8931	20.85	1.75	21.18	20.85	1.0161	21.18	1.42
6	5	18.5	19.80	17.05	0.4413	-8.7998	19.86	-0.06	21.18	20.85	1.0161	20.18	-0.38
6	6	27.9	21.90	22.33	0.5892	-8.6386	24.42	-2.52	21.18	20.85	1.0161	24.82	-2.92
7	1	30.4	29.10	23.39	0.6190	-8.4171	27.66	1.44	24.15	23.85	1.0127	28.01	1.09
7	2	19.2	18.50	17.53	0.4546	-8.8493	20.46	-1.96	24.15	23.85	1.0127	20.72	-2.22
7	3	20.8	21.80	18.56	0.4836	-8.9132	21.72	0.08	24.15	23.85	1.0127	21.99	-0.19
7	4	26.5	27.20	21.68	0.5710	-8.7428	25.54	1.66	24.15	23.85	1.0127	25.87	1.33
8	1	26.2	25.91	21.54	0.5670	-8.7627	22.84	3.07	20.34	20.36	0.9988	22.81	3.10
8	2	20.8	17.68	18.56	0.4836	-8.9132	19.03	-1.35	20.34	20.36	0.9988	19.01	-1.33
8	3	31.2	27.74	23.71	0.6279	-8.3391	25.76	1.98	20.34	20.36	0.9988	25.73	2.01
8	4	22.1	16.46	19.34	0.5056	-8.9217	20.02	-3.56	20.34	20.36	0.9988	19.99	-3.53
8	5	15.5	15.24	14.83	0.3792	-8.4021	14.58	0.66	20.34	20.36	0.9988	14.56	0.68
8	6	24.4	24.99	20.62	0.5414	-8.8615	21.65	3.34	20.34	20.36	0.9988	21.62	3.37
8	7	20.3	14.33	18.24	0.4748	-8.9001	18.64	-4.31	20.34	20.36	0.9988	18.62	-4.29

Random parameters	Plot							
	1	2	3	4	5	6	7	8
u_1	-3.3692	1.2083	-4.0190	-11.0924	-2.2174	-4.2304	7.9135	9.1461
u_2	-0.3163	0.3725	-0.0667	-0.0672	0.0083	-0.5310	0.0749	0.4431

Note: DBH (cm) and H (m) are observed tree diameter and height. All other variables are described in the main text.

5.3.1 Prediction of Random Parameters

To make a plot-specific prediction from the mixed model [4], the random parameters unique for each plot must be predicted first. Here, plot 1 is used as an example to demonstrate the computation, based on the estimated provincial coefficients listed in Table A4, Appendix 1 for white spruce:

$$b_1=35.6912,\ b_2=4.4737,\ b_3=-1.4524,\ \sigma^2_{u_1}=40.7210,\ \sigma^2_{u_2}=0.1348,\ \sigma_{u_1u_2}=1.5971,\ \sigma^2=3.0532.$$

where b_1, b_2 and b_3 are fixed parameters applicable to every plot in the population, $\sigma^2_{u_1}$, $\sigma^2_{u_2}$ and $\sigma_{u_1u_2}$ are variances and covariance for the two random parameters (u_1 and u_2) in model [4], and σ^2 is the residual variance.

The height predictions based on the fixed parameters only are listed in the 5th column of Table 9. They are obtained directly from the mixed model [4], with the random parameters set to zero:

$$\hat{\mathbf{H}}_{i\_fix} = 1.30 + \frac{b_1}{1+\exp[b_2 + b_3 \ln(\mathbf{DBH}_i)]}$$

The partial derivatives of [4] with respect to the two random parameters are:

$$der\_u_1 = der\_b_1 = \partial f(\mathbf{DBH}_i, \hat{\mathbf{b}}, \mathbf{0})/\partial b_1 = 1/(1+\exp(b_2 + b_3 \ln(\mathbf{DBH}_i)))$$

$$der\_u_2 = der\_b_2 = \partial f(\mathbf{DBH}_i, \hat{\mathbf{b}}, \mathbf{0})/\partial b_2 = -b_1 \exp(b_2 + b_3 \ln(\mathbf{DBH}_i))/(1+\exp(b_2 + b_3 \ln(\mathbf{DBH}_i)))^2$$

The derivatives (columns 6-7 of Table 9) constitute a design matrix, called the $\mathbf{Z}_i$ matrix, for the mixed model:

$$\mathbf{Z}_i = \begin{bmatrix} d\_u_{11} & d\_u_{21} \\ d\_u_{12} & d\_u_{22} \\ d\_u_{13} & d\_u_{23} \end{bmatrix} = \begin{bmatrix} 0.6416 & -8.2069 \\ 0.2939 & -7.4060 \\ 0.6803 & -7.7629 \end{bmatrix}$$

Given $\sigma^2_{u_1}$ = 40.7210, $\sigma_{u_1u_2}$ = 1.5971, $\sigma^2_{u_2}$ = 0.1348 and σ^2 = 3.0532, for plot 1 with three observations and two random parameters, the variance-covariance matrices $\hat{\mathbf{R}}_i$ and $\hat{\mathbf{D}}$ for the errors and random parameters are:

$$\hat{\mathbf{R}}_i = \begin{bmatrix} \sigma^2 & 0 & 0 \\ 0 & \sigma^2 & 0 \\ 0 & 0 & \sigma^2 \end{bmatrix} = \begin{bmatrix} 3.0532 & 0 & 0 \\ 0 & 3.0532 & 0 \\ 0 & 0 & 3.0532 \end{bmatrix}$$

$$\hat{\mathbf{D}} = \begin{bmatrix} \sigma^2_{u_1} & \sigma_{u_1u_2} \\ \sigma_{u_1u_2} & \sigma^2_{u_2} \end{bmatrix} = \begin{bmatrix} 40.7210 & 1.5971 \\ 1.5971 & 0.1348 \end{bmatrix}$$

Therefore, the two random parameters for plot 1 can be predicted using the standard random parameter prediction equation listed in Table 4 for the FO method ($\mathbf{H}_i = \mathbf{y}_i$, $\hat{\mathbf{H}}_{i\_fix} = f(\mathbf{x}_i, \hat{\mathbf{b}}, \mathbf{0})$):

$$\hat{\mathbf{u}}_i = \hat{\mathbf{D}}\mathbf{Z}_i'(\mathbf{Z}_i\hat{\mathbf{D}}\mathbf{Z}_i' + \hat{\mathbf{R}}_i)^{-1}(\mathbf{H}_i - \hat{\mathbf{H}}_{i\_fix})$$

which produces the following random parameter predictions for plot 1:

$$\hat{\mathbf{u}}_i = [u_1, u_2]' = [-3.3692, -0.3163]'$$

Random parameter predictions for other plots were obtained in a similar manner. Results are listed at the bottom of Table 9. Intermediate plot-by-plot computations are shown in the main body of Table 9.

5.3.2 Prediction of Tree Height

Once the $\hat{\mathbf{u}}_i$ for plot *i* are available (*i* = 1, 2, …, 8), plot-specific height predictions and associated prediction errors are calculated by (from Table 4):

[21] $\hat{\mathbf{H}}_i = \hat{\mathbf{H}}_{i\_fix} + \mathbf{Z}_i\hat{\mathbf{u}}_i$

[22] $\mathbf{e}_i = \mathbf{H}_i - (\hat{\mathbf{H}}_{i\_fix} + \mathbf{Z}_i\hat{\mathbf{u}}_i)$

Hence, for plot 1:

$$\hat{\mathbf{H}}_i = \hat{\mathbf{H}}_{i\_fix} + \mathbf{Z}_i\hat{\mathbf{u}}_i = \begin{bmatrix}24.20\\11.79\\25.58\end{bmatrix} + \begin{bmatrix}0.6416 & -8.2069\\0.2939 & -7.4060\\0.6803 & -7.7629\end{bmatrix} \times \begin{bmatrix}-3.3692\\-0.3163\end{bmatrix} = \begin{bmatrix}24.6350\\13.1405\\25.7434\end{bmatrix}$$

$$\mathbf{e}_i = \mathbf{H}_i - (\hat{\mathbf{H}}_{i\_fix} + \mathbf{Z}_i\hat{\mathbf{u}}_i) = \begin{bmatrix}24.70\\14.90\\25.00\end{bmatrix} - \begin{bmatrix}24.6350\\13.1405\\25.7434\end{bmatrix} = \begin{bmatrix}0.06504\\1.7595\\-0.7434\end{bmatrix}$$

Table 9 (columns 8 and 9) lists the predicted heights and prediction errors (to two decimal places) for all eight plots in the population.

Height predictions obtained from equation [21] represent the standard application of a mixed model fitted by the FO method. Some better predictions may be obtained if we could adjust the model through the proportional adjustment method. To use the proportional adjustment method for mixed model [4], a plot-specific proportional adjustment ratio (PAR_i) between the averages of the observed heights and predicted heights is calculated:

$$PAR_i = \overline{\mathbf{H}}_i / \overline{\hat{\mathbf{H}}}_i$$

where $\overline{\mathbf{H}}_i$ (column 10, Table 9) is the average of the observed heights for plot *i*, and $\overline{\hat{\mathbf{H}}}_i$ (column 11, Table 9) is the average of the unadjusted predicted heights for plot *i* from the FO method (equation [21]).

For plot 1, PAR_i = 21.53/21.17 = 1.0170 (column 12, Table 9). This plot-specific PAR_i can then be used to calculate adjusted height predictions and associated prediction errors for plot *i*:

[23] $\hat{\mathbf{H}}_{i\_adj} = PAR_i \cdot \hat{\mathbf{H}}_i$

[24] $\mathbf{e}_{i\_adj} = \mathbf{H}_i - \hat{\mathbf{H}}_{i\_adj}$

Results of the calculations for all eight plots are listed in columns 13 and 14 of Table 9.

For a graphic comparison, the unadjusted and adjusted predictions from the mixed model [4] are shown in Figure 10. Notice the closeness between the unadjusted and adjusted predictions across the DBH range for all plots except for plot 4, which has only one observation.

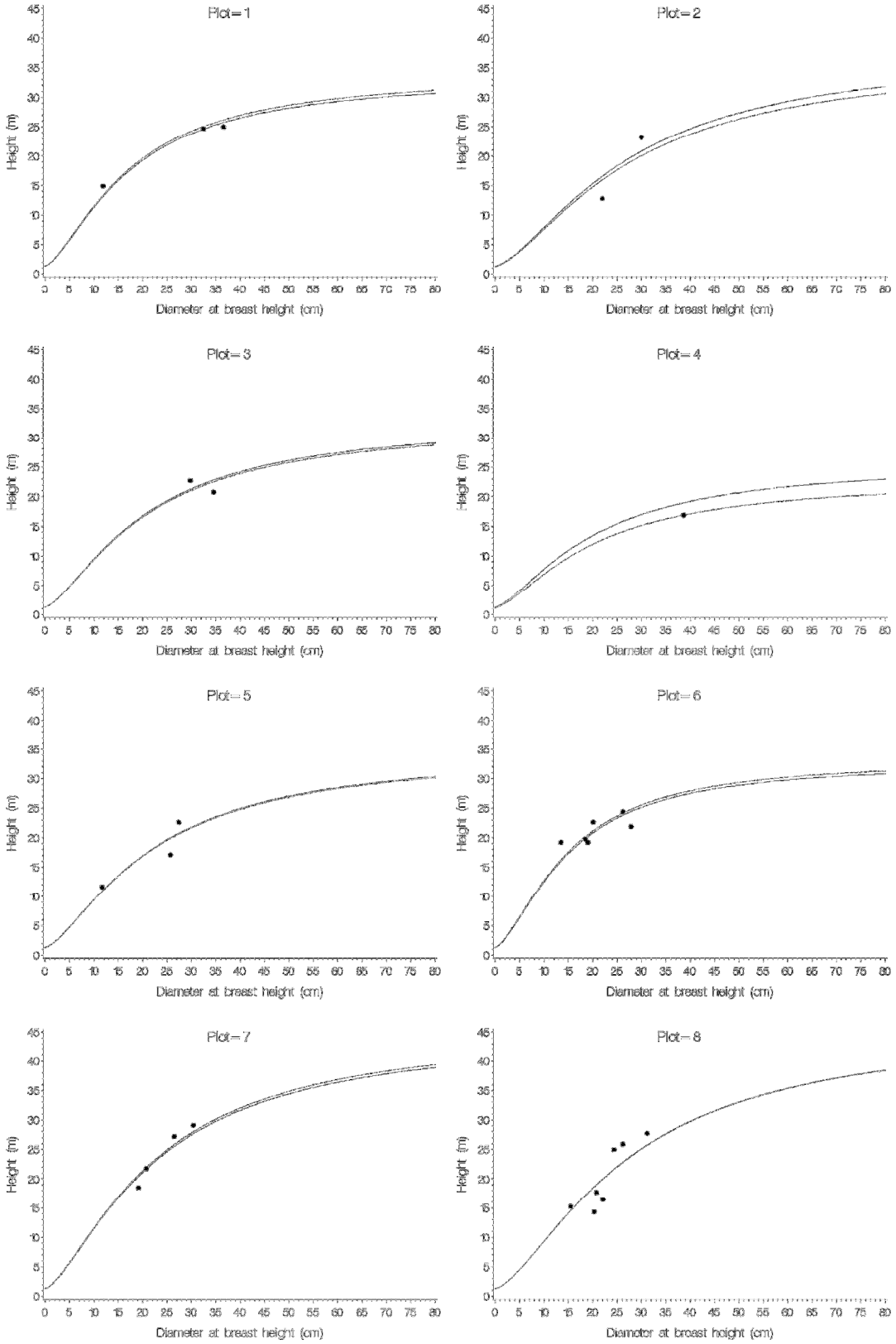

Figure 10. Unadjusted (solid lines) and adjusted (dashed lines) plot-specific predictions from the white spruce mixed model [4] with provincial coefficients. The original data and calculations are listed in Table 9. Plot-specific goodness-of-fit statistics are listed in Table 10.

The goodness-of-fit statistics associated with the unadjusted and adjusted predictions from the white spruce mixed model [4] are listed in Table 10 for each of the eight plots in the population. Summary goodness-of-fit statistics for all eight plots combined are also listed in Table 10 (bottom two rows).

Table 10. Goodness-of-fit statistics from the white spruce mixed model [4] on the example data.

Type	Plot	Goodness-of-fit measure										
		$\bar{e}$	$\bar{e}\%$	SD	MAD	MSE	R^2	CC	MPE	MAPE	e_{10}	δ
M-U	1	0.3604	1.6736	1.2773	0.8560	1.2176	0.9447	0.9777	3.03	5.02	33.33	1.7614
	2	-0.7081	-3.9361	4.3888	3.1033	10.1321	0.6296	0.6871	-9.78	20.09	100.00	19.7629
	3	-0.2860	-1.3153	2.4774	1.7518	3.1505	-2.3141	-0.9260	-1.68	8.13	0	6.2191
	4	-2.0671	-12.2313	N/A	2.0671	4.2728	N/A	0	-12.23	12.23	100.00	4.2728
	5	-0.1061	-0.6203	2.4678	1.8105	4.0713	0.7981	0.8985	-0.86	10.35	33.33	6.1013
	6	0.3356	1.5843	2.0188	1.5211	3.5091	0.0704	0.6982	1.68	7.31	33.33	4.1883
	7	0.3038	1.2580	1.6626	1.2842	2.1654	0.8784	0.9175	0.20	5.50	25.00	2.8564
	8	-0.0238	-0.1172	3.1098	2.6108	8.2900	0.6972	0.7825	-3.24	13.72	57.14	9.6716
M-A	1	0	0	1.3929	1.0239	1.2935	0.9412	0.9767	1.38	5.49	33.33	1.9402
	2	0	0	4.5027	3.1839	10.1371	0.6294	0.6786	-5.62	19.33	100.00	20.2742
	3	0	0	2.4631	1.7417	3.0334	-2.1910	-0.9718	-0.36	8.03	0	6.0669
	4	0	0	N/A	0	0	N/A	0	0	0	0	0
	5	0	0	2.4608	1.8345	4.0371	0.7998	0.8987	-0.24	10.49	33.33	6.0557
	6	0	0	2.0543	1.5314	3.5169	0.0683	0.7011	0.09	7.43	33.33	4.2203
	7	0	0	1.6250	1.2081	1.9804	0.8888	0.9249	-1.07	5.38	25.00	2.6405
	8	0	0	3.1117	2.6153	8.2992	0.6969	0.7822	-3.11	13.72	57.14	9.6824
M-U	All	-0.0082	-0.0398	2.2315	1.8684	4.8018	0.7725	0.8663	-1.44	9.88	42.86	4.9797
M-A	All	0	0	2.1896	1.8126	4.6232	0.7809	0.8755	-1.22	9.46	39.29	4.7945

Note: Type refers to prediction type. "M-U" denotes mixed model, unadjusted. "M-A" denotes mixed model, adjusted. "All" denotes all plots combined. "N/A" denotes "not applicable" (due to a zero denominator). The goodness-of-fit measures are defined in Table 5.

The results shown in Table 10 suggest that, in terms of the overall accuracy measure δ, the unadjusted predictions are more accurate for plots 1, 2, 6 and 8, whereas the adjusted predictions are more accurate for plots 3, 4, 5 and 7. When all plots are combined, the adjusted predictions (δ=4.7945) are slightly more accurate than the unadjusted predictions (δ=4.9797). Of course, the mean bias ($\bar{e}$) and the percent mean bias ($\bar{e}\%$) for adjusted predictions are guaranteed to be zeros.

Readers can make different types of comparisons based on the results listed in Tables 8 and 10. For instance, the summary statistics from all eight plots combined suggest that the overall accuracy (δ) has a ranking of:

M-A (most accurate, δ=4.7945) → M-U (δ=4.9797) → B-A (δ=6.5101) → B-U (least accurate, δ=15.4934).

5.3.3 Computation Program for Mixed Model Predictions

An example program for mixed model calculations is provided in Table 11. The program produced the results in Tables 9 and 10. It requires SAS BASIC, SAS STATS and SAS IML (interactive matrix language) to run. Interested readers who may have difficulties in accessing the SAS modules could use other programs or carry out the calculations manually, as demonstrated above. All calculations are algebraic (i.e., they are non-iterative). The only difficulty for some readers may be the derivation of the partial derivatives of a mixed model with respect to random parameters. Those who wish to seek possible help on derivatives may wish to contact the lead author, who will provide the partial derivatives for any model if such derivatives exist.

Table 11. An example program for mixed model calculations from the FO method.

NOTE: The example data for white spruce are identical to those listed in Table 9 (plotid = plot);

```
1   OPTIONS LS=118 PS=60;
2   data examp1;
3   input Plotid Tree DBH HT;
4   cards;
            1           1           32.5        24.70
            1           2           11.9        14.90
            1           3           36.6        25.00
            2           1           30.0        23.22
            2           2           22.0        12.76
            3           1           29.8        22.72
            3           2           34.6        20.77
            4           1           38.7        16.90
            5           1           25.7        17.10
            5           2           11.7        11.60
            5           3           27.4        22.60
            6           1           13.5        19.20
            6           2           19.0        19.20
            6           3           26.2        24.40
            6           4           20.1        22.60
            6           5           18.5        19.80
            6           6           27.9        21.90
            7           1           30.4        29.10
            7           2           19.2        18.50
            7           3           20.8        21.80
            7           4           26.5        27.20
            8           1           26.2        25.91
            8           2           20.8        17.68
            8           3           31.2        27.74
            8           4           22.1        16.46
            8           5           15.5        15.24
            8           6           24.4        24.99
            8           7           20.3        14.33
33  ;
34  run;
35
36  proc sort data=examp1; by plotid; run;
37
38  data examp2; set examp1; by plotid;
39    j+1; if first.plotid then do; i+1; j=1; end;
40  run;
```

NOTE: Fix parameters (b1, b2, b3) for the provincial white spruce model [4] are listed in Table A4, Appendix 1;
NOTE: z1 and z2 are partial derivatives with respect to random parameter u1 and u2, respectively;
NOTE: b1i and b2i are predicted random parameters;

```
41
42  data examp3; set examp2; by plotid  ;
43    b1=35.6912; b2= 4.4737; b3=-1.4524;
44    z1 = 1 / (1 + EXP(b2 + b3 * LOG(DBH)));
45    z2 = -  b1*EXP(b2 + b3 * LOG(DBH))  / (1 + EXP(b2 + b3 * LOG(DBH)))**2;
46    zu1=z1; zu2=z2;
47    ht_fix=1.30+ b1/(1+exp(b2 + b3*log(dbh+0) ));
48    res_fix=ht-ht_fix;
49  run;
50
51  filename random 'c:\_localdata\random.txt' ;
52  proc iml;
53    file random;
54    use examp3;
55    do k=1 to 8;
56      read all var {z1 z2} into Z where (i=k);
57      read all var {res_fix} into RES where (i=k);
58      read all var {j} into MM where (i=k);
59    ss=nrow(mm);
```

```
60    R=3.0532 * I(ss);
61    D= {40.7210 1.5971, 1.5971 0.1348};
62    b=D*Z`*  INV(Z * D * Z` + R)*RES;
63      bTrans = b`;
64      b1i= bTrans[1,1];
65      b2i= bTrans[1,2] ;
66      put k 5. +2 b1i 10.6 +2 b2i 10.6 +2;
67  end;
68  closefile random ;
69  quit ;
70
71  data prandom; infile random; input i b1i b2i; run ;
72
73  data all; merge examp3 Prandom; by i; run;
74
75  proc sort data=all; by plotid; run ;
NOTE: ht_pred is unadjusted height prediction, and ht_adj is adjusted height prediction;
77  data allpx; set all;
78    ht_pred = ht_fix+b1i*zu1+b2i*zu2;
79  run;
80
81  proc means data=allpx noprint;
82    var ht ht_pred;
83    output out=resu1 mean=m_ht_ss m_ht_pred_ss;
84    by plotid;
85  run;
86
87  data examp4; merge allpx resu1; by plotid; drop _TYPE_ _FREQ_; run;
88
89  data examp5; set examp4;
90    PAR=m_ht_ss/m_ht_pred_ss;
91    ht_adj=PAR*ht_pred;
92    diff_pred=ht-ht_pred; diff_adj =ht-ht_adj;
93  run;
94
NOTE: This produces the results shown in Table 9;
95  proc print data=examp5;
96    var Plotid Tree DBH HT ht_fix z1 z2 b1i b2i ht_pred diff_pred m_ht_ss m_ht_pred_ss PAR ht_adj diff_adj;
97  run;
98
NOTE: The follow statements calculate the goodness-of-fit measures defined in Table 5;
99  data allp;
100   set examp5;
101    y_pred = ht_pred;
102   *y_pred = ht_adj;
103    Y_res=ht-y_pred;
104    e_percent=(Y_res/ht)*100;
105   if e_percent> 5 or e_percent<-5  then do; e5 =1; end;
106   if e_percent>10 or e_percent<-10 then do; e10=1; end;
107   if e_percent>20 or e_percent<-20 then do; e20=1; end;
108     e_abs=abs(ht-Y_pred); e_percent1=abs(Y_res/ht)*100;
109   ident=1;
110 run;
111
112 proc means data=allP noprint;
113   id ident;
114   var ht y_pred Y_res e_percent e_abs e5 e10 e20  e_percent1;
115   by plotid;
116   output out=result  mean=m_ht m_pred m_res  MPE MAD m_e5 m_e10 m_e20  MAPE
117                 sum =s_ht s_pred s_res  s_e_percent s_e_abs s_e5 s_e10 s_e20  s_e_percent1 ;
118 run;
119
120 data res2; merge allP result;
121   by plotid;
122 run;
123
124 data res3; set res2;
```

```
125   diff0 = (ht - y_pred); diff1 = (ht - y_pred)**2; diff2 = (ht - m_ht)**2; diff3 = (y_pred - m_pred)**2;
126   diff4 = (y_res - m_res)**2;
127  run;
128
129  proc means data=res3 noprint;
130   id m_ht m_pred MPE MAD s_e5 s_e10 s_e20  MAPE;
131   var Y_res diff0 diff1 diff2 diff3 diff4;
132   by plotid;
133   output out=res4 sum=Y_sum sum0 sum1 sum2 sum3 sum4
134                mean=Y_mean mean0 mean1 mean2 mean3 mean4
135                std=std_r std_0 ;
136  run;
137
138  data ttk4;
139   set res4;
140   Mean_Bias=mean0;
141   Percent_Bias=100*mean0/m_ht;
142   bias2=100*(m_ht -m_pred)/m_ht;
143   rmse=sqrt(sum1/(_freq_-0));
144   MSE =    (sum1/(_freq_-0));
145   if _freq_>1 then do; Std_dev =sqrt(sum4/(_freq_-1)); end;
146   if _freq_=1 then do; Std_dev = 0                ; end;
147   if sum2>0 then do; R2 = 1 - sum1/sum2; end;
148   if sum2=0 then do; R2 = 0          ; end;
149   if (sum2+sum3)> 0     then do;  CC = 1- sum1/(sum2 + sum3 + _freq_*(m_ht-m_pred)**2); end;
150   if (sum2=0 and sum3=0) then do;  CC = 0                                    ; end;
151   if s_e5>0  then do;  e5 =100*s_e5/_freq_; end;
152   if s_e10>0 then do; e10=100*s_e10/_freq_; end;
153   if s_e20>0 then do; e20=100*s_e20/_freq_; end;
154   if s_e5=.  then do;  e5=0; end; if s_e10=. then do; e10=0; end; if s_e20=. then do; e20=0; end;
155
156   Accuracy=Mean_Bias**2+std_dev**2;
157
158  proc print data=ttk4;
159   var Plotid Mean_Bias Percent_Bias Std_dev MAD MSE R2 cc MPE MAPE e10 Accuracy;
160  run;
```

The final print statement produces the plot-specific M-U results listed in Table 10.

To produce the plot-specific M-A results listed in Table 10, one simply needs to remove the "*" in line 102 of the program (i.e., change "*y_pred = ht_adj" to "y_pred = ht_adj").

To produce the M-U and "Plot=All" results listed in Table 10, one needs to:

1). Remove line 113;
2). Change "by plotid" in lines 115, 121 and 132 to "by ident";
3). Remove variable "plotid" in line 159.

To produce the M-A and "Plot=All" results listed in Table 10, one needs to:

1). Remove the "*" in line 102;
2). Remove line 113;
3). Change "by plotid" in lines 115, 121 and 132 to "by ident";
4). Remove variable "plotid" in line 159.

The example program given in Table 11 was designed to demonstrate the logic and step-by-step computations involved in using the FO method of the NMM technique. More generalized and efficient programs could be written using SAS or other programming languages, so long as the logic embedded in the example program is fully understood. Several computationally more efficient programs are available to limited users.

6. Application Examples on Larger Data Sets

The performances of the fitted models and different types of predictions were evaluated on many independent data sets not used in modelling. Due to the amount of the work involved, here, only the procedures and results on two independent white spruce data sets collected after 1980 were presented. Interested readers could apply the same procedures to other data sets collected from other populations. Once again, the procedures presented here could also serve as an integral part of model validation.

The first data set (data I) was collected in the boreal forest natural region of Alberta. It consists of 4,396 height-diameter observations (trees) from 485 plots (subjects), with the number of observations per plot ranging from 1 to 56. The second data set (data II) was collected in the Rocky Mountain natural region of Alberta. It consists of 2,675 height-diameter observations from 183 plots, with the number of observations per plot ranging from 1 to 99. Figure 11 displays the two data sets. Due to the relatively large sample sizes, some figures presented in this section were cluttered. They were used to show patterns rather than details.

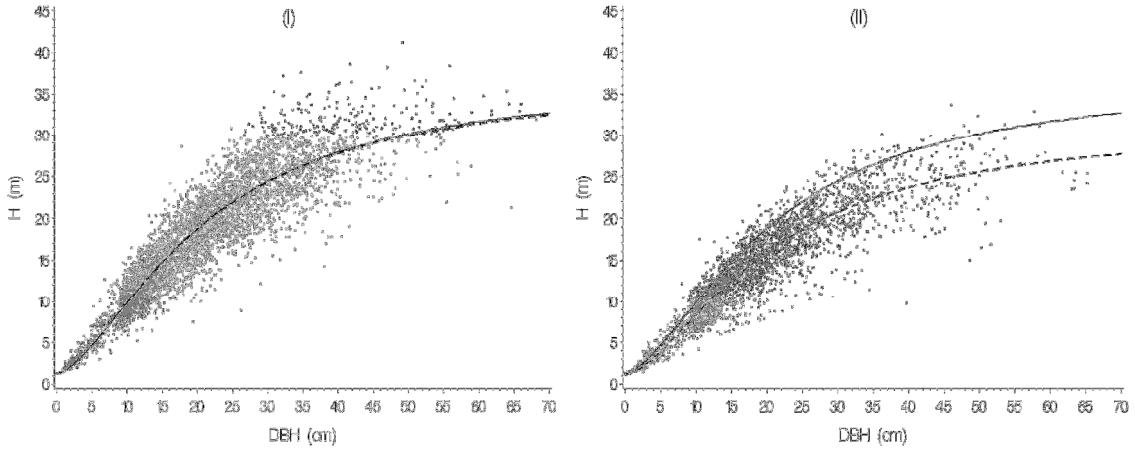

Figure 11. White spruce data sets I (left) and II (right). The solid lines are unadjusted population-based predictions from the white spruce base model [2] with the provincial coefficients. The dashed lines are adjusted population-based predictions from the proportional adjustment method.

6.1 Population-Based Predictions from the Base Model

For the two types of population-based predictions (B-U and B-A) from the base model [2] with the provincial coefficients (b_1=35.7854, b_2=4.8482 and b_3=-1.6040, Table A4, Appendix 1), their difference is minimal for data set I, but substantial for data set II. This can be seen in Figure 11 and from the calculated summary goodness-of-fit statistics in Table 12.

Table 12. Goodness-of-fit statistics for population-based predictions on two white spruce data sets.

Type	Goodness-of-fit measure										
	$\bar{e}$	$\bar{e}$%	SD	MAD	MSE	R^2	CC	MPE	MAPE	e_{10}	δ
					Data set I						
B-U	-0.1174	-0.6662	2.8235	2.1511	7.9839	0.8539	0.9224	-2.61	13.36	51.25	7.9857
B-A	0	0	2.8220	2.1480	7.9620	0.8543	0.9221	-1.93	13.28	51.48	7.9638
					Data set II						
B-U	-2.5253	-17.9770	2.6399	2.7888	13.3436	0.6496	0.8549	-20.25	22.45	68.90	13.3462
B-A	0	0	2.3609	1.7623	5.5719	0.8537	0.9228	-1.93	13.98	53.64	5.5740

Note: B-U = base model, unadjusted. B-A = base model, adjusted. The goodness-of-fit measures are defined in Table 5.

The results shown in Table 12 suggest that, for data set I, the percent mean bias for unadjusted predictions is small ($\bar{e}\% = -0.6662$), and the overall accuracies for unadjusted predictions (δ=7.9857) and adjusted predictions (δ=7.9638) are very close. These suggest that the base model performed well with or without adjustment on data set I. This is also evident in Figure 11 (left graph).

On data set II, the percent mean bias for unadjusted predictions is large ($\bar{e}\% = -17.9770$), and the overall accuracies for unadjusted predictions (δ=13.3462) and adjusted predictions (δ=5.5740) are very different. If unadjusted, the base model performed poorly on data set II. This can be seen in Figure 11 (right graph). The adjusted predictions provided much improvement over the unadjusted predictions on data set II.

6.2 Plot-Specific Predictions from the Base and Mixed Models

Using the base model [2] and mixed model [4] with the provincial coefficients listed in Table A4, Appendix 1, and following the procedures demonstrated earlier for the small population of eight plots, four types of plot-specific predictions were made on the two white spruce data sets. Summary goodness-of-fit statistics from the four types of predictions are listed in Table 13. Relevant spaghetti plots are shown in Figures 12 and 13.

Table 13. Summary goodness-of-fit statistics for plot-specific predictions on two white spruce data sets.

Type	Plot	Goodness-of-fit measure										
		$\bar{e}$	$\bar{e}\%$	SD	MAD	MSE	R^2	CC	MPE	MAPE	e_{10}	δ
						Data set I						
B-U	All	-0.1174	-0.6662	2.8235	2.1511	7.9839	0.8539	0.9224	-2.61	13.36	51.25	7.9857
B-A	All	0	0	1.7885	1.3381	3.1979	0.9415	0.9712	0.29	8.62	32.42	3.1986
M-U	All	0.0164	0.0929	1.5737	1.1792	2.4762	0.9547	0.9769	-0.58	7.69	26.68	2.4768
M-A	All	0	0	1.5635	1.1581	2.4438	0.9553	0.9773	-0.56	7.45	25.89	2.4444
						Data set II						
B-U	All	-2.5253	-17.9770	2.6399	2.7888	13.3436	0.6496	0.8549	-20.25	22.45	68.90	13.3462
B-A	All	0	0	1.5997	1.1960	2.5580	0.9328	0.9668	-0.07	9.48	36.82	2.5589
M-U	All	-0.0956	-0.6804	1.4873	1.1187	2.2203	0.9417	0.9700	-2.36	9.21	33.72	2.2211
M-A	All	0	0	1.4694	1.0899	2.1584	0.9433	0.9709	-1.23	8.73	32.04	2.1592

Note: B-U (base model, unadjusted), B-A (base model, adjusted), M-U (mixed model, unadjusted) and M-A (mixed model, adjusted) are prediction types, and "All" denotes all plots combined. The goodness-of-fit measures are defined in Table 5.

The goodness-of-fit statistics (Table 13) suggest that for both data sets, the overall accuracy (δ) for all plots combined follows a consistent ranking of:

M-A (most accurate) → M-U → B-A → B-U (least accurate)

The results in Table 13 also suggest that for the base model, the overall accuracies of the adjusted predictions (δ=3.1986 for data set I and δ=2.5589 for data set II) are much improved over their unadjusted counterparts, especially for data set II (δ=7.9857 for data set I and δ=13.3462 for data set II).

For the mixed model, the difference between the unadjusted (M-U) and adjusted (M-A) predictions is small. On data set I, δ=2.4768 for M-U and δ=2.4444 for M-A. On data set II, δ=2.2211 for M-U and δ=2.1592 for M-A. Both M-U and M-A are more accurate than the proportionally adjusted B-A predictions from the base model. But the improvement of either M-U or M-A over B-A is not as drastic as B-A over B-U for both data sets.

Note that the summary goodness-of-fit statistics for B-U in Table 13 are identical to those for the population-based B-U in Table 12. This is because the unadjusted plot-specific predictions were identical to the unadjusted population-based predictions. Both types of predictions were obtained directly from the base model [2].

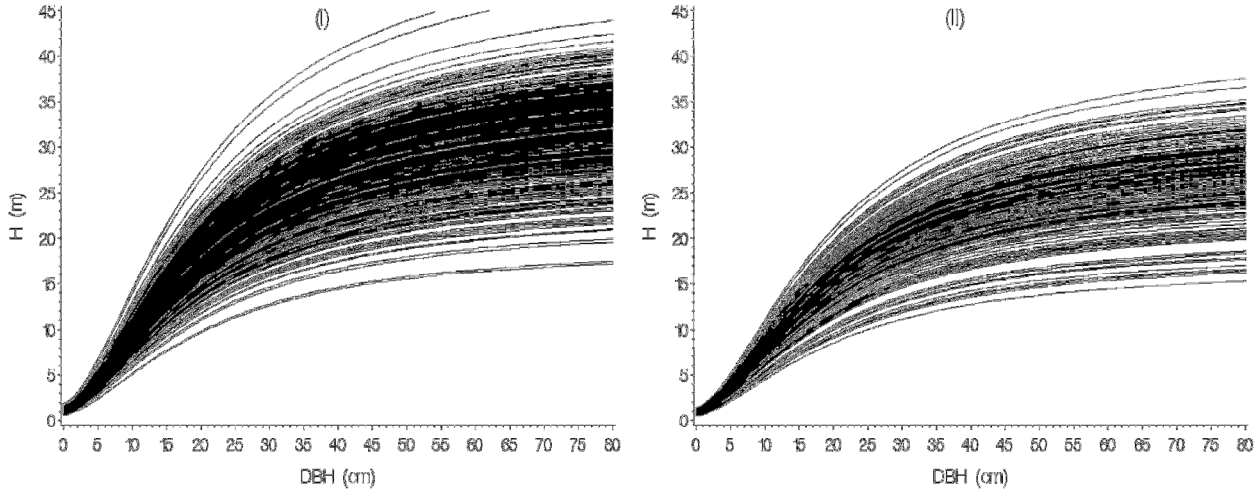

Figure 12. Plot-specific B-A predictions across the DBH range, from the white spruce base model [2] for data sets I and II shown in Figure 11. The curves are proportional to each other with no cross-overs.

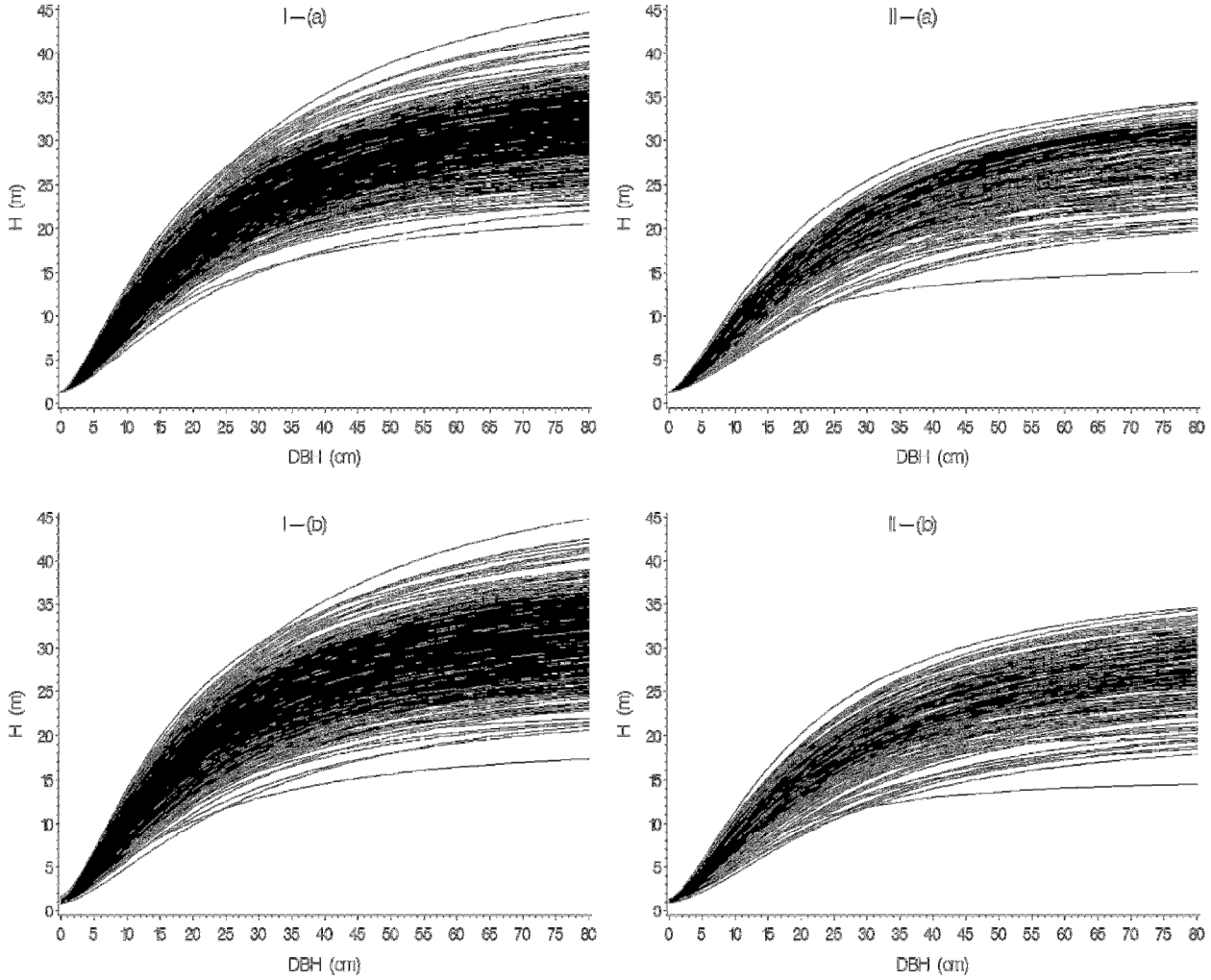

Figure 13. Plot-specific M-U (a) and M-A (b) predictions across the DBH range, from the white spruce mixed model [4] for data sets I and II shown in Figure 11. The curves may cross-over each other.

A comparison of the spaghetti plots in Figures 12 and 13 suggests that for data set I, some adjusted plot-specific height predictions from the base model could exceed 45 meters. This is unlikely for trees in Alberta. Hence, besides the better summary goodness-of-fit statistics, the mixed model predictions (M-U and M-A) are also considered biologically more reasonable than the adjusted predictions from the base model (B-A).

However, since only two plots (out of 485 for data set I) produced unrealistic predictions, and the rest of the plots (99.6%) for data set I and all plots for data set II produced reasonable predictions, the adjusted predictions from the base model can still be considered reasonable in the vast majority of cases, albeit they are somewhat inferior to the mixed model predictions when judged by the results in Table 13.

6.3 Further Assessment of Plot-Specific Predictions

Sometimes, in order to compare the plot-specific predictions in more details, or to reveal the relationships between prediction errors and other variables, some additional analyses are needed. These analyses could also be used to identify outlying observations, variables or plots that caused large prediction errors, which may require cautions or further actions. Here, we selected the plot-specific percent mean bias ($\bar{e}_i\%$), proportional adjustment ratio (PAR_i), overall accuracy (δ_i), number of observations per plot (n_i) and DBH_i as examples, and used some graphical means to briefly show how these analyses could be carried out on the white spruce data sets I and II.

Figure 14 shows the residual plots from the four types of predictions for data set II (the residual plots for data set I were all satisfactory). It can be seen that the "B-U" (base model, unadjusted) type of predictions (Figure 14(a)) shows an unreasonable pattern, whereas other types of predictions all appear to be reasonable.

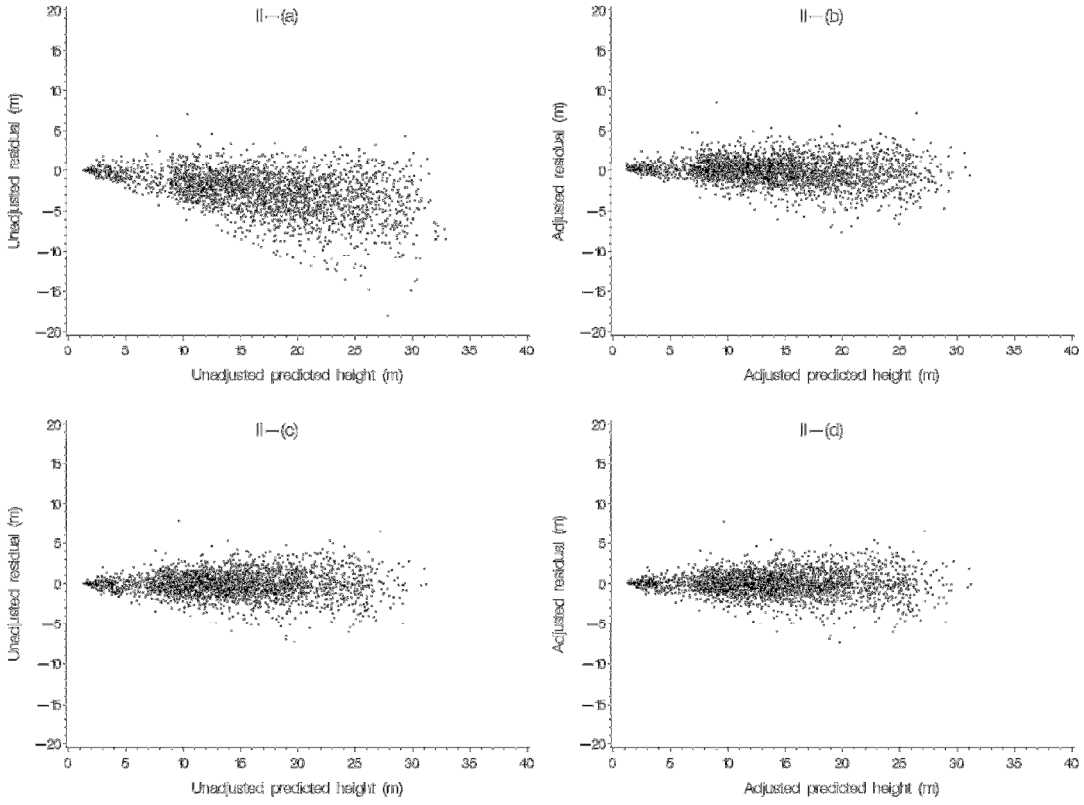

Figure 14. Residual plots for white spruce data set II from B-U (a), B-A (b), M-U (c) and M-A (d) types of predictions. Relevant summary goodness-of-fit statistics for the predictions are listed in Table 13.

Figure 15 shows the plot-specific $\bar{e}\%$ values for the unadjusted predictions from the base and mixed models. The $\bar{e}\%$ values for the adjusted predictions are not shown because they all equal to zeros. It can be seen from Figure 15 that the mixed model predictions are much better than the base model predictions, and that the predictions for data set I are better than those for data set II. For the mixed model predictions none of the absolute $\bar{e}\%$ values exceeds 10% when the number of observations in a plot is greater than five. For the base model predictions the absolute $\bar{e}\%$ values vary considerably across the number of observations per plot.

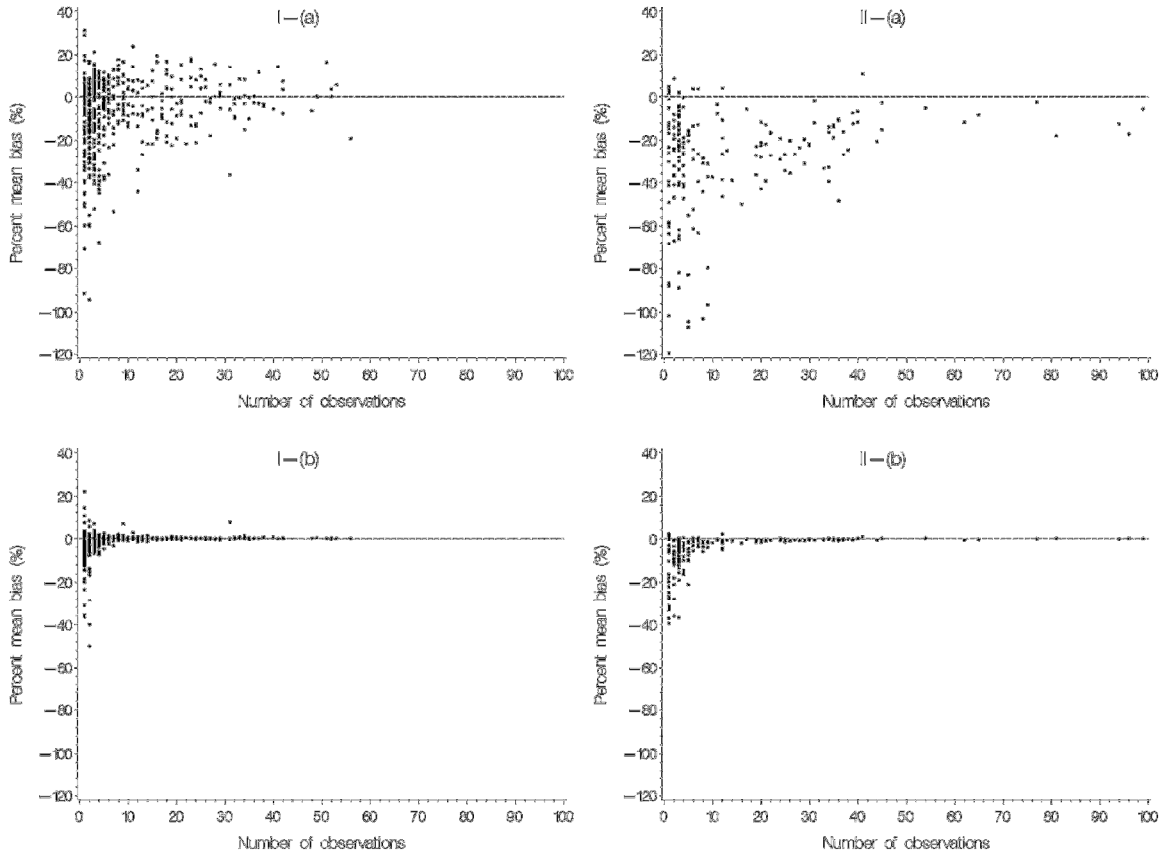

Figure 15. Plot-specific percent mean biases ($\bar{e}\%$) for the unadjusted predictions from the base model (a) and mixed model (b) for the two white spruce data sets (I and II), where the *x*-axis denotes the number of observations per plot.

Figure 15 also shows that, for data set I, except for a few points where the number of observations in the plots is small (i.e., ≤5), the $\bar{e}\%$ values are scattered more evenly around zero. This is not the case for data set II, where the $\bar{e}\%$ values are skewed downward, suggesting that the base and mixed models both over-predicted the data. The over-predictions from the mixed model are less profound than those from the base model.

Figure 16 shows the plot-specific PAR values against the number of observations per plot and DBH. Since a PAR value of one indicates a zero prediction bias, the closer the PAR value to one, the better the predictions. From Figure 16 it can be inferred that, for example:

- The predictions from the mixed model are better than those from the base model;
- The predictions for data set I are better than those for data set II;
- The spread of the PAR values across the gradients of the number of observations per plot (n_i) and DBH is much wider for the base model than that for the mixed model, especially when n_i>5.

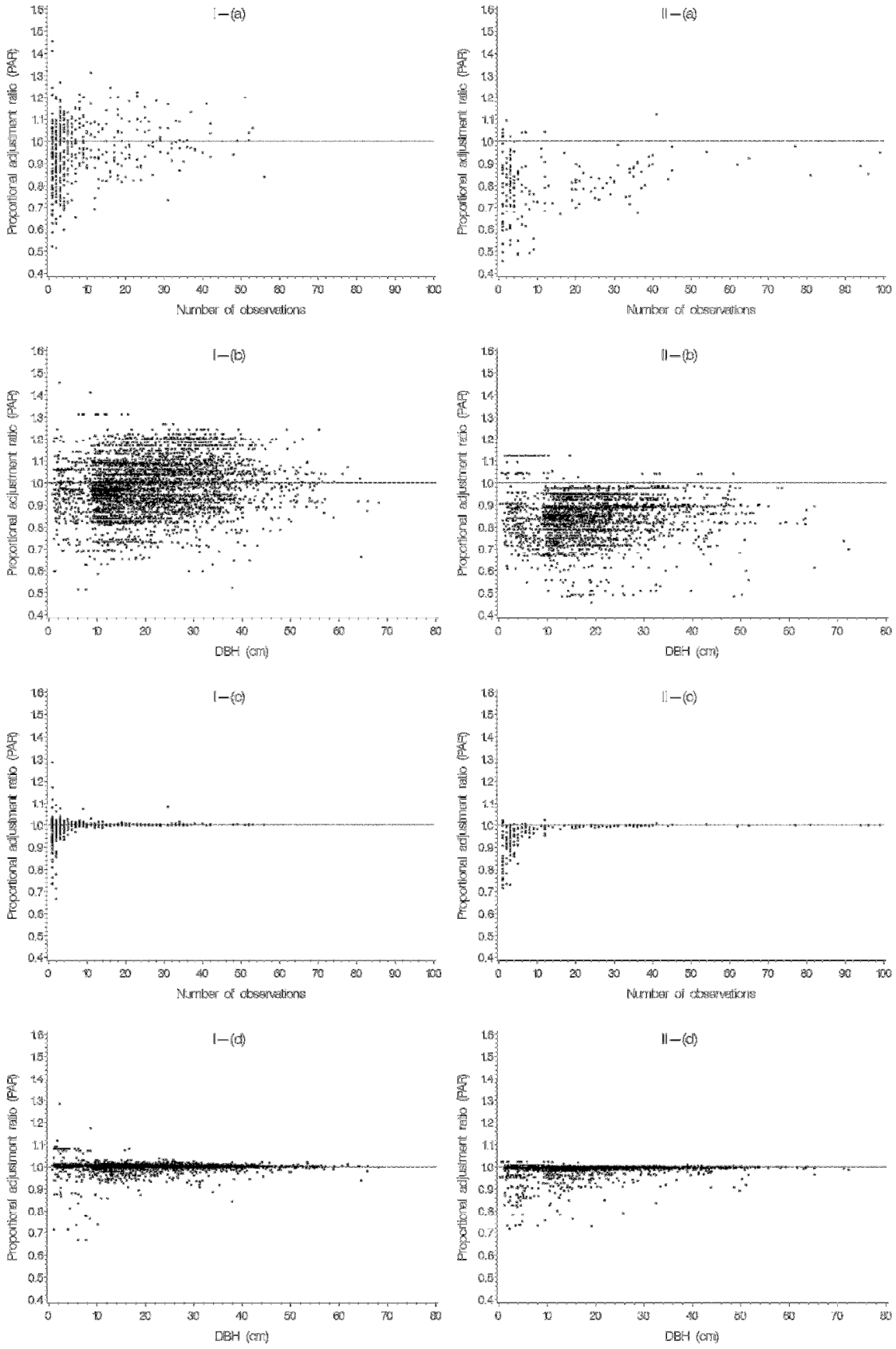

Figure 16. Plot-specific proportional adjustment ratios against the number of observations per plot and DBH from the base model (a, b) and mixed model (c, d) for the white spruce data sets I and II.

Figure 17 shows the plot-specific overall accuracy (δ) values against the number of observations per plot for data set II (the graphs for data set I showed similar patterns). It confirms that the differences between the B-A and B-U predictions are much larger than those between the M-A and M-U predictions. It also suggests that the M-U, M-A and B-A predictions are similar for most plots, and that they are typically better than the B-U predictions.

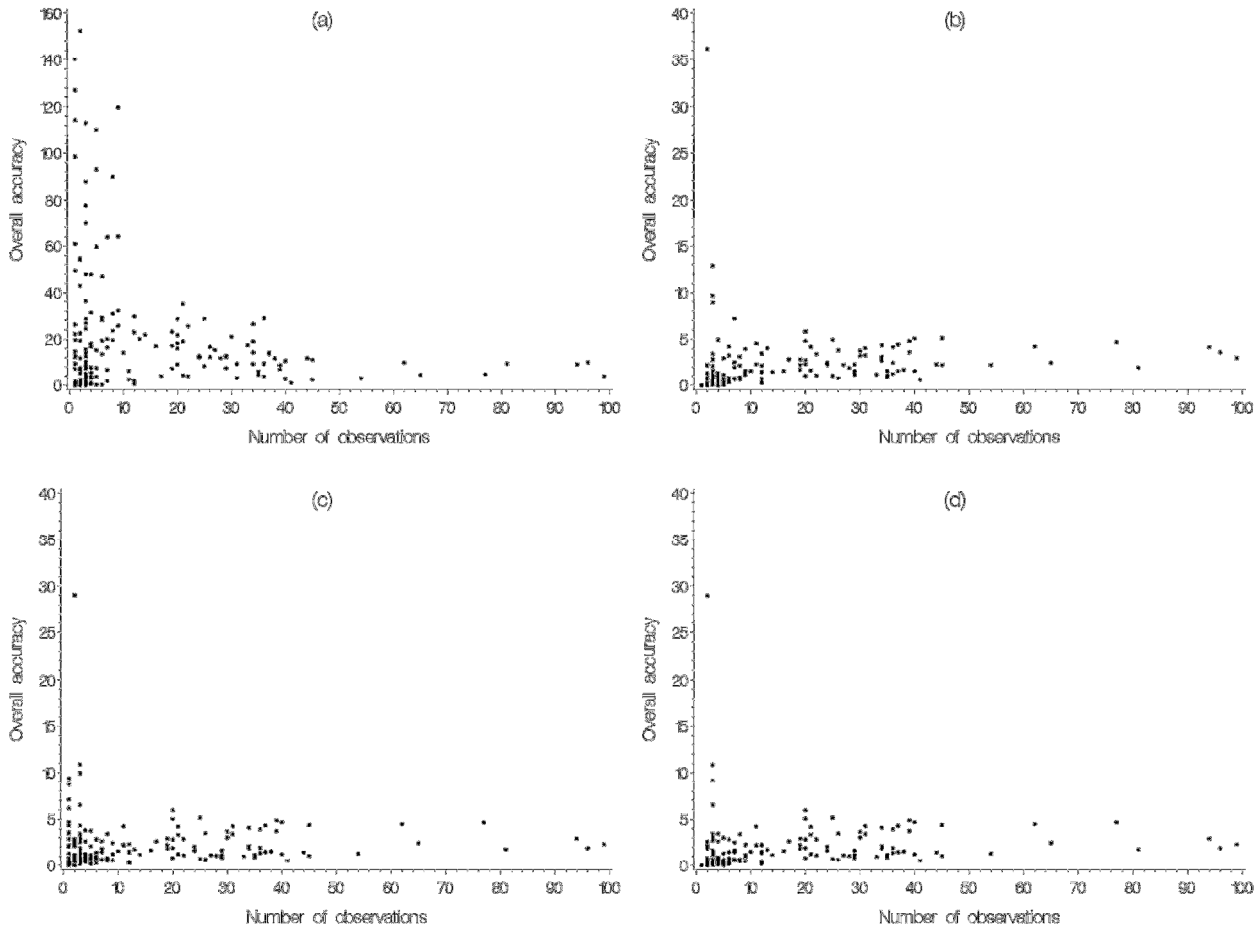

Figure 17. Plot-specific overall accuracy (δ) values for data set II (183 plots) from the unadjusted (a) and adjusted (b) base model, and the unadjusted (c) and adjusted (d) mixed model. Note: a different scale is used for graph (a) in order to include all data points.

More detailed numerical analyses revealed that, for the 183 plots for data set II, the M-U prediction was more accurate than the B-U and B-A predictions for 170 (92.9%) and 82 (44.8%) plots, respectively; and the M-A prediction was more accurate than the B-A and M-U predictions for 147 (80.3%) and 150 (82.0%) plots, respectively.

The fact that the M-U predictions were found to be more accurate than the B-A predictions for 82 (44.8%) out of 183 plots suggested that the B-A predictions were more accurate than the M-U predictions for 101 (55.2%) out of 183 plots. This implied that the computationally simpler B-A predictions were more accurate than the computationally more complex M-U predictions for more plots in data set II, even though the summary goodness-of-fit statistics obtained for all plots combined indicated that the M-U predictions (i.e., δ=2.2211, Table 13) were more accurate than the B-A predictions (δ=2.5589, Table 13). This explains why, in order to get a more complete picture about model performance, the goodness-of-fit statistics obtained from individual plots and the distribution of these statistics should be assessed jointly with the summary goodness-of-fit

statistics obtained for all plots combined (Huang et al. 2009a). Otherwise, some of the unique characteristics about model performance could be overlooked.

Further analyses of the white spruce data set II revealed that, among the 101 plots where the B-A predictions were more accurate than the M-U predictions, most had five or fewer observations per plot. The biggest differences between the M-U and M-A predictions also occurred when the number of observations per plot was five or less.

Based on Figures 14 to 17 and the spaghetti plots shown in Figures 12 and 13, one can identify the specific plots that produced large prediction errors and take appropriate actions if warranted. One can also identify the potential causes that resulted in the large prediction errors. For the two example data sets for white spruce, the potential causes for the large prediction errors could be: a small number of observations in a plot, a small *x*-value (i.e., a small diameter), a unique data set that a fitted model does not automatically apply to, or a combination of these.

7. Recommendations

Since making population-based predictions from a fitted base model is straightforward, the recommendations made here pertain to the plot-specific predictions.

1. For any species in a plot without any height measurement, the base model is used.

2. For any species in a plot with at least one observed height, there are four options for obtaining the predictions for missing heights: B-U (base model, unadjusted), B-A (base model, proportionally adjusted), M-U (mixed model, unadjusted) and M-A (mixed model, proportionally adjusted). Extensive evaluation on the model fitting data (Appendix 2) as well as on independent model application data suggested that under fairly general conditions, when all plots in a population were combined, the following ranking in terms of the overall accuracy measure δ ($\delta = \bar{e}^2 + SD^2$) always holds for the height-diameter models in Alberta:

   M-A (most accurate) → M-U → B-A → B-U (least accurate)

   Therefore, if a generic approach is to be used:

   - Mixed model, proportionally adjusted M-A prediction should be the first choice
   - Mixed model, unadjusted M-U prediction should be the second choice
   - Base model, proportionally adjusted B-A prediction should be the third choice
   - Base model, unadjusted B-U prediction should be the last choice

3. For any particular plot in a population, the "best" prediction could be obtained by comparing the δ values (and other goodness-of-fit measures) from the four types of predictions. Depending on the quantity, quality and the unique characteristics of the plot-specific data, the "best" prediction for a particular plot could be from any one of the four types of predictions.

4. For most practical purposes, it is recommended that adjusting mixed model predictions should be considered only when the percent mean bias of the unadjusted predictions exceeds ±5% (i.e., $|\bar{e}\%| > 5\%$). Otherwise, the gains from adjusting the predictions may not be substantial.

Some important additional notes with regard to the above recommendations are presented here:

1). For plot-specific predictions, the number of observed prior observations per plot (n_i) plays a critical role. When $n_i > 5$, M-U is generally as good as M-A. When $n_i \leq 5$, M-A is often better in more cases than M-U. Large values of prediction errors usually occur when $n_i \leq 5$. Of course, the spread or the distribution of the prior observations also plays an important role.

2). In many cases B-A and M-U predictions are similar, suggesting that B-A is a valid choice in many cases, especially when the mixed model option is unviable due to various reasons, or when the computational complexity and efficiency become blocking issues. However, in these cases the spaghetti plots of the B-A predictions should be checked to ensure that there is not unreasonable prediction across the DBH range in which a model is to be applied. In general whenever possible the M-U predictions are preferred over the B-A predictions.

3). In some cases where the predictions from a mixed model are unrealistic or biologically not meaningful (as usually revealed from spaghetti plots, assisted by a model users' knowledge and experience), the B-

A predictions should be considered. Or else the mixed model predictions must be constrained to biologically meaningful values. Similarly, if the B-A predictions are unrealistic or do not make any biological sense, they must be constrained or the mixed model predictions should be considered.

4). If there is a strong indication of proportionality among the individual plots within a population, the adjusted plot-specific predictions from a base model are viable options. Otherwise, the predictions from a mixed model are better in general as they can account for different shapes (proportionalities) of the plots. In cases where the proportional adjustment ratio obtained from some peculiar data may lead to extremely large or small adjustments that are unrealistic or biologically not meaningful (again, this can be checked from spaghetti plots), one should use the proportional adjustment method with caution. For instance, if the observed height is 5.0 m and the predicted height is 2.5 m at a small diameter, a PAR value of 2 will be obtained. If this ratio is used to multiply the predicted height across the diameter range, over-predictions at larger diameters are expected.

5). The goodness-of-fit of the base model predictions is largely determined by the validity of the proportionality assumption amongst the plots in a population. The goodness-of-fit of the mixed model predictions is highly influenced by the mixed model specification and the number of prior observations per plot.

6). Typically when there is only one observed prior observation for a species in a plot, the B-A prediction is simple and more efficient to obtain. They are also more accurate than the B-U prediction, and as accurate as the M-A prediction. However, since there is only one observation, spaghetti plots from different types of predictions must be assessed before determining the "best" prediction.

7). Since most of the large prediction errors occur when $n_i \leq 5$, whenever feasible, plot-specific predictions should be done with six or more reasonably spread observations per plot (even though a minimum of three observations were found to be enough in many cases).

8). For predictive models (e.g., height-diameter models developed in this study), examining the spaghetti plots is more important in most cases than evaluating the residual plots and goodness-of-fit statistics. Spaghetti plots provide a very powerful and intuitive tool for exposing model behaviors, not only within but also beyond the observed data range where predictions are likely to be made in model application. Standard residual plots and goodness-of-fit statistics (such as those used in this study) typically reflect model performance within the observed data range. Their utility for predictive models is limited.

9). The performance of the models and prediction types is highly dependent on the quality, quantity and relevance (e.g., data range and distribution) of the data. The prediction problems observed in this study were mostly caused by too few data in the plots and/or too narrow range of the data (e.g., the data were concentrated on the small DBH side or at the outer edges of the likely ranges). Improving the quality, quantity and distribution of the data shall be the focus of any data collection.

10). Short-term solutions for data and prediction problems may include: verifying the data to remove any potential measurement error, constraining the model to within some biologically meaningful bounds, adjusting the model predictions for some specific data ranges, or a combination of these.

8. Additional Notes

8.1 Alternative Model Specifications

There are many other base model forms suitable for describing the height-diameter relationship for different species in Alberta (e.g., Huang et al. 1992). Although exceptions do exist, models [1] and [2] were found to be better than other model forms in most cases. The mixed models shown in [3] and [4] were derived from [1] and [2], respectively. They could also be expanded to include a third random parameter (u_3) to become:

[25] $$H = 1.30 + (b_1 + u_1)[1 - \exp(-(b_2 + u_2)DBH)]^{(b_3+u_3)}$$

[26] $$H = 1.30 + \frac{(b_1 + u_1)}{1 + \exp[(b_2 + u_2) + (b_3 + u_3)\ln(DBH)]}$$

Indeed, estimations based on [25] and [26] were obtained for major Alberta tree species (they are available to interested readers). Furthermore, results showed that in many cases, the goodness-of-fit statistics indicated that models [25] and [26] were better than the recommended models [3] and [4]. For instance, Table 14 lists the fit statistics for aspen in subregions 1-6 and 12-21 from models [25] and [3]:

Table 14. Fit statistics for aspen in subregions 1-6 and 12-21 from models [25] and [3].

Parameter	Model [25]				Model [3]			
	Estimate	Std. Err.	*t* value	Pr > |*t*|	Estimate	Std. Err.	*t* value	Pr > |*t*|
b_1	24.4098	0.1450	168.37	<.0001	24.2731	0.1387	174.96	<.0001
b_2	0.09283	0.002303	40.31	<.0001	0.09511	0.001989	47.81	<.0001
b_3	1.4262	0.03557	40.10	<.0001	1.4440	0.03075	46.96	<.0001
$\sigma^2_{u_1}$	8.3669	0.7838	10.67	<.0001	8.6884	0.6614	13.14	<.0001
$\sigma_{u_1u_2}$	-0.04787	0.01017	-4.70	<.0001	-0.03443	0.004205	-8.19	<.0001
$\sigma^2_{u_2}$	0.001224	0.000193	6.35	<.0001	0.000305	0.000034	8.93	<.0001
$\sigma_{u_3u_1}$	-0.09473	0.1331	-0.71	0.4767				
$\sigma_{u_3u_2}$	0.01387	0.002791	4.97	<.0001				
$\sigma^2_{u_3}$	0.2047	0.04375	4.68	<.0001				
σ^2	2.3589	0.04404	53.57	<.0001	2.4253	0.04469	54.27	<.0001
AIC	31634.5				31717.5			
BIC	31686.7				31753.9			
δ	1.9570				2.0389			

Note: AIC, BIC and the overall accuracy measure δ are defined in Table 5.

Judging from the fit statistics listed in Table 14, it could be inferred that model [25], which has smaller AIC, BIC and δ values, is better than model [3] (except for a minor issue of insignificant $\sigma_{u_3u_1}$ at α=0.05). The residual plots from [25] and [3] are similar. They are shown in Figure 18.

However, a comparison of the spaghetti plots outputted from [25] and [3] indicated that [25] could be too flexible to be realistic when predictions were made across the DBH range. This is also shown in Figure 18.

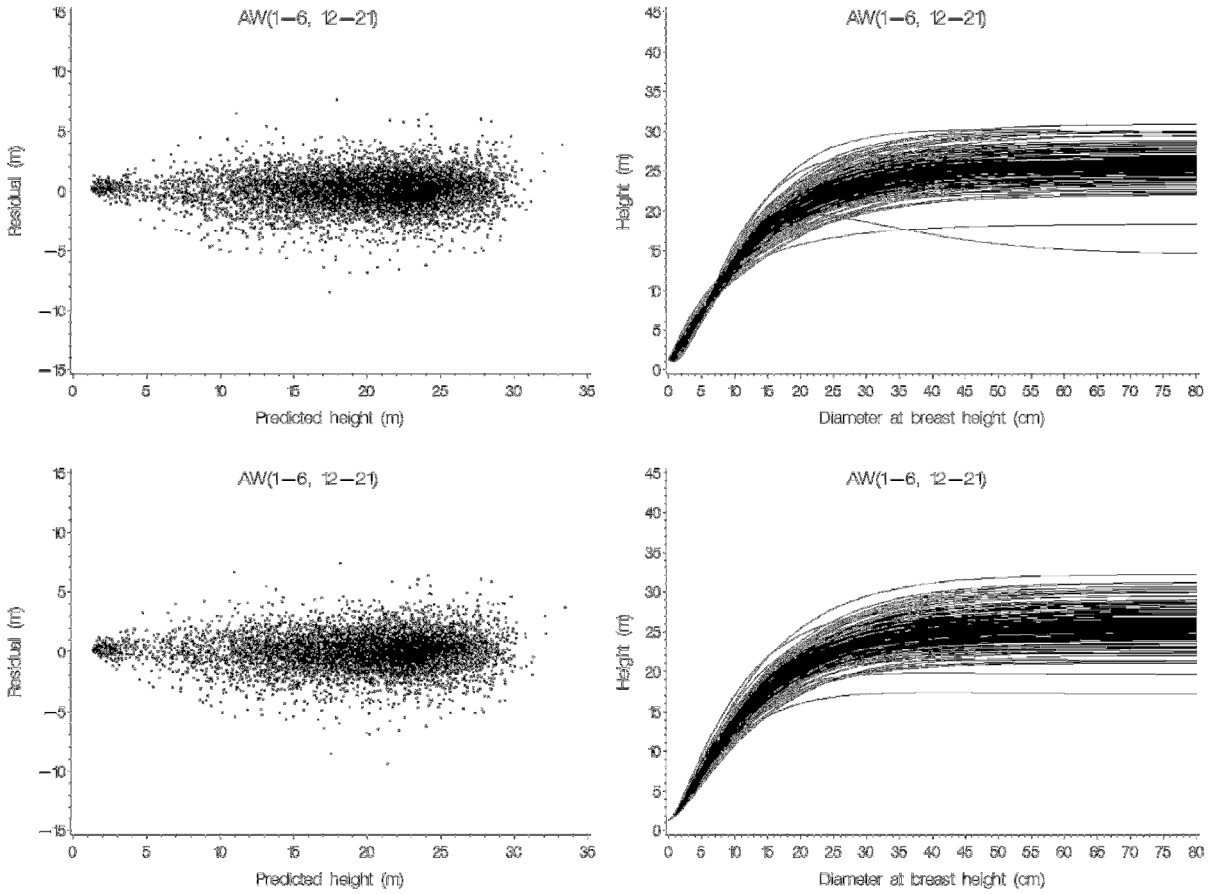

Figure 18. Residual and spaghetti plots from models [25] (top graphs) and [3] (bottom graphs) for aspen in subregions 1-6 and 12-21. 200 randomly selected plots (out of 1,358 plots) are shown in the spaghetti plots.

Similar to the results for aspen in subregions 1-6 and 12-21, the results for aspen in subregions 7-10 also showed that model [25] had better goodness-of-fit statistics than model [3], and the residual plots from both models are similar. However, the spaghetti plots showed that [25] could sometimes produce unrealistic predictions (Figure 19).

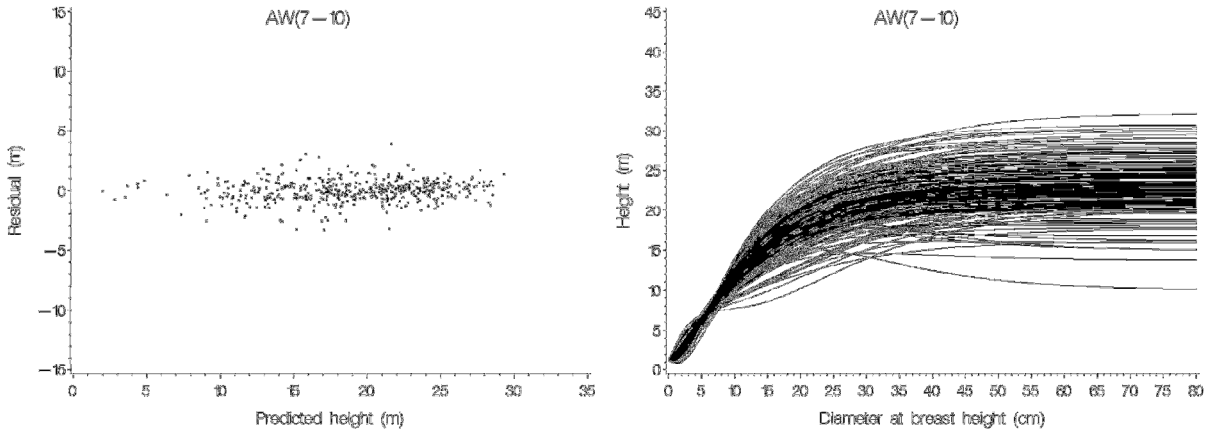

Figure 19. Residual and spaghetti plots from model [25] for aspen in subregions 7-10. Corresponding residual and spaghetti plots from model [3] are shown in Figure A2 of Appendix 1.

The recommended mixed models [3] and [4] can be considered reduced forms of models [25] and [26], respectively. For some species with unique data, further reductions may be necessary. For instance, for white birch (BW), when model [3] was estimated, fit statistics (Table 15) and residual plot (Figure 20(a)) indicated that [3] fitted the BW data well. However, an examination of the spaghetti plot (Figure 20(b)) suggested that [3] could produce biologically unrealistic predictions at many DBH ranges where predictions are likely to be made in model applications.

Table 15. Fit statistics for white birch from model [3].

Parameter	Estimate	Std. Err.	*t* value	Pr > |*t*|	AIC	BIC	δ
b_1	25.9979	1.9218	13.53	<.0001	6899.4	6930.3	0.749
b_2	0.04878	0.006109	7.99	<.0001			
b_3	1.0080	0.02590	38.92	<.0001			
$\sigma^2_{u_1}$	262.19	42.2803	6.2	<.0001			
$\sigma_{u_1u_2}$	-0.6393	0.1051	-6.08	<.0001			
$\sigma^2_{u_2}$	0.001636	0.000262	6.24	<.0001			
σ^2	0.9709	0.03649	26.61	<.0001			

Note: AIC, BIC and the overall accuracy measure δ are defined in Table 5. Corresponding fit statistics from model [27] (i.e., model [3] with one random parameter u_1 only) are listed in Table A2 of Appendix 1.

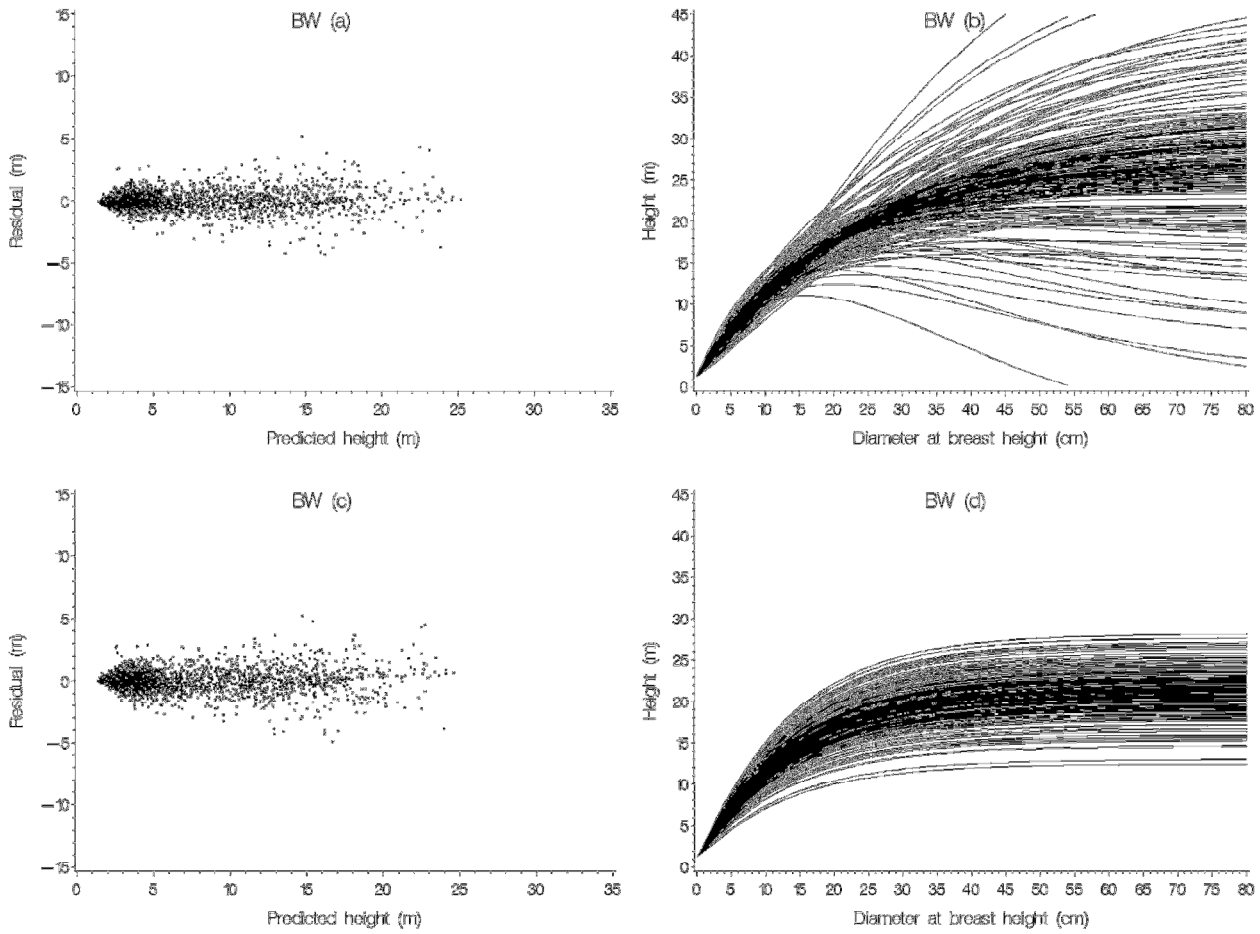

Figure 20. Residual and spaghetti plots from mixed models [3] (a and b) and [27] (c and d) for white birch (BW). The original data for BW are shown in Appendix 1 (Figure A3).

Since the spaghetti plot shown in Figure 20(b) indicated that [3] was too flexible to be a prediction model for BW, a reduced form of [3] expressed in [27] with only one random parameter u_1 was fitted (alternative model forms with one random parameter only were found to be inferior to [27]):

[27] $$H = 1.30 + (b_1 + u_1)[1 - \exp(-b_2 DBH)]^{b_3}$$

Residual and spaghetti plots from [27] are also shown in Figure 20 (fit statistics are listed in Table A2 of Appendix 1). Had the residual plot and goodness-of-fit statistics been used as criteria, one could have concluded that [3] were better. But the spaghetti plot clearly shows that, as a predictive model, [27] is more reasonable than [3] for BW.

8.2 Model Selection Criteria

As a general rule, when selecting the "best" model, three fundamental principles should be followed:

1. First and foremost, biological considerations and logical interpretations are of the greatest significance and importance when selecting the "best" model. No matter how good a model may appear to be, it matters little if the model does not make any biological sense, or if it divagates from the relevant subject matter considerations. The best model is the one that makes biological sense.

2. Second, appropriate graphical techniques are generally more important and more powerful than other diagnostic techniques and statistical measures in revealing model behaviors. Standard, studentized, normalized (and lagged if applicable) residual plots are most common in this regard. But they are not enough. Indeed, as demonstrated in the above examples, they could even be misleading if a satisfactory residual plot was interpreted to symbolize a good model. Residual plots or their variants (e.g., observed versus predicted values) are much less powerful in revealing model behaviors than spaghetti plots.

   For models developed for prediction purposes, which almost all models are in forest growth and yield studies, spaghetti plots are most capable in revealing model behaviors not only within but also beyond the observed data ranges where predictions are likely to be made in model application. They should be examined routinely before the "best" model is chosen, particularly when mixed models are involved. Depending on the NMM method used, mixed models can alter the inherent shapes of their fixed model counterparts. A sigmodal non-declining function such as the Chapman-Richards function can become a non-sigmodal declining function in predictions (e.g., Figures 19 and 20).

3. Third, once the last two principles are duly considered, many statistical measures can be used to quantify and to further assess the goodness-of-fit of models. For models fitted for prediction purposes (as oppose to for descriptive and/or hypothesis testing purposes), the goodness-of-fit on model application data is more important than that on model fitting data. The goodness-of-fit statistics obtained on model fitting data typically inflate the goodness-of-fit of a model. This could explain, at least in part, why sometimes the reported goodness-of-fit of a model was hard to achieve in real-world model applications.

In practice, when conflicts occur amongst the three fundamental principles, the recommended approach is that graphical techniques trump statistical measures, and biological and subject matter considerations trump graphical techniques. In fact, biological and subject matter considerations trump graphical techniques and statistical measures combined. It is the goal of and a challenge for model developers to develop models that make reasonable senses on all three fronts.

The fundamental principles also connote to some more specific caveats that practitioners should be aware of, for instance:

1). The "best" model cannot be selected or judged based on fit statistics (e.g., AIC, BIC, MSE, δ) alone;
2). The "best" model shall not be selected based on the significance of statistics alone;
3). A satisfactory residual plot does not necessarily mean a good model;
4). Fit statistics and residual plots usually only reflect model behaviors within the observed data range.

Indeed, simply based on the better fit statistics and residual plots to choose a prediction model could be very misleading. This is particularly true for mixed models.

It may be worth repeating that the shape of a mixed model could be very different from that of the base model the mixed model was derived from. The shape of a mixed model could also become too flexible to be biologically meaningful due to the inclusion of statistically significant random parameters and the use of the specific NMM technique such as the first-order method, even though the fit statistics and residual plots could still appear "nice" or "increasingly better" when more random parameters were included.

For predictive models, fit statistics and residual plots are not enough. Predictive models are frequently extrapolated beyond the observed data range (with some limits of course). The spaghetti plots fill in the gaps left by the fit statistics and residual plots. They can be used to ensure that a specified mixed model has enough but not too much flexibility to track the trends of specific data in a population, without resorting solely to some kind of statistical measures (or worse yet, statistical tests) to determine if additional random parameters are needed in model specification and selection.

In this study we recommended models [3] and [4] after an extensive evaluation of fit statistics, residual plots, and more importantly, spaghetti plots and biological considerations. Our emphasis was to find models that have broad application scopes and ranges for many species in many different geographical regions across the province, and that are robust under different conditions for different data sets, rather than to find the "best" model for a specific species in a specific region. Of course, for a specific data set, it is possible that a different model could be better than [3] and [4]. If that is the case, one could use the different model for the specific data set, so long as the model is appropriately evaluated across the likely model application range.

8.3 Expanded Height-Diameter Models

The height-diameter relationship developed in this study can be written in a general form as:

$$H = f(DBH)$$

where *f* denotes some nonlinear function. The height-diameter relationship can be expanded to include other variables (e.g., Huang and Titus 1994, Meng et al. 2008, Huang et al. 2009c):

$$H = f(DBH, \text{other variables})$$

where "other variables" can mean other tree and/or stand related variables, such as:

- Stand density and competition (e.g., basal area/ha, stems/ha, relative density index, stand density index, point densities, different spatial and/or a-spatial point densities and competition measures);
- Site quality (e.g., site index, dominant/co-dominant height, top height);
- Age (e.g., tree age, stand age);

- Bio-geo-climatic variables (e.g., ecosite, habitat type, slope, aspect, elevation, growing degree-days, mean annual temperature/precipitation, moisture/dryness index, solar radiation, GPS coordinates);
- Species composition (e.g., tree number proportion, basal area proportion);
- Crown size, crown length, crown ratio and/or crown closure class;
- Others (e.g., soil type/depth/salinity/nutrient regime, distance to the coastline, wind speed, etc.).

For instance, a study in Alberta showed that including wind speed and/or crown class into the height-diameter model of lodgepole pine significantly improved the fit of the model (Meng et al. 2008). Another study in Finland showed that including distance to/from the coastline into the height-diameter model significantly improved the height predictions for the Finnish National Forest Inventory (Eerikäinen 2009).

Since many of the variables listed above are related in one way or the other, unless the study objective is to assess their impacts on height-diameter relationship, it is uncommon to include more than three additional variables into a height-diameter model. Of course, when more (statistically significant) variables are added into a model, the model generally becomes more accurate (i.e., δ becomes smaller, $\delta = \bar{e}^2 + SD^2$).

However, in practical applications, the true benefit of including additional variables into height-diameter models must be weighed carefully against the time and costs associated with collecting and analyzing the additional variables. The increased costs of obtaining these variables need to be balanced out against the realized, palpable gains in practice. Otherwise, the prime purpose of model building, which is to use "simple and inexpensive" *x*-variable(s) to predict "difficult, complex and expansive" *y*-variable, could become a practically meaningless academic exercise.

It has generally been taken for granted that the inclusion of additional (statistically significant) variables into a model would automatically result in better predictions (helping to justify the increased time and costs associated with measuring the additional variables). But this is not always true for nonlinear mixed models, which have some capability to account for the impacts of known and unknown variables left-out by the models without actually requiring that these variables to be identified or measured.

For major Alberta tree species, expanded height-diameter models with up to seven *x*-variables were developed. They are currently available to interested readers and will be released to general public later.

8.4 Multi-level Height-Diameter Models

The height-diameter data used in this study are hierarchical in nature. They could also be modeled using multi-level NMM technique. In the simplest scenario, plots can be considered level one and trees within plots can be considered level two. Since some of the plots are clustered together, within known forest management units, and have been measured repeatedly, other levels of analyses could be added to represent, e.g.:

- Plot clusters
- Forest management units
- Measurement years

Interested readers could carry out such analyses following the procedures demonstrated in, e.g., Robinson and Wykoff (2004), and Yang and Huang (2011b). Our analyses (not presented here) suggested that the practical benefit in terms of accuracy gains of such analyses was very limited, relative to the substantially increased complexity in computation. In many cases, we had to reduce the sample size, the number of hierarchical levels, or the number of random parameters to achieve convergence. Hence, for practical purposes, we do not recommend fitting multi-level height-diameter models at this time.

References

Alberta Sustainable Resource Development. 2005. Permanent sample plot field procedures manual. Forest Management Branch, Alberta Sustainable Resource Development, Edmonton, Alberta. 127 total pages.

Davidian, M., and Giltinan, D.M. 1995. Nonlinear models for repeated measurement data. Chapman & Hall, New York. 360 p.

Eerikäinen, K. 2009. A multivariate linear mixed-effects model for the generalization of sample tree heights and crown ratios in the Finnish National Forest Inventory. Forest Science 55: 480-493.

Huang, S. 2002. Validating and localizing growth and yield models: procedures, problems and prospects. An invited presentation given at "Reality, Models and Parameter Estimation", June 2-5, 2002, jointly sponsored by IUFRO and Instituto Superior de Agronomia, Sesimbra, Portugal, 28 p.

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 p.

Huang, S., and Titus, S.J. 1994. An age-independent individual tree height prediction model for boreal spruce-aspen stands in Alberta. *Canadian Journal of Forest Research* 24: 1295-1301.

Huang, S., Titus, S.J., and Wiens, D.P. 1992. Comparison of nonlinear height-diameter functions for major Alberta tree species. *Canadian Journal of Forest Research* 22: 1297-1304.

Huang, S., Titus, S.J., Lakusta, T.W., and Held, R.J. 1994. Ecologically based individual tree height-diameter models for major Alberta tree species. Land and Forest Service, Alberta Environmental Protection, Tech. Rep. Pub. No. T/291. 27 p.

Huang, S., Price, D., and Titus, S.J. 2000. Development of ecoregion-based height-diameter models for white spruce in boreal forests. *Forest Ecology and Management* 129: 125-141.

Huang, S., Meng, S.X., and Yang, Y. 2009a. Assessing the goodness-of-fit of forest models estimated by nonlinear mixed model methods. *Canadian Journal of Forest Research* 39: 2418-2436.

Huang, S., Meng, S.X. and Yang, Y. 2009b. Prediction implications of nonlinear mixed-effects forest biometric models estimated with a generalized error structure. Proceedings of Joint Statistical Meetings, Section on Statistics and the Environment, August 1–6, 2009, Washington, D.C. American Statistical Association, Alexandria, Va. pp. 1174–1188.

Huang, S., Wiens, D.P., Yang, Y., Meng, S.X., and VanderSchaaf C.L. 2009c. Assessing the impacts of species composition, top height and density on individual tree height prediction of quaking aspen in boreal mixedwoods. *Forest Ecology and Management* 258: 1235-1247.

Meng, S.X., and Huang, S. 2009. Improved calibration of nonlinear mixed-effects models demonstrated on a height growth function. *Forest Science* 55: 238-248.

Meng, S.X., Huang, S., Lieffers, V., Nunifu, T., and Yang, Y. 2008. Wind speed and crown class influence the height-diameter relationship of lodgepole pine: Nonlinear mixed effects modeling. *Forest Ecology and Management* 256: 570-577.

Natural Regions Committee 2006. Natural Regions and Subregions of Alberta. Compiled by D.J. Downing and W.W. Pettapiece. Government of Alberta. Pub. No. T/852. 254 pp.

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

Robinson, A.P., and Wykoff, W.R. 2004. Imputing missing height measures using a mixed-effects modeling strategy. *Canadian Journal of Forest Research* 34:2492–2500.

Vonesh, E.F. and Chinchilli, V.M. 1997. Linear and nonlinear models for the analysis of repeated measurements. Marcel Dekker, New York. 560 p.

Yang, Y. and Huang, S. 2011a. Comparison of different methods for fitting nonlinear mixed forest models and for making predictions. *Canadian Journal of Forest Research* 41: 1671-1686.

Yang, Y., and Huang, S. 2011b. Estimating a multilevel dominant height-age model from nested data with generalized errors. *Forest Science* 57:102–116.

Appendix 1. Fit Statistics and Diagnostic Plots for Height-Diameter Models.

Table A1. Parameter estimates and fit statistics for aspen and deciduous species combined.

	Aspen								Deciduous combined	
	Model [1]				Model [3]				Model [1]	Model [3]
Sub-region	7 to 10	11	Others	Prov.	7 to 10	11	Others	Prov.	Prov.	Prov.
b_1	27.8444	26.9085	25.9208	26.6988	21.5750	25.3145	24.2731	24.7105	26.7947	25.1911
b_2	0.0554	0.0717	0.0780	0.0707	0.09508	0.08479	0.09511	0.08791	0.0693	0.07710
b_3	1.3352	1.2526	1.2472	1.1996	1.4822	1.4096	1.4440	1.3934	1.2518	1.2939
$\sigma^2_{u_1}$					16.4815	7.7043	8.6884	8.9402		10.7569
$\sigma_{u_1u_2}$					-0.05070	-0.02313	-0.03443	-0.02921		-0.03519
$\sigma^2_{u_2}$					0.000530	0.000248	0.000305	0.000295		0.000306
σ^2					1.5407	2.7003	2.4253	2.5523		2.7128
AIC					2127.0	34645.7	31717.5	68759.1		94109.2
BIC					2149.8	34682.1	31753.9	68800.8		94152.1
N	490	8323	7802	16615	490	8323	7802	16615	22616	22616
m	192	1330	1358	2880	192	1330	1358	2880	3431	3431
$\bar{e}$	0.0058	-0.0036	-0.0034	-0.0049	0.0012	0.0016	0.0040	0.0017	-0.0007	0.0084
$\bar{e}\%$	0.0306	-0.0166	-0.0174	-0.0238	0.0063	0.0076	0.02077	0.0081	-0.0035	0.0447
SD	2.9245	2.4038	2.2880	2.4093	0.9614	1.5140	1.4279	1.4604	2.4448	1.5209
MAD	2.3299	1.8694	1.7482	1.8525	0.7136	1.1393	1.0766	1.0980	1.8636	1.1343
MSE	8.5880	5.7798	5.2364	5.8055	0.9224	2.2918	2.0386	2.1325	5.9774	2.3130
R^2	0.6994	0.8324	0.8562	0.8400	0.9675	0.9335	0.9440	0.9412	0.89	0.9563
CC	0.8238	0.9083	0.9224	0.9128	0.9831	0.9652	0.9709	0.9694	0.94	0.9776
MPE	-2.76	-1.69	-1.87	-1.97	-0.68	-0.77	-0.76	-0.82	-2.43	-0.98
MAPE	13.90	9.59	10.04	10.15	4.67	6.11	6.58	6.30	11.97	7.81
e_{10}	54.49	36.54	37.63	38.04	9.39	16.54	18.47	17.32	43.74	24.05
δ	8.5529	5.7784	5.2351	5.8048	0.9243	2.2921	2.0389	2.1327	5.9768	2.3131

Note: Subregions are defined in Table 2, "Prov." denotes provincial (i.e., all subregions combined), *N* is the total number of observations, and *m* is the number of plots, $\sigma^2_{u_1}$, $\sigma^2_{u_2}$ and $\sigma_{u_1u_2}$ are variances and covariance for the random parameters, σ^2 is residual variance. The goodness-of-fit measures (AIC, BIC, $\bar{e}$, $\bar{e}\%$, SD, MAD, MSE, R^2, CC, MPE, MAPE, e_{10} and δ) are defined in Table 5.

Table A2. Parameter estimates and fit statistics for balsam poplar and white birch.

	Balsam poplar						White birch	
	Model [1]			Model [3]			Model [1]	Model [3]
Subregion	7 to 11	Others	Prov.	7 to 11	Others	Prov.	Prov.	Prov.
b_1	27.0288	31.0457	28.3642	26.2016	27.1012	26.5900	22.4814	19.7369
b_2	0.0577	0.0405	0.0510	0.05536	0.05746	0.05622	0.0673	0.08067
b_3	1.2407	0.9862	1.1338	1.0899	1.1380	1.1094	1.1496	1.1258
$\sigma^2_{u_1}$				14.1752	19.3487	17.4954		11.2387
$\sigma_{u_1u_2}$				-0.03296	-0.05695	-0.04696		
$\sigma^2_{u_2}$				0.000204	0.000256	0.000240		
σ^2				2.1735	2.9184	2.4426		1.1304
AIC				9833.6	6498.3	16371.4		7027.5
BIC				9863.6	6526.5	16405.4		7049.6
N	2412	1501	3913	2412	1501	3913	2088	2088
m	537	415	952	537	415	952	608	608
$\bar{e}$	0.0313	-0.0224	-0.0023	0.0118	-0.0065	.0057	0.0253	0.0121
$\bar{e}\%$	0.1905	-0.1244	-0.0137	0.0720	-0.0362	0.0335	0.3288	0.1573
SD	2.3211	2.4120	2.3911	1.3134	1.5141	1.3852	1.6307	0.9612
MAD	1.7593	1.8738	1.8250	0.9609	1.1379	1.0230	1.1691	0.6991
MSE	5.3930	5.8258	5.7201	1.7246	2.2911	1.9184	2.6623	0.9236
R^2	0.9123	0.8779	0.8992	0.9719	0.9519	0.9662	0.9034	0.9664
CC	0.9545	0.9344	0.9469	0.9857	0.9749	0.9827	0.9496	0.9829
MPE	-1.48	-4.05	-2.98	-1.09	-2.33	-1.53	-3.49	-2.19
MAPE	13.21	13.22	13.22	8.24	8.76	8.41	18.00	12.01
e_{10}	49.25	44.37	47.94	24.75	24.18	24.58	62.40	44.06
δ	5.3885	5.8180	5.7172	1.7253	2.2926	1.9189	2.6598	0.9240

Note: Subregions are defined in Table 2, "Prov." denotes provincial (i.e., all subregions combined), *N* is the total number of observations, *m* is the number of plots, $\sigma^2_{u_1}$, $\sigma^2_{u_2}$ and $\sigma_{u_1u_2}$ are variances and covariance for the random parameters, and σ^2 is residual variance. The goodness-of-fit measures (AIC, BIC, $\bar{e}$, $\bar{e}\%$, SD, MAD, MSE, R^2, CC, MPE, MAPE, e_{10} and δ) are defined in Table 5. For white birch (BW), only one random parameter u_1 was used in the mixed model [3] due to data and prediction issues.

Table A3. Parameter estimates and fit statistics for lodgepole pine.

	Model [2]					Model [4]				
Subregion	7 to 9	10	11	Others	Prov.	7 to 9	10	11	Others	Prov.
b_1	21.2291	28.6413	36.3029	38.9208	35.7547	18.8448	26.0553	29.2765	33.4784	27.5659
b_2	4.1138	3.7850	3.7620	3.8340	3.8234	3.4986	3.0418	3.0320	3.0063	3.0715
b_3	-1.6672	-1.4192	-1.2816	-1.2172	-1.2824	-1.5232	-1.2077	-1.1951	-1.0743	-1.2163
$\sigma^2_{u_1}$						18.0265	21.7236	17.6153	30.9620	23.5958
$\sigma_{u_1u_2}$						1.2058	1.2360	0.7772	2.0049	1.0124
$\sigma^2_{u_2}$						0.1391	0.1900	0.1136	0.2050	0.1452
σ^2						0.9720	1.3496	1.5973	1.9026	1.4718
AIC						11461.7	67424.1	89362.3	11553.6	181256.9
BIC						11485.5	67460.9	89400.9	11581.5	181300.7
N	3721	19458	24814	2964	50957	3721	19458	24814	2964	50957
m	221	1411	1830	401	3863	221	1411	1830	401	3863
$\overline{e}$	0.0074	-0.0020	-0.0005	0.0028	-0.0013	-0.0006	-0.0007	-0.0007	0.0002	-0.0008
$\overline{e}\%$	0.0533	-0.0108	-0.0025	0.0133	-0.0066	-0.0041	-0.0038	-0.0036	0.0008	-0.0039
SD	1.8436	2.5894	2.3057	2.4822	2.6058	0.9438	1.0999	1.2014	1.2656	1.1487
MAD	1.4332	2.0678	1.7963	1.9717	2.0542	0.7091	0.8361	0.9156	0.9784	0.8724
MSE	3.4008	6.7058	5.3165	6.1653	6.7906	0.8906	1.2098	1.4433	1.6013	1.3195
R^2	0.7421	0.6496	0.7538	0.6955	0.7175	0.9324	0.9368	0.9332	0.9208	0.9451
CC	0.8532	0.7871	0.8596	0.8210	0.8353	0.9648	0.9671	0.9651	0.9581	0.9716
MPE	-1.66	-2.41	-1.55	-1.44	-2.23	-0.66	-0.57	-0.51	-0.48	-0.55
MAPE	10.96	12.26	9.40	9.78	11.64	5.58	5.01	4.73	4.88	4.91
e_{10}	44.72	48.80	34.77	38.93	44.96	13.73	10.78	9.12	10.56	10.16
δ	3.3990	6.7051	5.3161	6.1612	6.7903	0.8908	1.2099	1.4433	1.6018	1.3196

Note: Subregions are defined in Table 2, "Prov." denotes provincial (i.e., all subregions combined), *N* is the total number of observations, *m* is the number of plots, $\sigma^2_{u_1}$, $\sigma^2_{u_2}$ and $\sigma_{u_1u_2}$ are variances and covariance for the random parameters, and σ^2 is residual variance. The goodness-of-fit measures (AIC, BIC, $\overline{e}$, $\overline{e}\%$, SD, MAD, MSE, R^2, CC, MPE, MAPE, e_{10} and δ) are defined in Table 5.

Table A4. Parameter estimates and fit statistics for white spruce.

	Model [2]					Model [4]				
Subregion	7 to 9	10	11	Others	Prov.	7 to 9	10	11	Others	Prov.
b_1	32.3269	34.9926	34.6665	36.2971	35.7854	24.2756	32.2041	38.6180	35.5340	35.6912
b_2	3.9573	4.2859	5.3149	4.6580	4.8482	4.2848	4.3066	4.5318	4.5852	4.4737
b_3	-1.2246	-1.3676	-1.7700	-1.5743	-1.6040	-1.5075	-1.3920	-1.4163	-1.5398	-1.4524
$\sigma^2_{u_1}$						28.6345	45.5216	36.9042	25.2084	40.7210
$\sigma_{u_1u_2}$						1.3682	1.9412	1.5107	1.3413	1.5971
$\sigma^2_{u_2}$						0.09606	0.1394	0.1284	0.1211	0.1348
σ^2						1.4274	2.1279	3.2690	3.4560	3.0532
AIC						6109.7	35225.2	61177.8	88061.1	193005.9
BIC						6130.6	35258.5	61215.5	88099.1	193050.5
N	1766	9063	14150	20314	45293	1766	9063	14150	20314	45293
m	146	864	1604	1684	4298	146	864	1604	1684	4298
$\overline{e}$	-0.0241	-0.0182	0.0446	-0.0013	0.0074	-0.0103	-0.0032	0.0033	-0.0007	-0.0003
$\overline{e}\%$	-0.1887	-0.1101	0.2326	-0.0060	0.0377	-0.0808	-0.0194	0.0170	-0.0032	-0.0014
SD	1.7383	2.4395	2.9029	2.6301	2.9481	1.1361	1.3748	1.6937	1.7554	1.6423
MAD	1.2528	1.8550	2.2101	2.0055	2.2681	0.8683	1.0315	1.2644	1.3321	1.2319
MSE	3.0256	5.9529	8.4300	6.9182	8.6915	1.2900	1.8899	2.8686	3.0814	2.6971
R^2	0.9012	0.8770	0.8897	0.8537	0.8606	0.9578	0.9609	0.9624	0.9348	0.9567
CC	0.9475	0.9341	0.9422	0.9210	0.9253	0.9783	0.9800	0.9808	0.9660	0.9778
MPE	-3.26	-3.32	-1.05	-1.90	-2.38	-1.64	-1.28	-1.14	-1.00	-1.12
MAPE	11.49	12.85	13.44	10.33	13.02	8.05	7.51	8.10	6.95	7.48
e_{10}	42.64	49.27	50.98	37.58	49.54	29.11	24.47	27.60	21.09	24.06
δ	3.0222	5.9516	8.4288	6.9175	8.6911	1.2907	1.8901	2.8688	3.0815	2.6972

Note: Subregions are defined in Table 2, "Prov." denotes provincial (i.e., all subregions combined), *N* is the total number of observations, *m* is the number of plots, $\sigma^2_{u_1}$, $\sigma^2_{u_2}$ and $\sigma_{u_1u_2}$ are variances and covariance for the random parameters, and σ^2 is residual variance. The goodness-of-fit measures (AIC, BIC, $\overline{e}$, $\overline{e}\%$, SD, MAD, MSE, R^2, CC, MPE, MAPE, e_{10} and δ) are defined in Table 5.

Table A5. Parameter estimates and fit statistics for black spruce.

	Model [2]				Model [4]			
Subregion	7 to 10	11	Others	Prov.	7 to 10	11	Others	Prov.
b_1	24.1828	31.0767	41.1666	32.3393	30.1593	30.3027	36.9007	31.4656
b_2	4.1123	3.9215	4.0910	3.9894	3.8672	3.7906	3.8145	3.8184
b_3	-1.5121	-1.3339	-1.2291	-1.3141	-1.2362	-1.2628	-1.1753	-1.2312
$\sigma^2_{u_1}$					86.0875	112.06	103.29	121.78
$\sigma_{u_1u_2}$					4.1653	5.7989	4.1131	5.8217
$\sigma^2_{u_2}$					0.2578	0.3506	0.2097	0.3361
σ^2					1.5639	1.3796	1.6199	1.4595
AIC					18567.1	29853.7	4958.2	53577.7
BIC					18599.0	29887.6	4982.7	53616.5
N	5168	8597	1321	15086	5168	8597	1321	15086
m	703	935	245	1883	703	935	245	1883
$\bar{e}$	-0.0004	-0.0064	-0.0140	-0.0083	-0.0031	-0.0035	-0.0064	-0.0037
$\bar{e}\%$	-0.0033	-0.0506	-0.0998	-0.0675	-0.0278	-0.0277	-0.0455	-0.0305
SD	1.9530	1.8462	2.0435	1.9844	1.1642	1.0985	1.1516	1.1226
MAD	1.5021	1.4491	1.5702	1.5396	0.8793	0.8260	0.8790	0.8467
MSE	3.8156	3.4092	4.1826	3.9383	1.3551	1.2065	1.3252	1.2603
R^2	0.7869	0.8303	0.8571	0.8121	0.9243	0.9399	0.9546	0.9399
CC	0.8807	0.9069	0.9225	0.8958	0.9601	0.9687	0.9765	0.9686
MPE	-3.30	-2.85	-3.22	-3.42	-1.71	-1.28	-1.51	-1.45
MAPE	14.68	12.75	12.83	14.04	8.89	7.32	7.49	7.89
e_{10}	55.82	50.27	49.36	53.84	33.42	24.01	23.62	27.02
δ	3.8141	3.4084	4.1762	3.9378	1.3554	1.2066	1.3262	1.2603

Note: Subregions are defined in Table 2, "Prov." denotes provincial (i.e., all subregions combined), *N* is the total number of observations, *m* is the number of plots, $\sigma^2_{u_1}$, $\sigma^2_{u_2}$ and $\sigma_{u_1u_2}$ are variances and covariance for the random parameters, and σ^2 is residual variance. The goodness-of-fit measures (AIC, BIC, $\bar{e}$, $\bar{e}\%$, SD, MAD, MSE, R^2, CC, MPE, MAPE, e_{10} and δ) are defined in Table 5.

Table A6. Parameter estimates and fit statistics for balsam fir.

	Model [2]					Model [4]				
Subregion	7 to 9	10	11	Others	Prov.	7 to 9	10	11	Others	Prov.
b_1	22.3112	34.2277	33.7327	35.0921	32.2788	19.8891	29.8705	38.7029	36.3526	32.5355
b_2	4.8034	4.6921	4.7704	4.8799	4.6623	4.7659	4.7651	4.5386	4.7330	4.6422
b_3	-1.8182	-1.4838	-1.5809	-1.5840	-1.5370	-1.8645	-1.5958	-1.4116	-1.5003	-1.5239
$\sigma^2_{u_1}$						17.8958	42.0409	171.90	53.0408	64.3154
$\sigma_{u_1u_2}$						1.1288	2.0883	6.3316	1.6889	2.5988
$\sigma^2_{u_2}$						0.08223	0.1282	0.2736	0.08218	0.1401
σ^2						1.4992	1.9746	1.4513	2.2005	1.7860
AIC						4637.7	16335.0	16453.3	9211.6	47027.8
BIC						4655.4	16363.6	16482.9	9236.3	47063.9
N	1369	4337	4748	2399	12853	1369	4337	4748	2399	12853
m	93	440	507	251	1291	93	440	507	251	1291
$\overline{e}$	0.0024	0.0072	0.0789	0.0218	0.0274	0.0062	0.0256	0.0199	0.0100	0.0225
$\overline{e}\%$	0.0249	0.0530	0.8649	0.1896	0.2460	0.0638	0.1875	0.2184	0.0874	0.2013
SD	1.4293	1.9832	1.7408	2.0718	1.9337	1.1877	1.3284	1.1412	1.4132	1.2684
MAD	1.0597	1.4274	1.1824	1.5081	1.36373	0.9106	0.9674	0.7783	1.0461	0.9112
MSE	2.0459	3.9349	3.0380	4.2966	3.7403	1.4095	1.7648	1.3024	1.9965	1.6092
R^2	0.8892	0.9233	0.9351	0.8944	0.9220	0.9235	0.9656	0.9722	0.9509	0.9665
CC	0.9414	0.9602	0.9669	0.9446	0.9597	0.9602	0.9825	0.9859	0.9747	0.9830
MPE	-2.08	-1.39	0.78	-1.72	-1.22	-1.51	0.21	-0.99	-1.21	-0.52
MAPE	11.98	10.94	14.58	14.03	13.24	10.66	8.45	10.62	10.55	9.97
e_{10}	45.43	44.09	56.02	55.94	51.87	40.03	30.30	40.04	41.52	37.28
δ	2.0429	3.9331	3.0367	4.2930	3.7398	1.4106	1.7652	1.3027	1.9973	1.6094

Note: Subregions are defined in Table 2, "Prov." denotes provincial (i.e., all subregions combined), *N* is the total number of observations, *m* is the number of plots, $\sigma^2_{u_1}$, $\sigma^2_{u_2}$ and $\sigma_{u_1u_2}$ are variances and covariance for the random parameters, and σ^2 is residual variance. The goodness-of-fit measures (AIC, BIC, $\overline{e}$, $\overline{e}\%$, SD, MAD, MSE, R^2, CC, MPE, MAPE, e_{10} and δ) are defined in Table 5.

Table A7. Parameter estimates and fit statistics for Douglas-fir, larch, jack pine, and all coniferous combined.

	Model [2]				Model [4]			
Species	FD	LT	PJ	All coniferous combined	FD	LT	PJ	All coniferous combined
b_1	31.3354	25.1942	29.6471	37.0245	21.6981	25.6393	20.2435	30.2546
b_2	3.5952	4.6404	3.7192	4.4310	4.2180	3.3332	3.0007	4.2341
b_3	-1.1138	-1.7314	-1.3144	-1.4401	-1.5347	-1.0836	-1.3632	-1.4934
$\sigma^2_{u_1}$					32.7207	68.5390	21.5268	60.8756
$\sigma_{u_1u_2}$					3.1756	2.7006	1.0492	3.5585
$\sigma^2_{u_2}$					0.3557	0.2629	0.1248	0.2911
σ^2					1.7938	0.7084	1.2315	2.6934
AIC					3047.8	4015.5	12158.0	529208.9
BIC					3060.5	4036.5	12181.7	529256.5
N	841	1378	3681	130089	841	1378	3681	130089
m	45	150	217	6625	45	150	217	6625
$\bar{e}$	-0.0082	0.0425	-0.0021	-0.0061	0.0017	0.0002	0.0001	0.0015
$\bar{e}\%$	-0.0598	0.4989	-0.0143	-0.0350	0.0123	0.0023	0.0005	0.0084
SD	1.9165	2.4700	1.9721	2.7271	1.2922	0.7867	1.0675	1.5742
MAD	1.5078	1.7728	1.5340	2.1090	0.9805	0.5349	0.8028	1.1679
MSE	3.6817	6.1117	3.8914	7.4374	1.6678	0.6184	1.1392	2.4782
R^2	0.7850	0.7848	0.8148	0.8533	0.9023	0.9782	0.9457	0.9511
CC	0.8787	0.8809	0.8978	0.9207	0.9483	0.9889	0.9720	0.9748
MPE	-3.01	-3.68	-2.35	-3.18	1.17	-1.02	-0.72	-1.19
MAPE	12.69	20.42	11.81	13.69	7.93	7.12	6.04	8.07
e_{10}	45.90	69.45	45.72	51.89	26.04	23.51	16.49	26.10
δ	3.6730	6.1028	3.8893	7.4373	1.6698	0.6189	1.1395	2.4782

Note: only provincial models are fitted for Douglas-fir (FD), larch (LT), jack pine (PJ), and all coniferous species combined, *N* is the total number of observations, *m* is the number of plots, $\sigma^2_{u_1}$, $\sigma^2_{u_2}$ and $\sigma_{u_1u_2}$ are variances and covariance for the random parameters, and σ^2 is residual variance. The goodness-of-fit measures (AIC, BIC, $\bar{e}$, $\bar{e}\%$, SD, MAD, MSE, R^2, CC, MPE, MAPE, e_{10} and δ) are defined in Table 5.

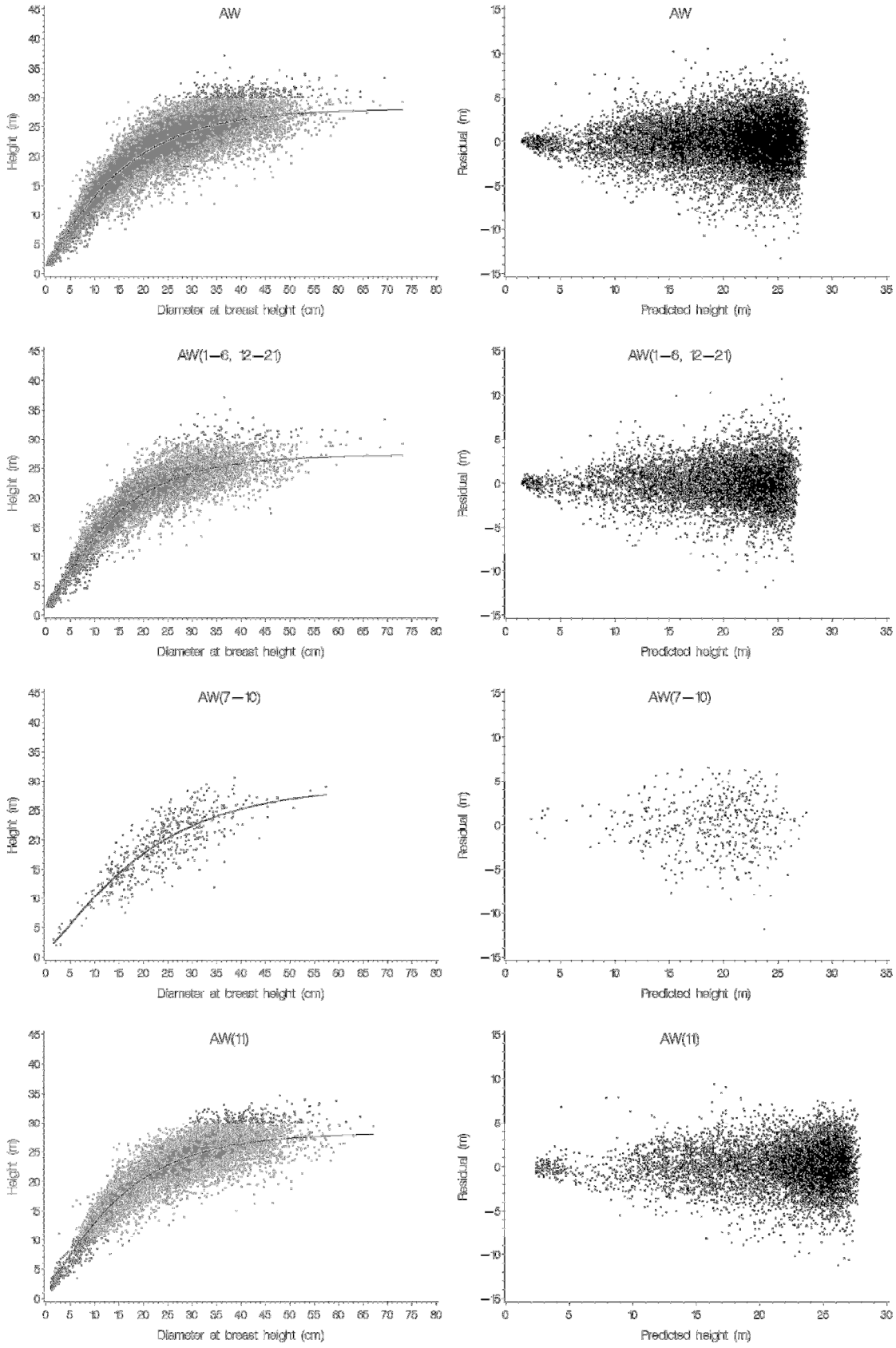

Figure A1. Fitted curves and residual plots from the base model [1] for aspen. The number in parentheses refers to the natural subregion defined in Table 2. The top two graphs are provincial.

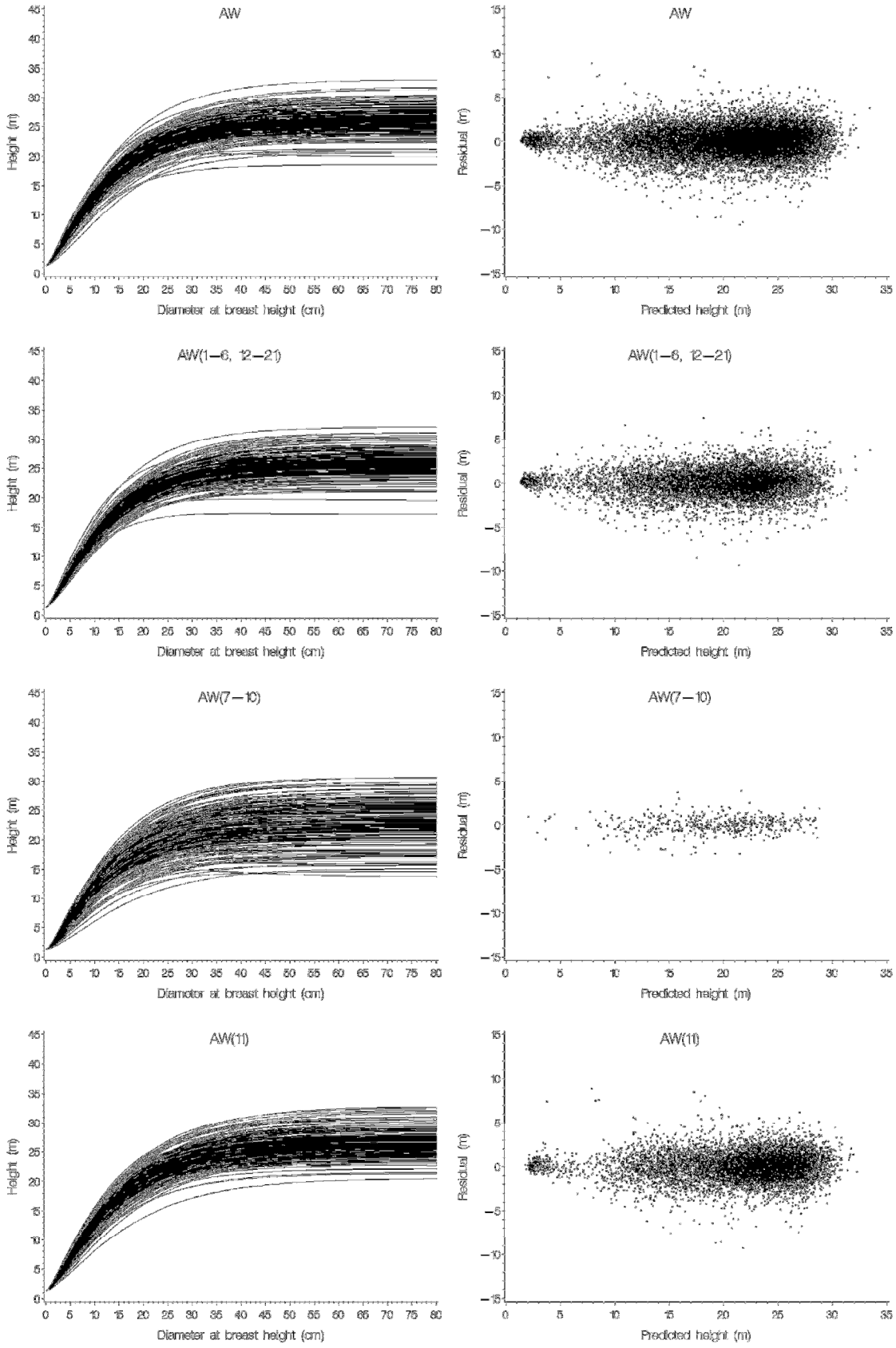

Figure A2. Spaghetti and residual plots from the mixed model [3] for aspen. The number in parentheses refers to the natural subregion defined in Table 2. Up to 200 randomly selected plots are shown in each spaghetti plot. The top two graphs are provincial.

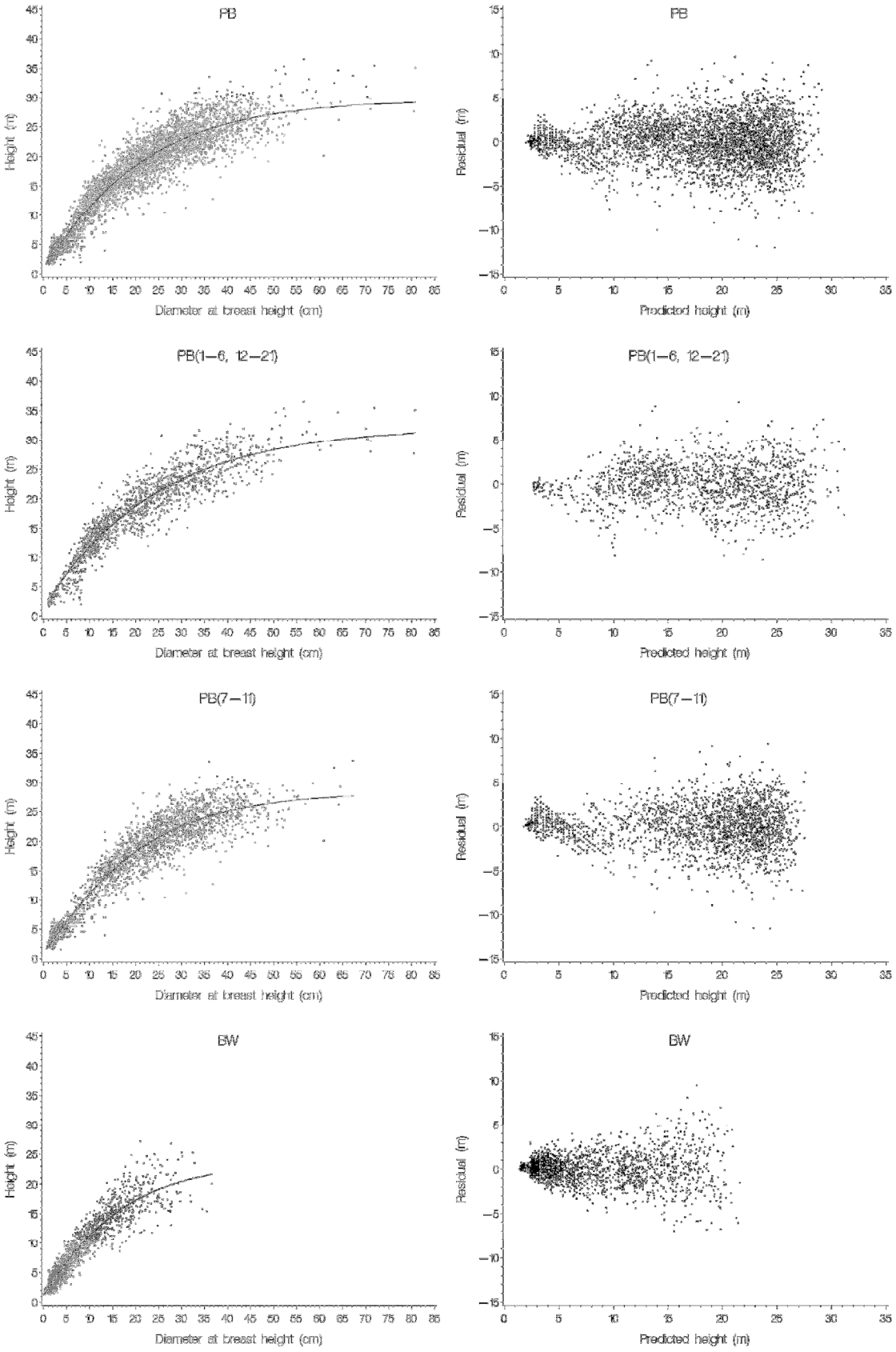

Figure A3. Fitted curves and residual plots from the base model [1] for balsam poplar and white birch. The number in parentheses refers to the natural subregion defined in Table 2. Otherwise, the graphs are provincial.

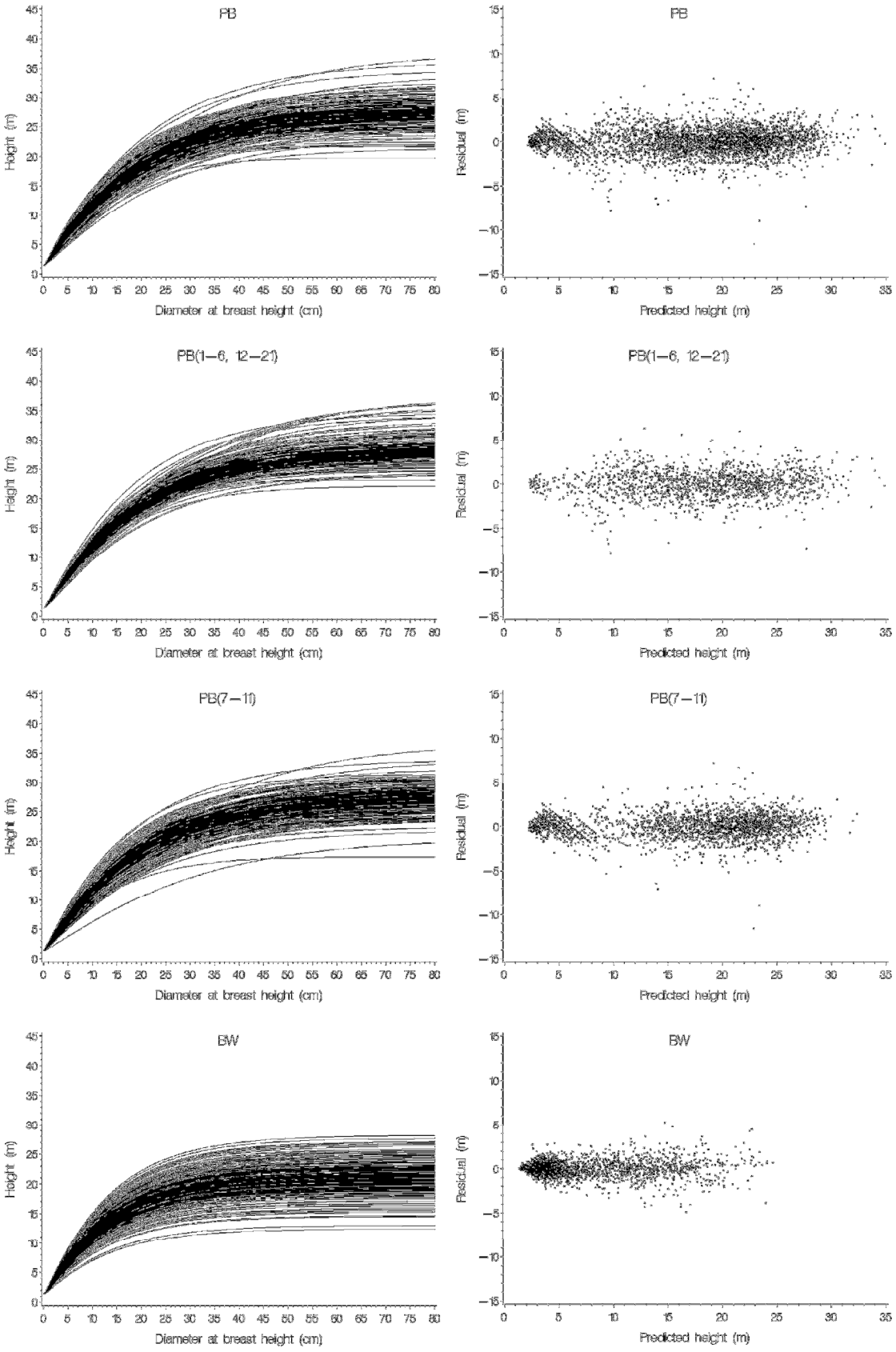

Figure A4. Spaghetti and residual plots from the mixed models [3] for PB and BW. The number in parentheses refers to the natural subregion defined in Table 2. Otherwise, the graphs are provincial. Up to 200 randomly selected plots are shown in each spaghetti plot.

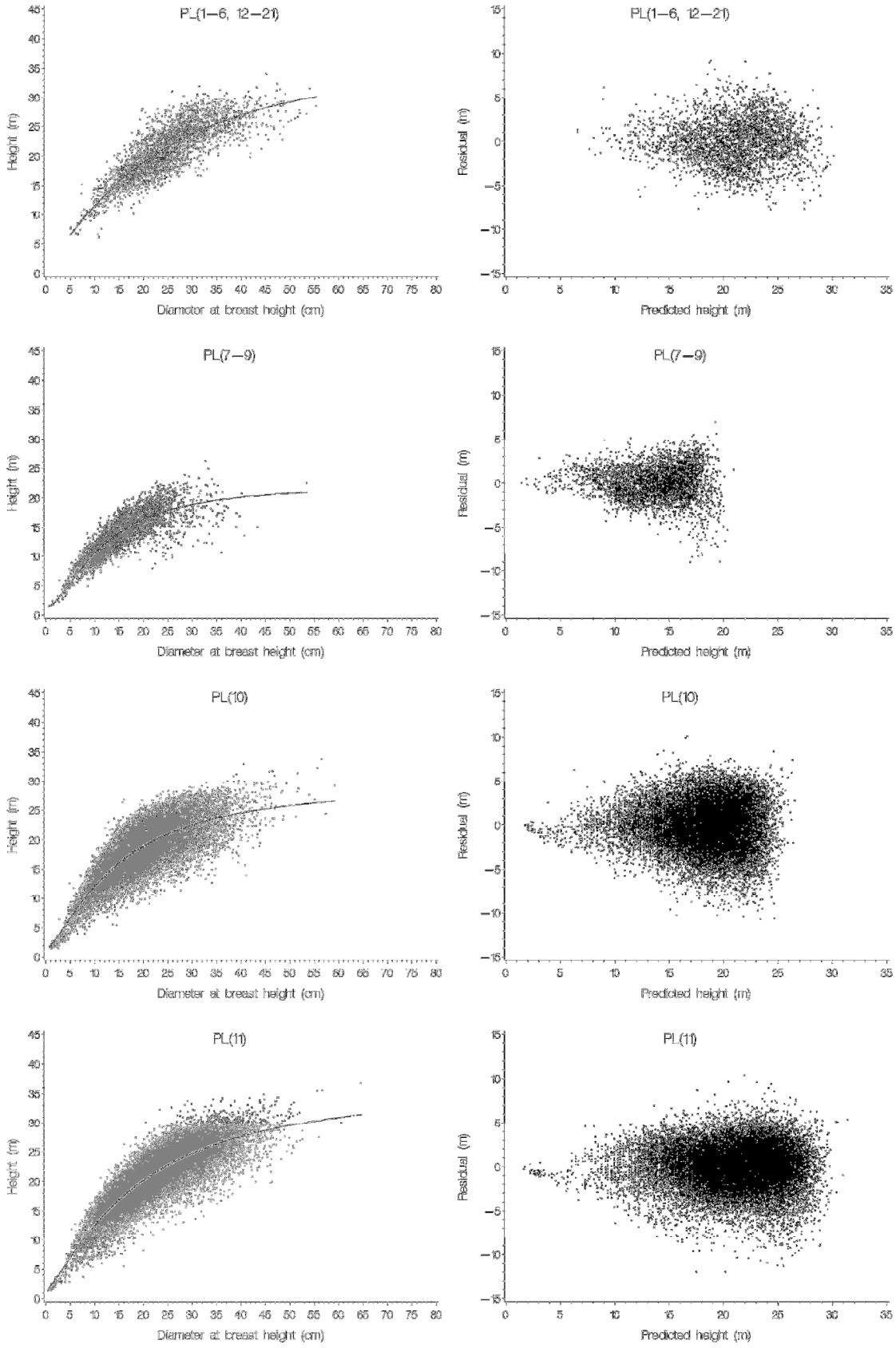

Figure A5. Fitted curves and residual plots from the base model [2] for lodgepole pine. The number in parentheses refers to the natural subregion defined in Table 2.

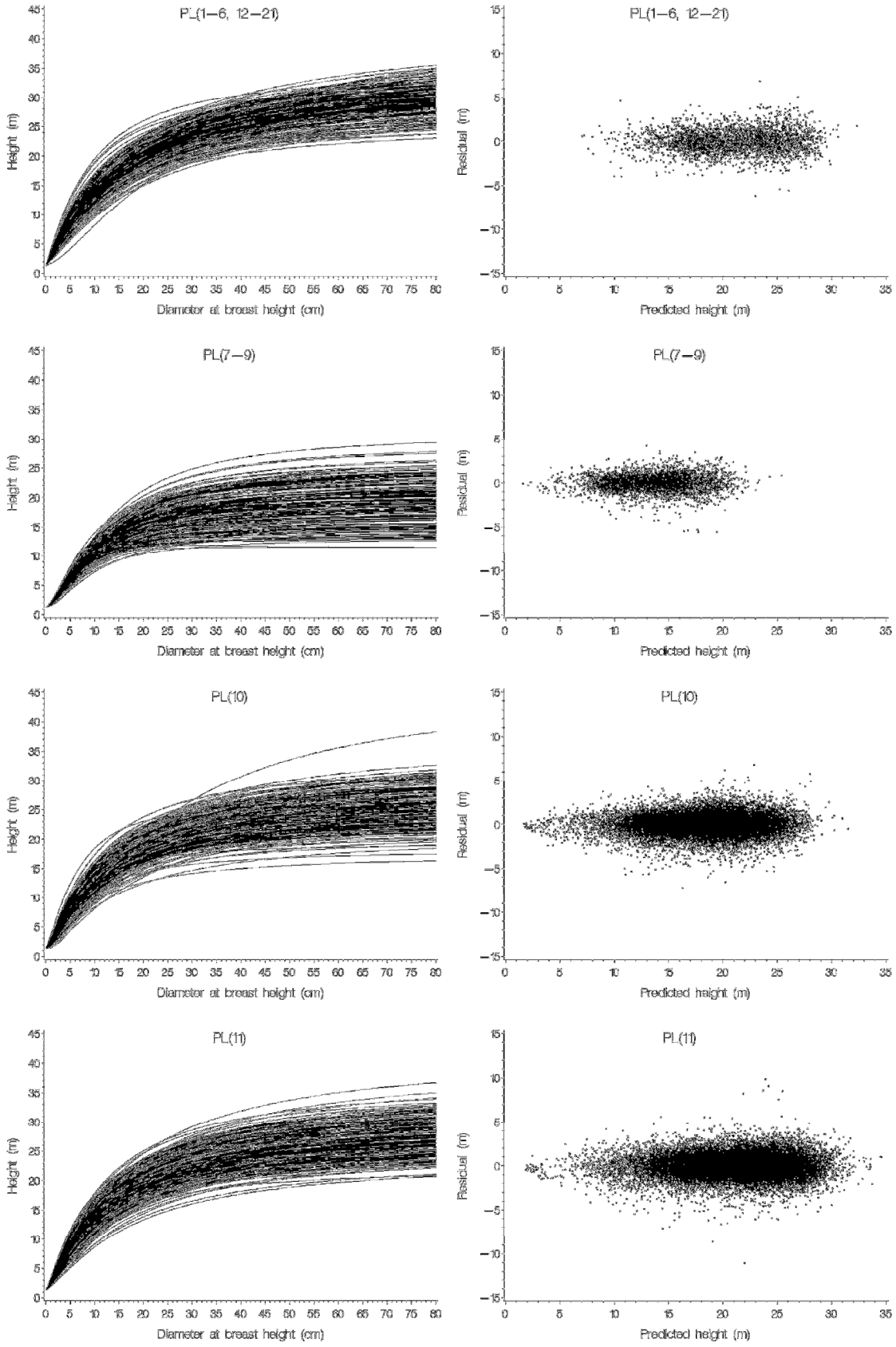

Figure A6. Spaghetti and residual plots from the mixed model [4] for lodgepole pine. The number in parentheses refers to the natural subregion defined in Table 2. Up to 200 randomly selected plots are shown in each spaghetti plot.

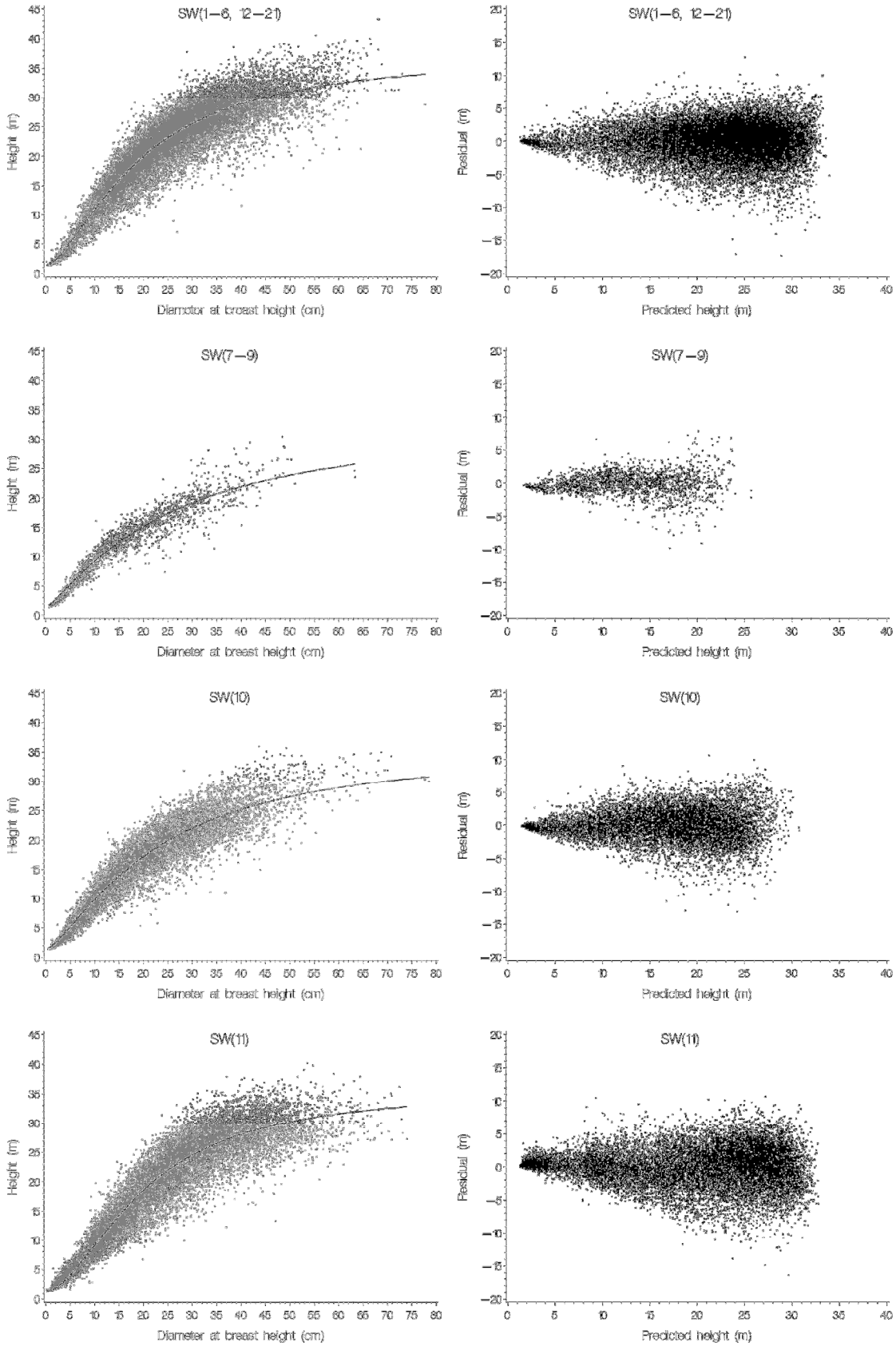

Figure A7. Fitted curves and residual plots from the base model [2] for white spruce. The number in parentheses refers to the natural subregion defined in Table 2.

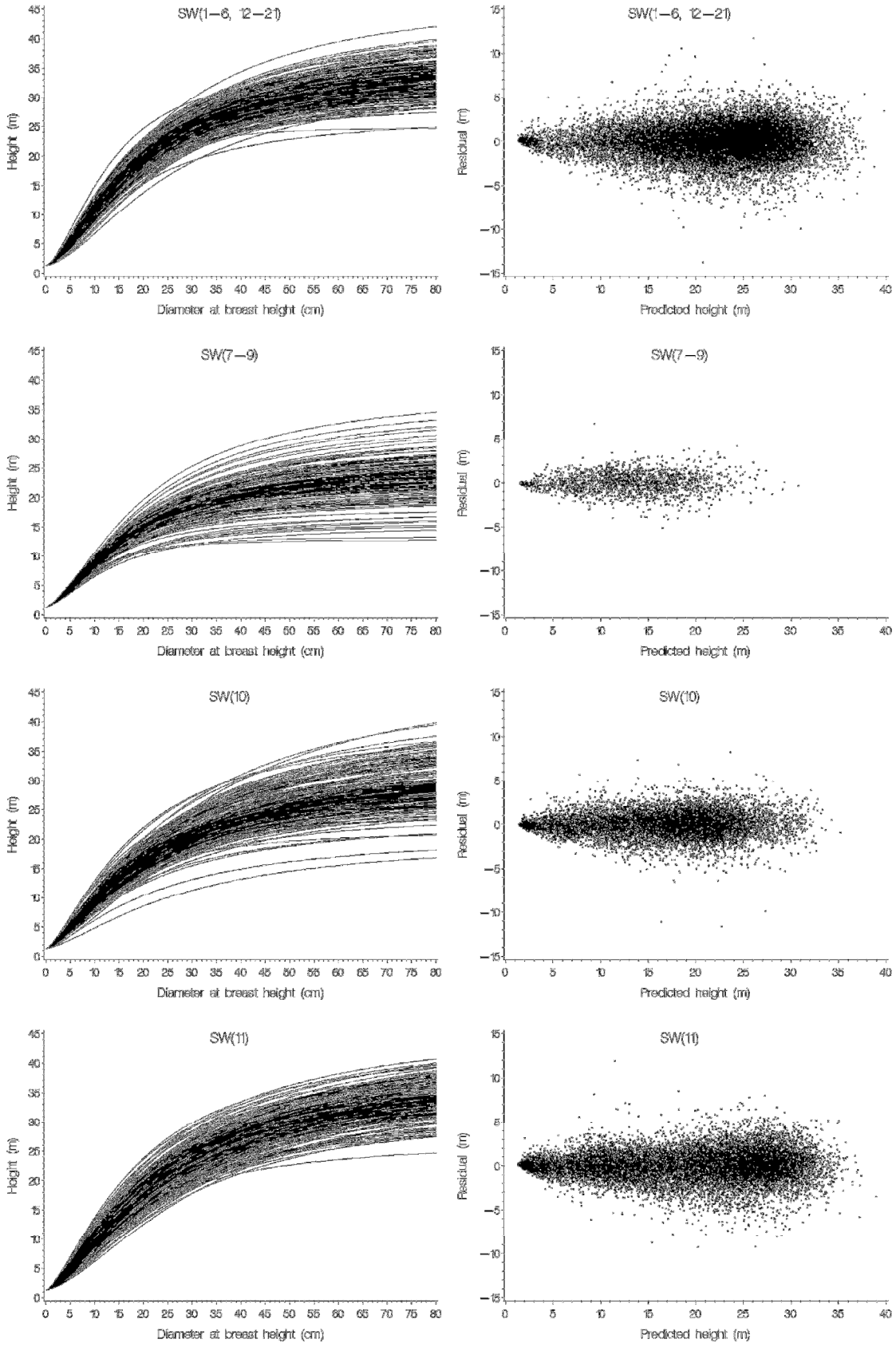

Figure A8. Spaghetti and residual plots from the mixed model [4] for white spruce. The number in parentheses refers to the natural subregion defined in Table 2. Up to 200 randomly selected plots are shown in each spaghetti plot.

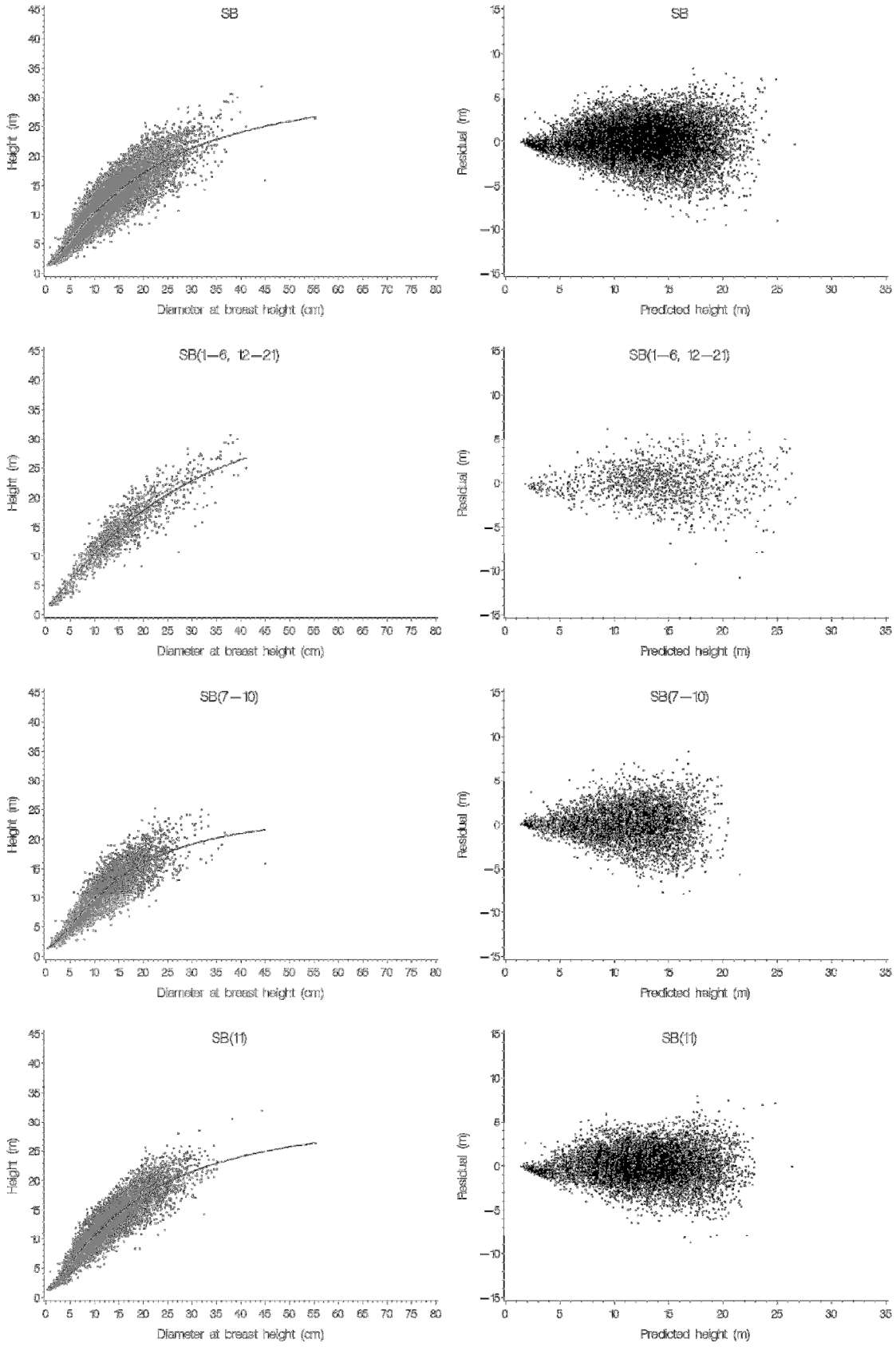

Figure A9. Fitted curves and residual plots from the base model [2] for black spruce. The number in parentheses refers to the natural subregion defined in Table 2. The top two graphs are provincial.

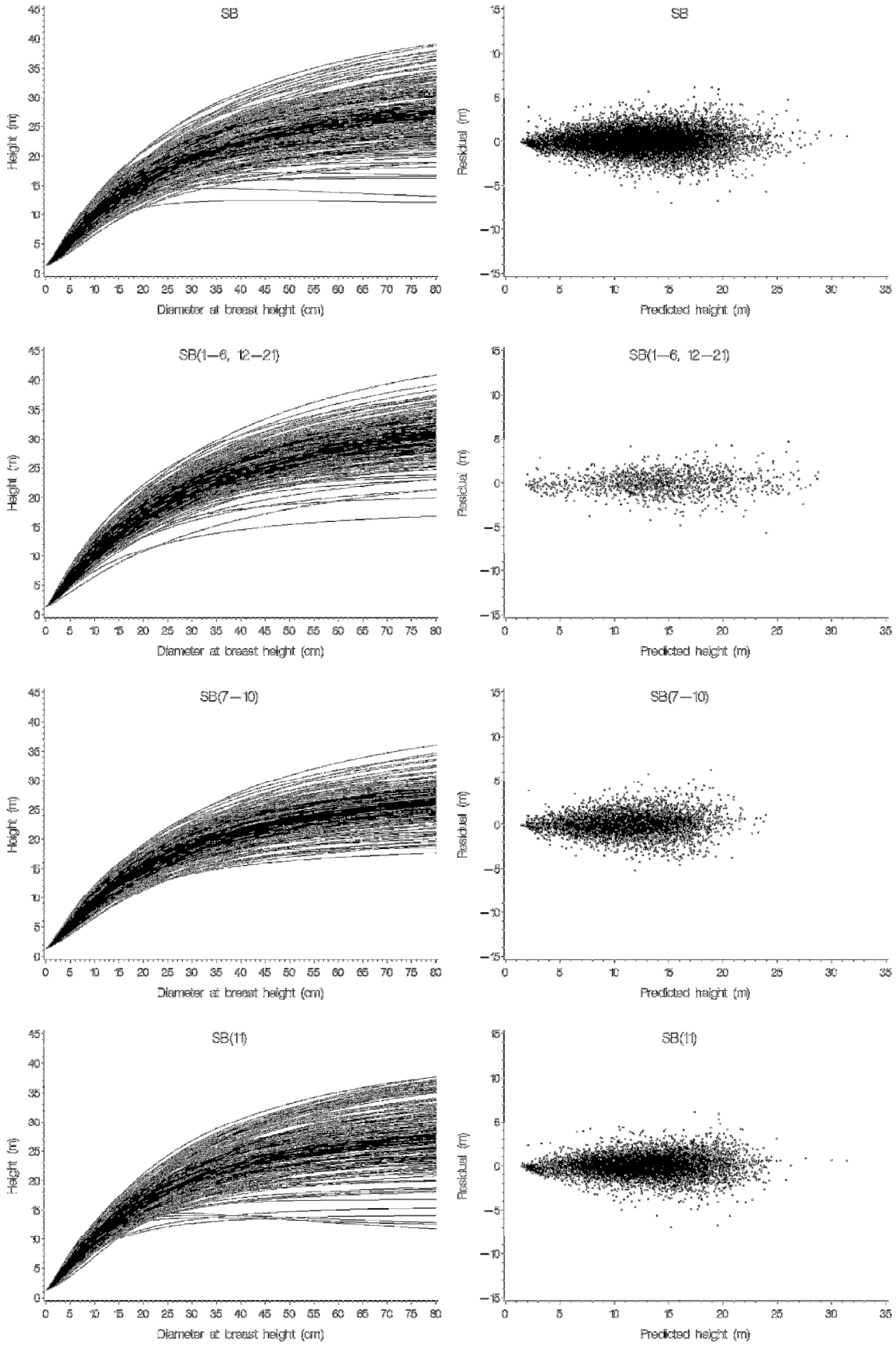

Figure A10. Spaghetti and residual plots from the mixed model [4] for black spruce. The number in parentheses refers to the natural subregion defined in Table 2. The top two graphs are provincial. Up to 200 randomly selected plots are shown in each spaghetti plot.

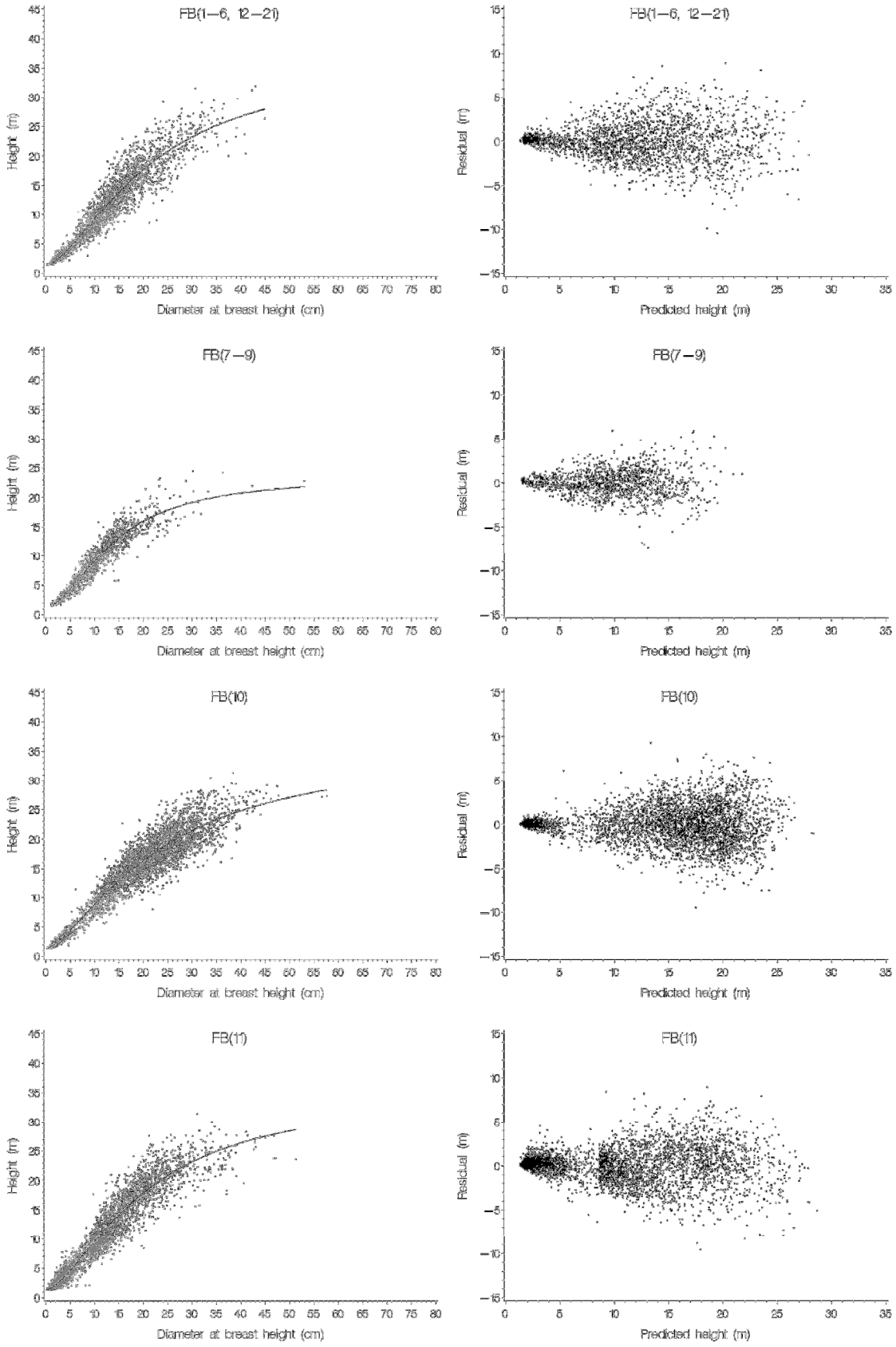

Figure A11. Fitted curves and residual plots from the base model [2] for balsam fir. The number in parentheses refers to the natural subregion defined in Table 2.

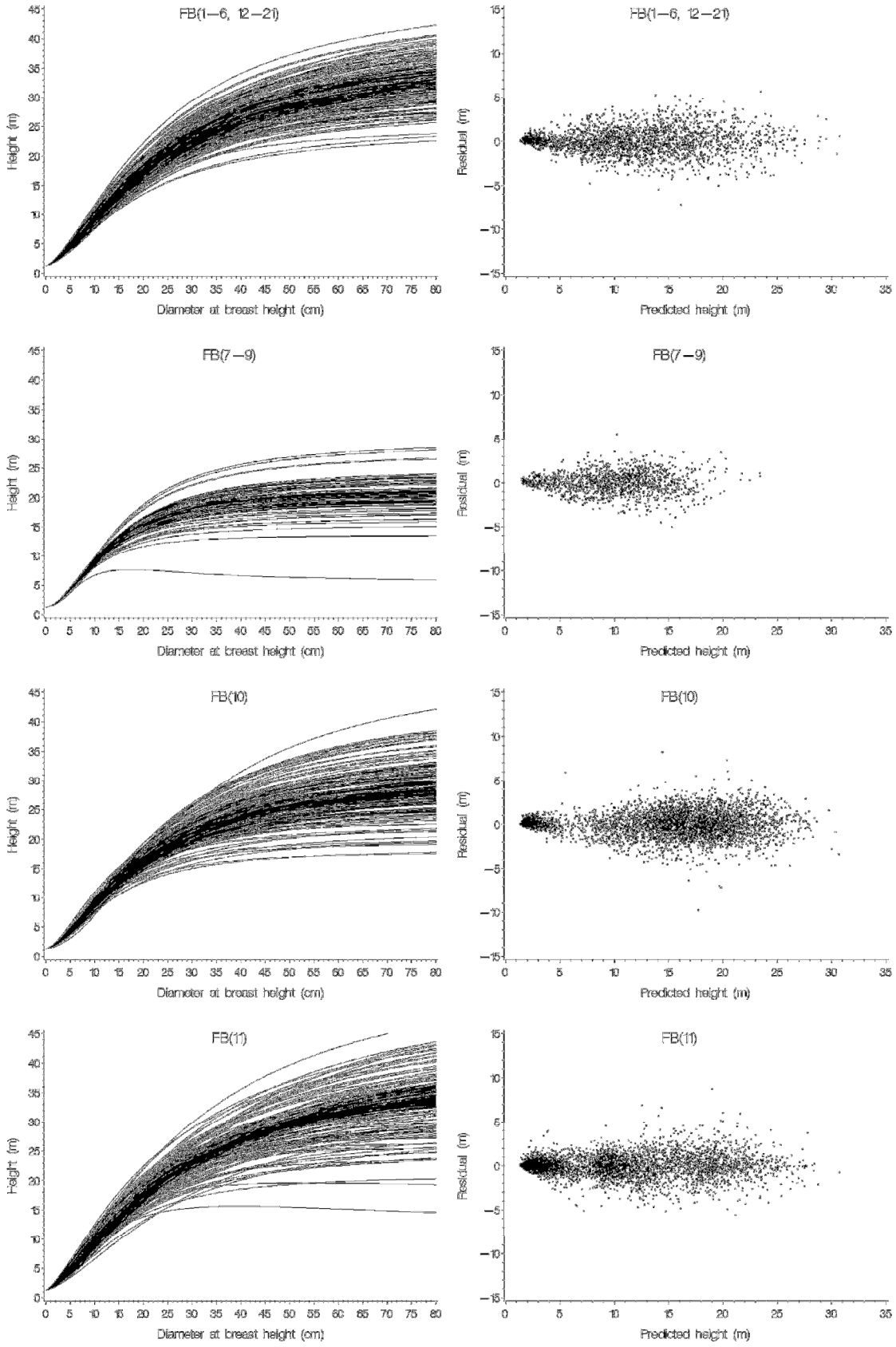

Figure A12. Spaghetti and residual plots from the mixed model [4] for balsam fir. The number in parentheses refers to the natural subregion defined in Table 2. Up to 200 randomly selected plots are shown in each spaghetti plot.

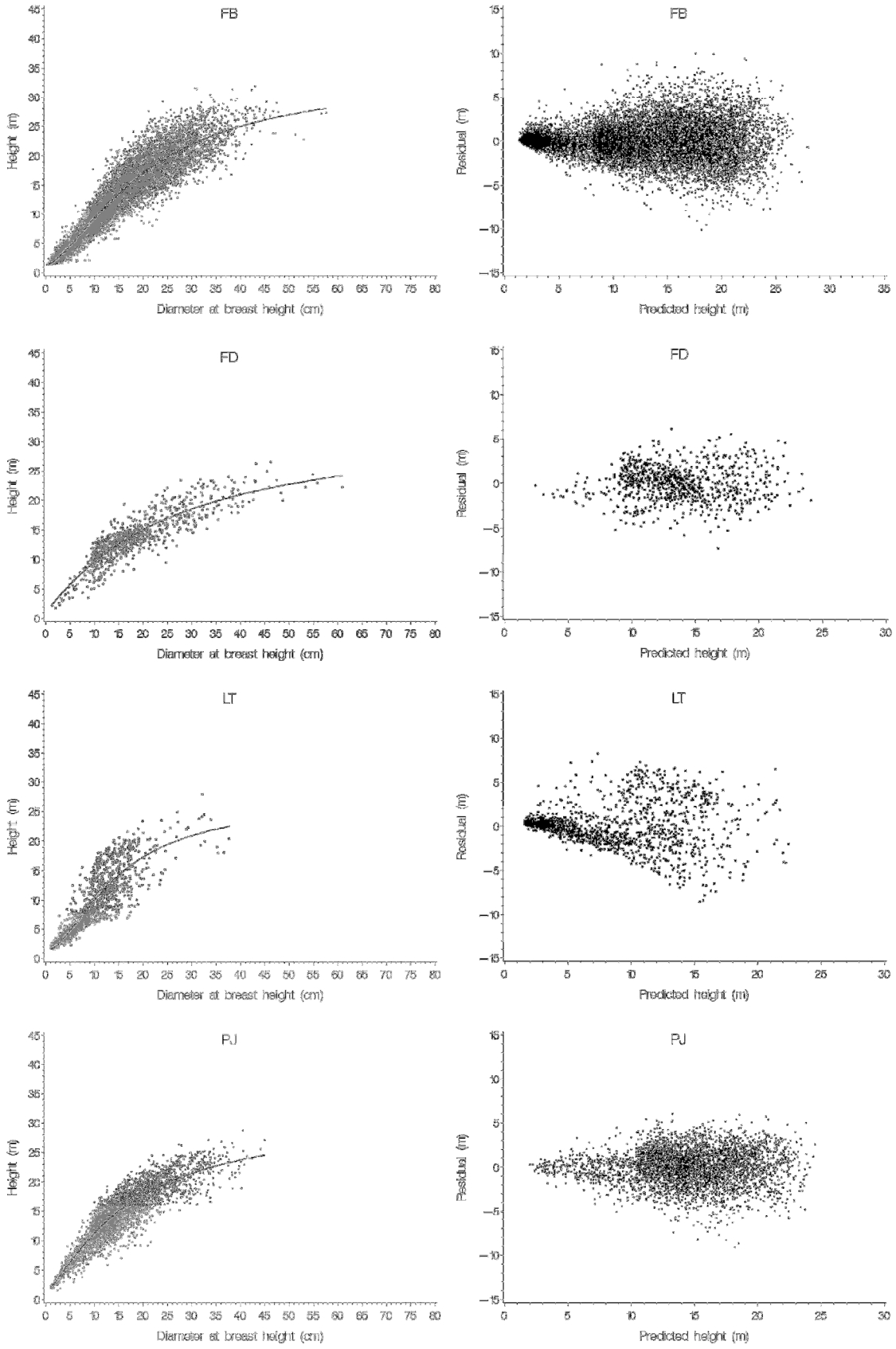

Figure A13. Fitted curves and residual plots from the base model [2] for balsam fir, Douglas-fir, larch, and jack pine. All are provincial.

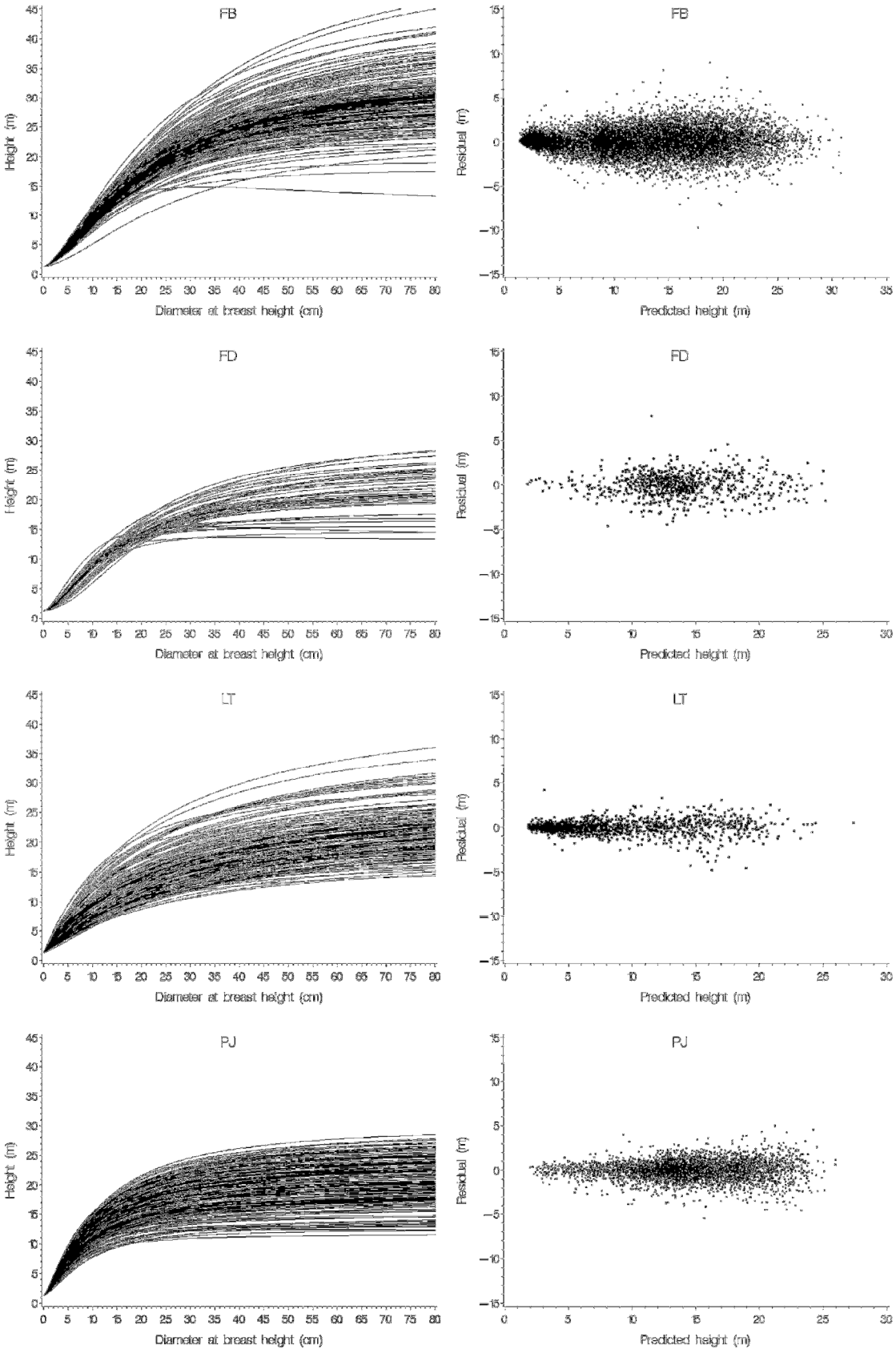

Figure A14. Spaghetti and residual plots from the mixed model [4] for balsam fir, Douglas-fir, larch, and jack pine. All are provincial. Up to 200 randomly selected plots are shown in each spaghetti plot.

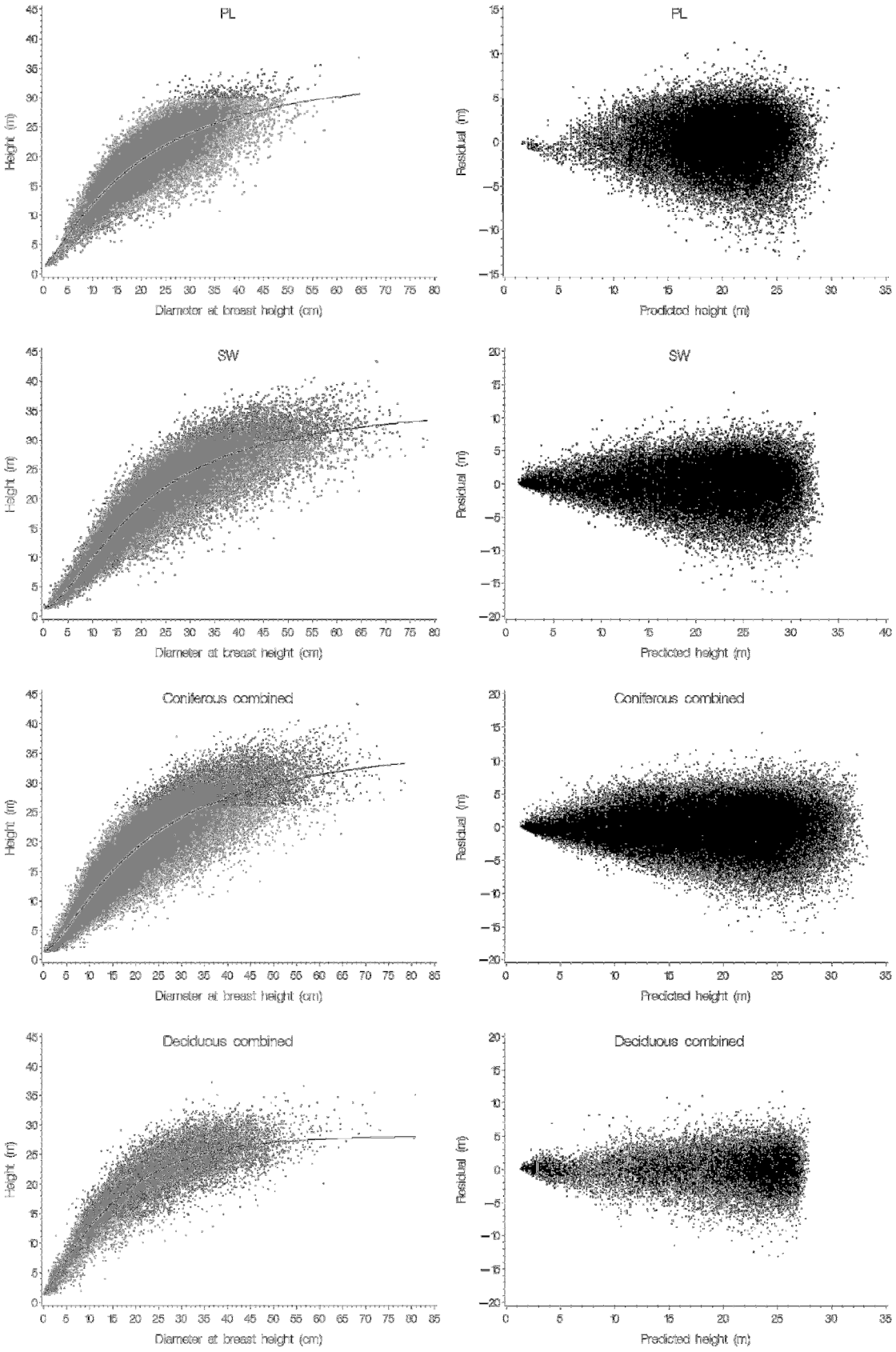

Figure A15. Fitted curves and residual plots for lodgepole pine (model [2]), white spruce (model [2]), coniferous combined (model [2]), and deciduous combined (model [1]). All are provincial.

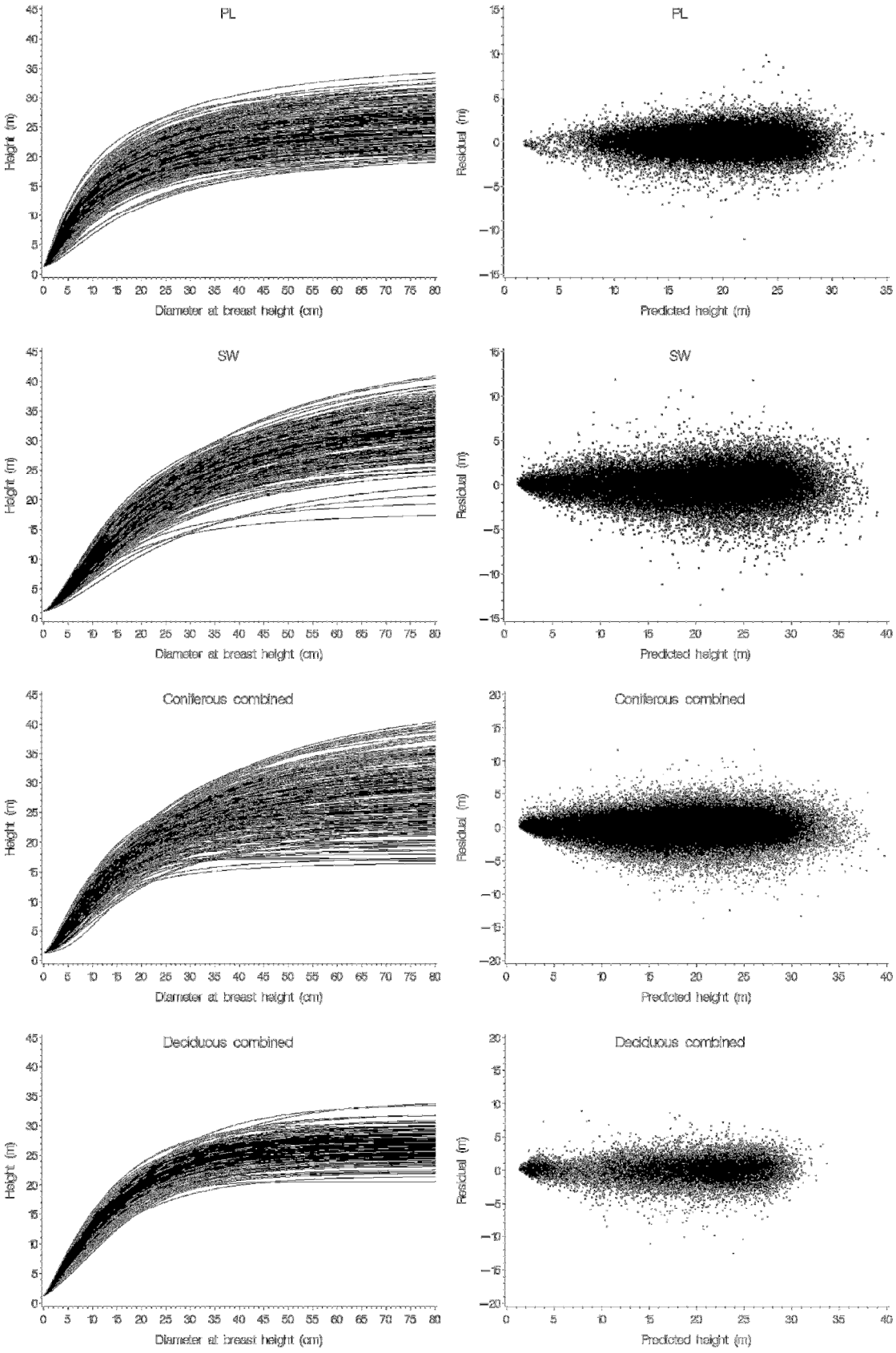

Figure A16. Spaghetti and residual plots for lodgepole pine (model [4]), white spruce (model [4]), coniferous combined (model [4]), and deciduous combined (model [3]). All are provincial. Up to 200 randomly selected plots are shown in each spaghetti plot.

Appendix 2. Summary Statistics from Different Types of Predictions (part 1 of 3).

	Type	Goodness-of-fit measure										
		$\bar{e}$	$\bar{e}$%	SD	MAD	MSE	R^2	CC	MPE	MAPE	e_{10}	δ
AW	B-U	-0.0049	-0.0238	2.4093	1.8525	5.8055	0.8400	0.9128	-1.97	10.15	38.04	5.8048
	B-A	0	0	1.5623	1.1538	2.4409	0.9327	0.9657	-0.82	6.49	19.55	2.4406
	M-U	0.0017	0.0081	1.4604	1.0980	2.1325	0.9412	0.9694	-0.82	6.30	17.32	2.1327
	M-A	0	0	1.4297	1.0438	2.0440	0.9437	0.9710	-0.70	5.90	16.32	2.0442
AW1	B-U	0.0058	0.0306	2.9245	2.3299	8.5880	0.6994	0.8238	-2.76	13.9	54.49	8.5529
	B-A	0	0	1.4029	0.9765	1.9761	0.9308	0.9665	-0.04	5.96	18.57	1.9680
	M-U	0.0012	0.0063	0.9614	0.7136	0.9224	0.9675	0.9831	-0.68	4.67	9.39	0.9243
	M-A	0	0	0.9373	0.6515	0.8767	0.9691	0.9844	-0.42	4.22	9.18	0.8785
AW2	B-U	-0.0036	-0.0166	2.4038	1.8694	5.7798	0.8324	0.9083	-1.69	9.59	36.54	5.7784
	B-A	0	0	1.5989	1.1875	2.5570	0.9259	0.9621	-0.71	6.27	18.25	2.5563
	M-U	0.0016	0.0076	1.5140	1.1393	2.2918	0.9335	0.9652	-0.77	6.11	16.54	2.2921
	M-A	0	0	1.4849	1.0887	2.2046	0.9360	0.9670	-0.65	5.77	15.79	2.2049
AW3	B-U	-0.0034	-0.0174	2.2880	1.7482	5.2364	0.8562	0.9224	-1.87	10.04	37.63	5.2351
	B-A	0	0	1.5287	1.1263	2.3376	0.9358	0.9673	-0.86	6.72	20.76	2.3370
	M-U	0.0040	0.0208	1.4279	1.0766	2.0386	0.9440	0.9709	-0.76	6.58	18.47	2.0389
	M-A	0	0	1.3933	1.0192	1.9411	0.9467	0.9726	-0.68	6.12	17.26	1.9413
PB	B-U	-0.0023	-0.0137	2.3911	1.8250	5.7201	0.8992	0.9469	-2.98	13.22	47.94	5.7172
	B-A	0	0	1.4919	1.0477	2.2268	0.9608	0.9804	-0.96	8.29	25.81	2.2257
	M-U	0.0057	0.0335	1.3852	1.0230	1.9184	0.9662	0.9827	-1.53	8.41	24.58	1.9189
	M-A	0	0	1.3382	0.9322	1.7904	0.9684	0.9840	-1.18	7.63	22.41	1.7908
PB1	B-U	0.0313	0.1905	2.3211	1.7593	5.3930	0.9123	0.9545	-1.48	13.21	49.25	5.3885
	B-A	0	0	1.4533	1.0183	2.1137	0.9656	0.9830	-0.08	8.45	27.11	2.1120
	M-U	0.0118	0.072	1.3134	0.9609	1.7246	0.9719	0.9857	-1.09	8.24	24.75	1.7253
	M-A	0	0	1.2814	0.8915	1.6414	0.9733	0.9865	-0.84	7.58	23.22	1.6421
PB2	B-U	-0.0224	-0.1244	2.4120	1.8738	5.8258	0.8779	0.9344	-4.05	13.22	44.37	5.8180
	B-A	0	0	1.5731	1.1185	2.4778	0.9481	0.9736	-2.06	8.49	25.12	2.4745
	M-U	-0.0065	-0.0362	1.5141	1.1379	2.2911	0.9519	0.9749	-2.33	8.76	24.18	2.2926
	M-A	0	0	1.4316	1.0027	2.0480	0.9570	0.9780	-1.79	7.71	21.59	2.0493
BW	B-U	0.0253	0.3288	1.6307	1.1691	2.6623	0.9034	0.9496	-3.49	18.00	62.40	2.6598
	B-A	0	0	0.9739	0.6452	0.9494	0.9656	0.9828	-0.70	10.28	39.03	0.9485
	M-U	0.0066	0.0858	0.8654	0.6188	0.7487	0.9728	0.9861	-2.11	10.71	38.60	0.7490
	M-A	0	0	0.8329	0.5557	0.6934	0.9748	0.9873	-1.24	9.23	33.62	0.6938
Dec	B-U	-0.0007	-0.0035	2.4448	1.8636	5.9774	0.8872	0.9402	-2.43	11.97	43.74	5.9768
	B-A	0	0	1.6418	1.2044	2.6958	0.9491	0.9742	-0.95	8.05	26.54	2.6955
	M-U	0.0084	0.0447	1.5209	1.1343	2.3130	0.9563	0.9776	-0.98	7.81	24.05	2.3131
	M-A	0	0	1.4962	1.0868	2.2385	0.9578	0.9785	-0.91	7.38	22.75	2.2386
Con	B-U	-0.0061	-0.0350	2.7271	2.1090	7.4374	0.8533	0.9207	-3.18	13.69	51.89	7.4373
	B-A	0	0	1.8237	1.3905	3.3260	0.9344	0.9667	-1.07	9.34	33.99	3.3259
	M-U	0.0015	0.0084	1.5742	1.1679	2.4782	0.9511	0.9748	-1.19	8.07	26.10	2.4782
	M-A	0	0	1.5732	1.1656	2.4749	0.9512	0.9749	-1.18	8.02	25.86	2.4749

Note: Type refers to the prediction types: B-U (base model, unadjusted), B-A (base model, adjusted), M-U (mixed model, unadjusted), and M-A (mixed model, adjusted). The goodness-of-fit measures are defined in Table 5. The base and mixed models are given in [1] to [4] for different species. "Dec" denotes all deciduous species combined. "Con" denotes all coniferous species combined. The number after the species code indicates the group of natural subregions. Otherwise, it is provincial (e.g., AW=aspen provincial). AW1=subregions 7-10, AW2=subregion 11, AW3=other subregions, PB1=subregions 7-11, and PB2=other subregions. Subregions are defined in Table 2.

Appendix 2. Summary Statistics from Different Types of Predictions (part 2 of 3).

	Type	Goodness-of-fit measure										
		$\bar{e}$	$\bar{e}$%	SD	MAD	MSE	R^2	CC	MPE	MAPE	e_{10}	δ
PL	B-U	-0.0013	-0.0066	2.6058	2.0542	6.7906	0.7175	0.8353	-2.23	11.64	44.96	6.7903
	B-A	0	0	1.4534	1.1382	2.1125	0.9121	0.9575	-0.09	6.30	19.49	2.1124
	M-U	-0.0008	-0.0039	1.1487	0.8724	1.3195	0.9451	0.9716	-0.55	4.91	10.16	1.3196
	M-A	0	0	1.1470	0.8692	1.3156	0.9453	0.9718	-0.49	4.88	10.06	1.3157
PL1	B-U	0.0028	0.0133	2.4822	1.9717	6.1653	0.6955	0.8210	-1.44	9.78	38.93	6.1612
	B-A	0	0	1.5388	1.1945	2.3694	0.8830	0.9429	-0.16	5.91	17.51	2.3678
	M-U	0.0002	0.0008	1.2656	0.9784	1.6013	0.9208	0.9581	-0.48	4.88	10.56	1.6018
	M-A	0	0	1.2572	0.9628	1.5799	0.9219	0.9595	-0.39	4.80	10.26	1.5805
PL2	B-U	0.0074	0.0533	1.8436	1.4332	3.4008	0.7421	0.8532	-1.66	10.96	44.72	3.3990
	B-A	0	0	1.1079	0.8529	1.2281	0.9069	0.9546	-0.08	6.73	21.90	1.2274
	M-U	-0.0006	-0.0041	0.9438	0.7091	0.8906	0.9324	0.9648	-0.66	5.58	13.73	0.8908
	M-A	0	0	0.9427	0.7075	0.8885	0.9326	0.9650	-0.62	5.55	13.79	0.8887
PL3	B-U	-0.0020	-0.0108	2.5894	2.0678	6.7058	0.6496	0.7871	-2.41	12.26	48.80	6.7051
	B-A	0	0	1.3043	1.0176	1.7013	0.9111	0.9561	-0.26	6.05	17.39	1.7011
	M-U	-0.0007	-0.0038	1.0999	0.8361	1.2098	0.9368	0.9671	-0.57	5.01	10.78	1.2099
	M-A	0	0	1.0985	0.8336	1.2067	0.9369	0.9674	-0.52	4.98	10.71	1.2067
PL4	B-U	-0.0005	-0.0025	2.3057	1.7962	5.3165	0.7538	0.8596	-1.55	9.40	34.77	5.3161
	B-A	0	0	1.4605	1.1431	2.1333	0.9012	0.9521	-0.12	5.82	16.46	2.1331
	M-U	-0.0007	-0.0036	1.2014	0.9156	1.4433	0.9332	0.9651	-0.51	4.73	9.12	1.4433
	M-A	0	0	1.1984	0.9107	1.4362	0.9335	0.9656	-0.45	4.69	8.93	1.4362
SW	B-U	0.0074	0.0377	2.9481	2.2681	8.6915	0.8606	0.9253	-2.38	13.02	49.54	8.6911
	B-A	0	0	1.8390	1.3826	3.3822	0.9458	0.9728	-0.16	8.25	29.03	3.3821
	M-U	-0.0003	-0.0014	1.6423	1.2319	2.6971	0.9567	0.9778	-1.12	7.48	24.06	2.6972
	M-A	0	0	1.6341	1.2162	2.6701	0.9572	0.9781	-0.92	7.27	23.37	2.6702
SW1	B-U	-0.0013	-0.0060	2.6301	2.0055	6.9182	0.8537	0.9210	-1.90	10.33	37.58	6.9175
	B-A	0	0	1.9111	1.4517	3.6528	0.9227	0.9601	-0.88	7.51	24.07	3.6524
	M-U	-0.0007	-0.0032	1.7554	1.3321	3.0814	0.9348	0.9660	-1.00	6.95	21.09	3.0815
	M-A	0	0	1.7427	1.3178	3.0367	0.9358	0.9667	-0.88	6.83	20.59	3.0369
SW2	B-U	-0.0241	-0.1887	1.7383	1.2528	3.0256	0.9012	0.9475	-3.26	11.49	42.64	3.0222
	B-A	0	0	1.2149	0.9444	1.4777	0.9517	0.9752	-2.17	9.12	33.98	1.4761
	M-U	-0.0103	-0.0808	1.1361	0.8683	1.2900	0.9578	0.9783	-1.64	8.05	29.11	1.2907
	M-A	0	0	1.1297	0.8604	1.2754	0.9583	0.9786	-1.46	7.93	28.77	1.2761
SW3	B-U	-0.0182	-0.1101	2.4395	1.8550	5.9529	0.8770	0.9341	-3.32	12.85	49.27	5.9516
	B-A	0	0	1.5199	1.1343	2.3105	0.9523	0.9759	-0.92	8.08	27.88	2.3100
	M-U	-0.0032	-0.0194	1.3748	1.0315	1.8899	0.9609	0.9800	-1.28	7.51	24.47	1.8901
	M-A	0	0	1.3667	1.0157	1.8676	0.9614	0.9803	-1.06	7.27	23.44	1.8678
SW4	B-U	0.0446	0.2326	2.9029	2.2101	8.4300	0.8897	0.9422	-1.05	13.44	50.98	8.4288
	B-A	0	0	1.9071	1.4236	3.6377	0.9524	0.9763	0.43	9.18	33.64	3.6372
	M-U	0.0033	0.0170	1.6937	1.2644	2.8686	0.9624	0.9808	-1.14	8.10	27.60	2.8688
	M-A	0	0	1.6812	1.2381	2.8263	0.9630	0.9812	-0.86	7.75	26.39	2.8265
PJ	B-U	-0.0021	-0.0143	1.9721	1.5340	3.8914	0.8148	0.8978	-2.35	11.81	45.72	3.8893
	B-A	0	0	1.3262	1.0210	1.7599	0.9163	0.9591	-0.09	7.63	27.28	1.7589
	M-U	0.0001	0.0005	1.0675	0.8028	1.1392	0.9457	0.9720	-0.72	6.04	16.49	1.1395
	M-A	0	0	1.0662	0.8003	1.1365	0.9459	0.9722	-0.62	5.99	16.35	1.1368

Note: B-U (base model, unadjusted), B-A (base model, adjusted), M-U (mixed model, unadjusted), and M-A (mixed model, adjusted). The goodness-of-fit measures are defined in Table 5. The base and mixed models are given in [1] to [4] for different species. The number after the species code indicates the group of natural subregions. Otherwise, it is provincial (e.g., PL=PL provincial). PL1=other subregions, PL2=subregions 7-9, PL3=subregion 10, PL4=subregion 11, SW1=other subregions, SW2=subregions 7-9, SW3=subregion 10, and SW4=subregion 11. Subregions are defined in Table 2.

Appendix 2. Summary Statistics from Different Types of Predictions (part 3 of 3).

	Type	Goodness-of-fit measure										
		$\bar{e}$	$\bar{e}$%	SD	MAD	MSE	R^2	CC	MPE	MAPE	e_{10}	δ
SB	B-U	-0.0083	-0.0675	1.9844	1.5396	3.9383	0.8121	0.8958	-3.42	14.04	53.84	3.9378
	B-A	0	0	1.2262	0.9176	1.5037	0.9283	0.9638	-0.73	8.26	30.63	1.5035
	M-U	-0.0037	-0.0305	1.1226	0.8466	1.2603	0.9399	0.9686	-1.45	7.89	27.02	1.2603
	M-A	0	0	1.1115	0.8228	1.2353	0.9411	0.9696	-1.13	7.53	25.59	1.2354
SB1	B-U	-0.0140	-0.0998	2.0435	1.5702	4.1826	0.8571	0.9225	-3.22	12.83	49.36	4.1762
	B-A	0	0	1.2614	0.9327	1.5936	0.9455	0.9729	-0.50	7.43	26.95	1.5912
	M-U	-0.0064	-0.0455	1.1516	0.8790	1.3252	0.9546	0.9765	-1.51	7.49	23.62	1.3262
	M-A	0	0	1.1289	0.8311	1.2733	0.9564	0.9778	-0.95	6.88	21.35	1.2743
SB2	B-U	-0.0004	-0.0033	1.9530	1.5021	3.8156	0.7869	0.8807	-3.30	14.68	55.82	3.8141
	B-A	0	0	1.2348	0.9182	1.5253	0.9148	0.9565	-0.83	9.14	34.77	1.5247
	M-U	-0.0031	-0.0278	1.1642	0.8793	1.3551	0.9243	0.9601	-1.71	8.98	33.42	1.3554
	M-A	0	0	1.1494	0.8515	1.3209	0.9262	0.9616	-1.39	8.58	31.89	1.3211
SB3	B-U	-0.0064	-0.0506	1.8462	1.4491	3.4092	0.8303	0.9069	-2.85	12.75	50.27	3.4084
	B-A	0	0	1.2146	0.9143	1.4757	0.9265	0.9629	-0.64	7.89	28.60	1.4753
	M-U	-0.0035	-0.0277	1.0985	0.8260	1.2065	0.9399	0.9687	-1.28	7.32	24.01	1.2066
	M-A	0	0	1.0883	0.8057	1.1842	0.9410	0.9696	-0.97	7.00	22.43	1.1844
FB	B-U	0.0274	0.2460	1.9337	1.3637	3.7403	0.9220	0.9597	-1.22	13.24	51.87	3.7398
	B-A	0	0	1.3650	0.9630	1.8634	0.9612	0.9802	-1.07	10.16	38.40	1.8631
	M-U	0.0225	0.2013	1.2684	0.9112	1.6092	0.9665	0.9830	-0.52	9.97	37.28	1.6094
	M-A	0	0	1.2601	0.8895	1.5877	0.9669	0.9832	-0.89	9.51	35.36	1.5878
FB1	B-U	0.0218	0.1896	2.0718	1.5081	4.2966	0.8944	0.9446	-1.72	14.03	55.94	4.2930
	B-A	0	0	1.4888	1.0809	2.2183	0.9455	0.9725	-0.60	10.66	43.02	2.2165
	M-U	0.01	0.0874	1.4132	1.0461	1.9965	0.9509	0.9747	-1.21	10.55	41.52	1.9973
	M-A	0	0	1.4033	1.0205	1.9685	0.9515	0.9752	-1.31	10.08	39.18	1.9693
FB2	B-U	0.0024	0.0249	1.4293	1.0597	2.0459	0.8892	0.9414	-2.08	11.98	45.43	2.0429
	B-A	0	0	1.2229	0.9273	1.4978	0.9189	0.9579	-1.52	10.72	40.91	1.4956
	M-U	0.0062	0.0638	1.1877	0.9106	1.4095	0.9235	0.9602	-1.51	10.66	40.03	1.4106
	M-A	0	0	1.1716	0.8866	1.3716	0.9256	0.9615	-1.39	10.32	38.50	1.3726
FB3	B-U	0.0072	0.0530	1.9832	1.4274	3.9349	0.9233	0.9602	-1.39	10.94	44.09	3.9331
	B-A	0	0	1.4391	1.0307	2.0720	0.9596	0.9795	-0.73	8.68	32.21	2.0711
	M-U	0.0256	0.1875	1.3284	0.9673	1.7648	0.9656	0.9825	0.21	8.45	30.30	1.7652
	M-A	0	0	1.3183	0.9474	1.7374	0.9661	0.9828	-0.07	8.09	28.64	1.7378
FB4	B-U	0.0789	0.8649	1.7408	1.1824	3.0380	0.9351	0.9669	0.78	14.58	56.02	3.0367
	B-A	0	0	1.2685	0.8504	1.6098	0.9656	0.9827	-0.38	11.04	41.83	1.6091
	M-U	0.0199	0.2184	1.1412	0.7783	1.3024	0.9722	0.9859	-0.99	10.62	40.04	1.3027
	M-A	0	0	1.1351	0.7567	1.2881	0.9725	0.9861	-1.13	10.09	37.68	1.2884
FD	B-U	-0.0082	-0.0598	1.9165	1.5078	3.6817	0.7850	0.8787	-3.01	12.69	45.90	3.6730
	B-A	0	0	1.4927	1.1672	2.2335	0.8696	0.9313	-1.57	9.56	35.91	2.2282
	M-U	0.0017	0.0123	1.2922	0.9805	1.6678	0.9023	0.9483	-1.17	7.93	26.04	1.6698
	M-A	0	0	1.2876	0.9729	1.6560	0.9029	0.9489	-1.10	7.86	25.92	1.6580
LT	B-U	0.0425	0.4989	2.4700	1.7728	6.1117	0.7848	0.8809	-3.68	20.42	69.45	6.1028
	B-A	0	0	1.2987	0.9495	1.6890	0.9405	0.9720	2.77	13.68	51.96	1.6866
	M-U	0.0002	0.0023	0.7867	0.5349	0.6184	0.9782	0.9889	-1.02	7.12	23.51	0.6189
	M-A	0	0	0.7859	0.5285	0.6171	0.9782	0.9890	-0.78	6.97	22.42	0.6176

Note: B-U (base model, unadjusted), B-A (base model, adjusted), M-U (mixed model, unadjusted), and M-A (mixed model, adjusted). The goodness-of-fit measures are defined in Table 5. The base and mixed models are given in [1]-[4] for different species. The number after the species code indicates the group of natural subregions. Otherwise, it is provincial (e.g., SB=SB provincial). SB1=other SB subregions, SB2=subregions 7-10, SB3=subregion 11, FB1=other FB subregions, FB2=subregions 7-9, FB3=subregion 10, and FB4=subregion 11. Subregions are defined in Table 2.

Appendix 3. Metric Conversion Chart

1 cm	=	0.39370 in.
1 m	=	3.28083 ft.
1 m	=	1.09361 yards
1 ha	=	2.47105 acres
1 m^2	=	10.76385 ft^2
1 m^3	=	35.31435 ft^3
1 m^2/ha	=	4.3560 ft^2/acre
1 m^3/ha	=	14.2913 ft^3/acre
1 ha	=	10000 m^2
1 km	=	1000 m
1 km	=	0.62137 miles
1 km^2	=	100 ha
1 km^2	=	0.3861 $miles^2$
1 in.	=	2.5400 cm
1 ft.	=	0.3048 m
1 acre	=	0.4047 ha
1 ft^2	=	0.09290 m^2
1 ft^3	=	0.02832 m^3
1 ft^2/acre	=	0.2296 m^2/ha
1 ft^3/acre	=	0.06997 m^3/ha
1 mile	=	1.6093 km
1 $mile^2$	=	2.5898 km^2
1 $mile^2$	=	258.9846 ha
1 fbm	=	1 ft. $\times$ 1 ft. $\times$ 1 in.
1 fbm	=	0.0023597 m^3
1 Mfbm	=	1000 foot board measure (fbm)
1 Mfbm	=	2.3597 m^3
1 township	=	6 miles $\times$ 6 miles = 36 $mile^2$
1 township	=	9.6558 km $\times$ 9.6558 km = 93.2345 km^2
1 township	=	9323.45 ha
1 m^3 log	$\approx$	233 board feet lumber (provincial average conversion factor)
1 Mfbm	$\approx$	4.3 m^3 log (provincial average conversion factor)


Alberta
Freedom To Create. Spirit To Achieve.