Forest Ecology and Management 498 (2021) 119574

Contents lists available at ScienceDirect

Forest Ecology and Management

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

Imputation to predict height to crown base for trees with predicted heights

Elijah Allensworth [a], Hailemariam Temesgen [b], Bryce Frank [b,*], Andrew Gray [c]

[a] Port Orchard, WA, United States
[b] Department of Forest Engineering, Resources and Management, 140 Peavy Hall, 3100 SW Jefferson Way, Oregon State University, Corvallis, OR, United States
[c] United States Department of Agriculture Forest Service, Pacific Northwest Research Station, Corvallis, OR, United States

ARTICLE INFO

Keywords:
Crown size
Allometric model
Error propagation
Mixed effects model
Calibration

ABSTRACT

Height-to-crown-base (HTCB) measurements are frequently used as inputs for growth and yield models. They are essential for reliable projections of stand structure over time required for sustainable forest management. Often, HTCB is measured alongside total height (HT) for only a subsample of trees and must be imputed for unobserved measurements in the sample. Typically, HTCB models require HT as a covariate, requiring a "double" imputation for missing HTCB measurements, leading to potentially substantial error propagation issues. We compared the effects of subsample size, imputation method, and use of imputed rather than measured HT on the accuracy of HTCB predictions. HT and HTCB were imputed using a nonlinear fixed-effects model (NFEM), a NFEM with a correction factor estimated using an ordinary least squares (OLS) regression on the subsampled measurements, and a nonlinear mixed-effects model (NMEM) with stand- and plot-level random parameters. Using cross-validated bias and root mean squared error (RMSE), our results indicated that NMEM obtained the smallest RMSE at subsample sizes greater than one, with RMSEs ranging between 2.19 and 2.51 m. However, while NMEM provided generally smaller RMSEs, biases ranging between −0.92 and −0.29 m existed at subsample sizes less than three trees per plot, and −2.86 m when no subsample was available. We observed negative bias when using imputed rather than measured HT for both the correction factor and mixed-effects model that is mitigated to a range of −0.51 m to −0.08 m with subsample sizes of at least two trees. Following this, we recommend using NMEMs for HTCB imputation with at least four trees per plot.

1. Introduction

Forest inventory projections are a necessary tool for sustainable forest management, allowing decision-makers and managers to confidently plan silviculture activities, increase the time required between field surveys, estimate recoverable timber products and associated cashflows during tactical planning, and appraise property values during long-term planning. Individual-tree growth and yield models often require accurate estimates of tree crown size for reliable long-term projections of forest stand development (e.g., Valentine and Makela, 2005). The size of a tree crown is strongly correlated with its photosynthetic capacity and is regarded as a valuable attribute for modeling individual tree growth (Maguire et al., 1988; Bianchi et al., 2020). Consequently, crown size is a strong predictor of basal area increment (Ritchie and Hann, 1987), height increment (Ritchie and Hamann, 2008), and crown recession (Maguire and Hann, 1990), which, in turn, drives the simulation of stand growth, mortality, and response to thinning over time. Crown size is generally characterized by height to crown base (HTCB), crown length, or crown ratio. HTCB is defined as the length along the main stem of a tree from the bottom of the tree to the height of the live crown base and crown ratio is the crown length divided by the total height of the tree (HT).

Often, HTCB is only measured for a subset of trees on a given field plot. For individual tree growth models that require HTCB measurements, these unobserved values must be imputed to facilitate their use. Imputation methods are frequently used in forest inventories to predict values missing in the set of sample observations (e.g., Garber et al., 2009). Although various methods have been proposed to impute crown ratios, regional species-specific models that predict either crown ratio (Temesgen et al., 2005; Sharma et al., 2017) or, to a lesser extent, height to crown base (HTCB) (e.g. Ritchie and Hann, 1987), are commonly used. These models are often nonlinear and formulated in a way that constrains predictions of the crown ratio between zero and one and predictions of HTCB between zero and HT. Many modelers favor the logistic model because it constrains the upper range of predictions to one by definition (e.g. Hasenauer and Monserud, 1996; Soares and Tome,

* Corresponding author.
E-mail address: bryce.frank@oregonstate.edu (B. Frank).

https://doi.org/10.1016/j.foreco.2021.119574
Received 2 February 2021; Received in revised form 4 June 2021; Accepted 27 July 2021
Available online 4 August 2021
0378-1127/Published by Elsevier B.V.

2001; Temesgen et al., 2005; Ducey, 2009) but other model forms including exponential (Soares and Tome, 2001; Ducey, 2009), the Chapman-Richards function (Soares and Tome, 2001), and the Weibull function (Soares and Tome, 2001) have been used as well.

Broadly, models will include covariates such as individual tree size, competition indices, and stand productivity to predict HTCB. Tree size is often characterized by HT alone (Lynch et al., 2012) but some models include diameter at breast height (DBH) in addition to HT (Temesgen et al., 2005) or in place of it (Soares and Tome, 2001). Recently, remote sensing covariates from ALS data have also been explored as a substitute for a field-measured height (Yang et al., 2020). In some studies, the ratio between diameter at breast height (DBH) to HT is included to account for the effect of the live crown on tree taper (Dyer and Burkhart, 1987; Colin and Houllier, 1992; Hynynen, 1995; Hasenauer and Monserud, 1996; Temesgen et al., 2005); a tree with a greater DBH-to-HT ratio is predicted to have a larger crown than a tree with a lesser DBH-to-HT ratio and equal height. Competition is regularly accounted for using local and stand measures. Local competition experienced by a tree has been indicated with the sum of the basal area of larger trees in the same stand (Hanus et al., 2000; Temesgen et al., 2005) or the sum of the crown competition factor of larger trees (CCFL) in the same stand (Hanus et al., 2000). Stand-level measures of competition have included basal area (BA; Hynynen, 1995; Lynch et al., 2012), crown competition factor (Hasenauer and Monserud, 1996; Temesgen et al., 2005), relative spacing (Ducey, 2009; Zhao et al., 2012) and number of trees per unit area (Soares and Tome, 2001; Fu et al., 2017).

HTCB is often measured alongside HT on a subsample of trees during field surveys. When available, this subsample of measurements from the field plot can be used to calibrate regional equations to the stand- or plot-level and improve the accuracy of imputed HT and HTCB for unmeasured trees. A commonly used method is to calculate a proportional calibration factor by regressing the measured tree attribute (i.e., HT or HTCB) on the predicted tree attribute through the origin using ordinary least squares (OLS) for the subsample (Garber et al., 2009; Özçelik et al., 2018). The resulting slope parameter estimate is then multiplied by the predicted attribute for unmeasured trees in the sample. Often, the calibration factor is further constrained to prevent large multiplicative errors. For example, the ORGANON growth and yield model constrains predictions between 0.5 and 2.0 (Hann, 2011).

A newer method of localizing HTCB equations is the use of mixed-effects models. Mixed-effects models include one or more random effects that incorporate the hierarchical structure of the data that accommodates potential variation among groups in the hierarchy. Where fixed effects do not vary regardless of the individual observation, random effects vary across groupings. This allows the random effects to partition the residual error of a model by group, increasing the precision when significant differences occur between groups. In forestry, random effects are often specified at the level of forest management units (e.g. Fu et al., 2017), referred to as "stands", that are typically homogenous in terms of individual tree characteristics, such as HT and HTCB. Additional random effects can be obtained at finer scales as well, such as field plots belonging to a specific stand. Mixed-effects models in forestry have frequently been used for HT equations (Calama and Montero, 2004; Robinson and Wykoff, 2004; Temesgen et al., 2008) but some work has been done for HTCB as well (Zhao et al., 2012). Although the ability to partition error by group is a powerful tool, mixed-effects models display poor predictive performance when extrapolated to groups not included in the original fitting dataset (Robinson and Wykoff, 2004; Temesgen et al., 2008; Garber et al., 2009). To rectify this issue, a random effect for the novel group may be predicted using a subsample of measured attributes. For example, Temesgen et al. (2008) and Garber et al. (2009) used predicted random effects to improve predictive performance of HT beyond those achieved using an OLS calibration factor when four or more heights were measured within the new stand. Raptis et al. (2021) used mixed effects models with random effects specified at the plot-level and observed declining RMSE for plot-level predictions of HT with an increase in subsample size. Sharma et al. (2017) used mixed effects models for HTCB predictions of Norway spruce and European beech and found that a subsample size of four trees at the plot-level was sufficient for calibration.

HT is often one of the most influential predictors in HTCB equations (e.g. Temesgen et al., 2008; Fu et al., 2017). HT and HTCB are typically measured together for the same tree and, when combined with subsampling, this implies that HT and HTCB must be imputed together for these missing observations. Given that HT is an important input for most HTCB equations, the use of an imputed value that is subject to model error (i.e., the error due to the imperfect prediction of HT) has the potential to propagate further error into HTCB predictions. Unfortunately, while "double" imputation of HT and HTCB together is a practical approach to resolve missing HTCB measurements, the magnitude of this source of error for HTCB predictions, to our knowledge, has received little to no exploration in the current body of literature.

The objectives of this study are to compare the accuracy of predictions of HTCB and HT using three methods: (1) a non-linear fixed effects model (NFEM) without a correction factor, (2) a NFEM with a correction factor, and (3) a non-linear mixed effects model (NMEM) with stand-level and plot-level random effects. These methods will be assessed for the purposes of predicting HTCB using measured heights and imputed heights using subsamples of different sizes, leading to plot-level subsample size recommendations for each imputation method we consider.

2. Methods

2.1. Field data

Two separate data sets were used for this analysis. The first data set was collected between 1981 and 1983 to develop the southwestern variant of the ORGANON individual-tree growth and yield model (SWO-ORGANON). The second data set was collected between 1992 and 1996 to extend the capabilities of SWO-ORGANON to old-growth and young hardwood forests (Hann, 2011). The study area for both data sets is bounded by Cow Creek (43°00′ N) in the north, the crest of the Cascades (122°15′ W) in the east, just north of the California-Oregon border (42°10′ N) in the south, and 15 miles west of Glendale, OR (123°50′ W) in the west, with elevations between 275 and 450 m (Hanus et al., 2000).

The candidate stands for this study were identified by restricting the study area to homogeneous, accessible mixed species stands less than 120 years old with at least 50% of total basal area occupied by conifers located on the ownerships of study participants (Hann, 1982). For the second data set, the selection criteria were modified to allow stands older than 250 years and younger stands with a greater hardwood component (Hanus et al., 2000). In total, 529 stands were considered in this study. Field plot positions were located within each stand using a two-dimensional systematic sampling design (e.g., D'Orazio, 2003). Plots were positioned using an equilateral triangular grid of points spaced 45.7 m apart. The initial sample point was located 45.7 m from the point on the stand boundary nearest to the road and along the approximate major axis of the stand, rounded to the nearest 10° (Hann, 1982). At each sample point, four subplots were measured: a fixed-radius plot of 17.7 m^2 for trees with a DBH less than or equal to 10.2 cm; a fixed-radius plot of 71 m^2 for trees with a DBH greater than 10.2 and less than or equal to 20.3 cm; a variable-radius plot using a BAF of 3.7 m^2 for trees with a DBH greater than 20.3 cm and less than 91.4 cm; and a variable-radius plot using a BAF of 5.6 m^2 for trees with a DBH exceeding 91.4 cm.

The DBH, HT, and HTCB was recorded for all trees within any of the four subplots (Hann, 1982). The base of the live crown was determined as the height about the ground where the live crown would begin if the live branches were evenly distributed on either side of the stem, ignoring short branches less than one meter long and epicormic branches. Additionally, crown competition (CCF) factor was calculated for each tree using the maximum crown width equations detailed in Paine and Hann (1982) and site index (SI) was calculated for each stand using the equation created by Hann and Scrivani (1987).

For both data sets combined, 34,972 trees were measured among 529 stands with four to 25 plots within each stand and one to 76 trees within each plot. Given that Douglas-fir accounted for most of the data, with 19,779 trees among 339 stands, and is a commercially important species for the region, it was selected as the focus of this analysis. Thus, DBH, HT, HTCB, and CCF for each undamaged Douglas-fir tree was separated from the combined parent data sets, as well as basal area – in square meters per hectare – and site index for associated stands. Since examining the effects of subsampling and imputing HT and HTCB within a plot is the goal of this study, plots with fewer than seven Douglas-fir trees and stands with fewer than three plots containing seven Douglas-fir trees each were discarded. The resultant data set comprised of 2,884 Douglas-fir trees among 87 stands covering a large range of densities, site productivity, and size attributes (Table 1). Figs. 1 and 2 display the covariates plotted against their respective response variable.

2.2. Total height imputation methods

Imputation of HT and HTCB can be conducted in various ways

Table 1
Mean, standard deviation, and ranges for the HT and HTCB, as well as the covariates used in the models for each of these variables.

	Attribute	Mean	Standard Deviation	Minimum	Maximum
Tree-level (n = 2, 884)	DBH(cm)	36.54	21.56	0.25	178.82
	HT(m)	27.21	11.32	1.40	62.06
	HTCB(cm)	15.77	7.79	0.37	40.81
	CCFL (m^2ha^{-1})	49.72	33.83	0.00	196.52
	PCCFL(m^2ha^{-1})	52.55	41.41	0.00	301.84
Stand-level (m = 87)	SI45(m)	29.35	5.22	15.45	42.06
	BA(m^2ha^{-1})	54.89	13.55	9.83	101.01
	H5D5 ($\frac{m \times cm}{10,000}$)	0.08	0.01	0**	0.62

*H5D5 refers to the product of the average height and diameter of the trees with the five largest diameters in the stand, scaled by a constant of 10,000.
**Refers to a value less than 1e−03.

depending on the level of subsampling within a new stand. In some cases, HT of some trees may be unobserved due to subsampling. For all HTCB imputation methods we consider, HT is used as a covariate, therefore, we first require methods to impute unobserved tree heights. We consider three such methods.

Non-linear Fixed Effects Model (NFEM-HT): We used a Chapman-Richards NFEM previously developed by Temesgen et al. (2007) for HT imputation on the same 1981–1983 Douglas-fir data sets. Assume the total heights of trees follow the model:

$$\mathrm{HT}_{ijk} = 1.37 + \left(\alpha_1 + \alpha_2 \mathrm{CCFL}_{ijk} + \alpha_3 \mathrm{BA}_i\right)\left(1 - \exp\left(\alpha_4 \mathrm{DBH}_{ijk}\right)\right)^{\alpha_5} + \varepsilon_{ijk} \quad (1)$$

where HT_{ijk} is the total height of tree k in plot j in stand i; CCFL_{ijk} is the CCF of trees larger in diameter than tree k in plot j in stand i; BA_i is the basal area per hectare of stand i; DBH_{ijk} is the diameter at breast height of tree k in plot j in stand i; α_1, α_2, α_3, α_4, and α_5 are parameters to be estimated; and ε_{ijk} is the residual error term for tree j in stand i, assumed to be independently distributed $N(0, \sigma^2_{HT}\left|DBH_{ijk}\right|^{2\delta})$ where δ is a parameter controlling heteroskedasticity. We obtain the prediction for the total height:

$$\widehat{\mathrm{HT}}_{1,ijk} = 1.37 + \left(\widehat{\alpha}_1 + \widehat{\alpha}_2 \mathrm{CCFL}_{ij} + \widehat{\alpha}_3 \mathrm{BA}_i\right)\left(1 - \exp\left(\widehat{\alpha}_4 \mathrm{DBH}_{ijk}\right)\right)^{\widehat{\alpha}_5} \quad (2)$$

where $\widehat{\alpha}_1$, $\widehat{\alpha}_2$, $\widehat{\alpha}_3$, $\widehat{\alpha}_4$, and $\widehat{\alpha}_5$ are the estimated parameters obtained from the *gnls* function in the R programming language (R Core Team, 2020) from the package *nlme* (Pinherio et al., 2021). Parameter estimates were obtained using the Broyden-Fletcher-Goldfarb-Shann algorithm.

Non-linear Fixed Effects Model with Plot-level Correction (NFEMwC-HT): HT imputations can be calibrated using a correction factor constructed from an OLS regression using n_{ij} trees from plot j in stand i. The correction factor is the slope estimate of the OLS regression through the origin with measured heights as the dependent variable and heights predicted with Eq. (3) as the independent variable. We obtain the correction factor:

$$\widehat{k}_{ij} = \frac{\sum_{k=1}^{n_{ij}}\left[\left(\widehat{HT}_{1,ijk} - 1.37\right)\left(HT_{ijk} - 1.37\right)\right]}{\sum_{j=1}^{n_{ij}}\left[\left(\widehat{HT}_{1,ijk} - 1.37\right)^2\right]} \quad (3)$$

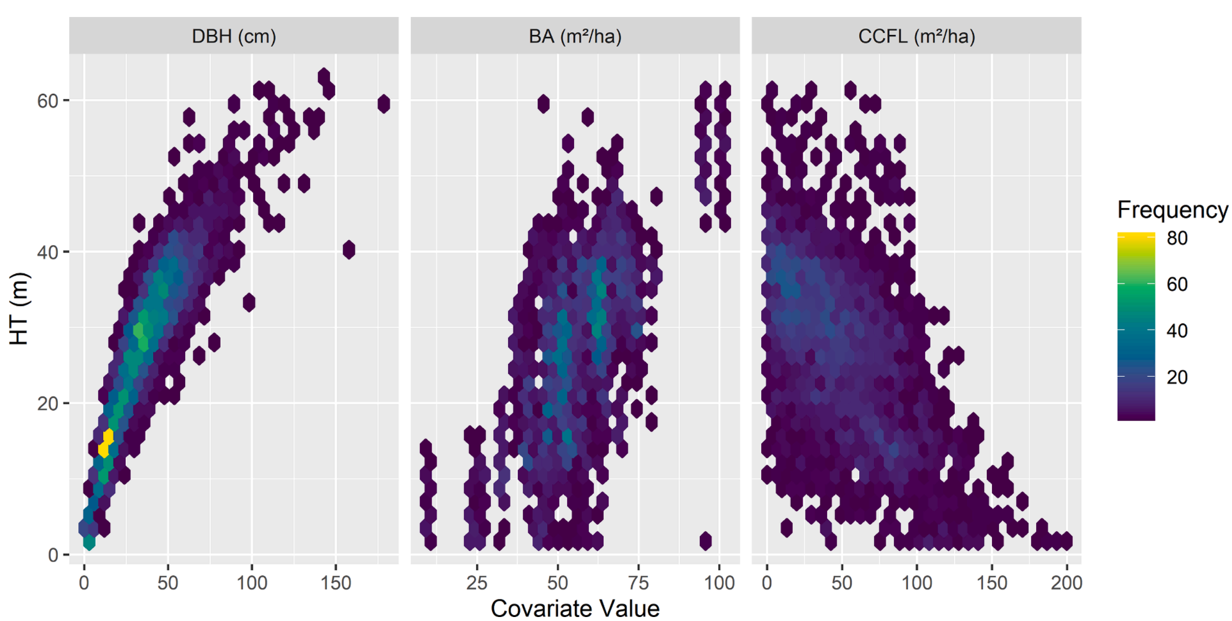

Fig. 1. Covariates used in the HT models plotted against HT, cells are colored by the number of trees in the discrete hexagonal bin.

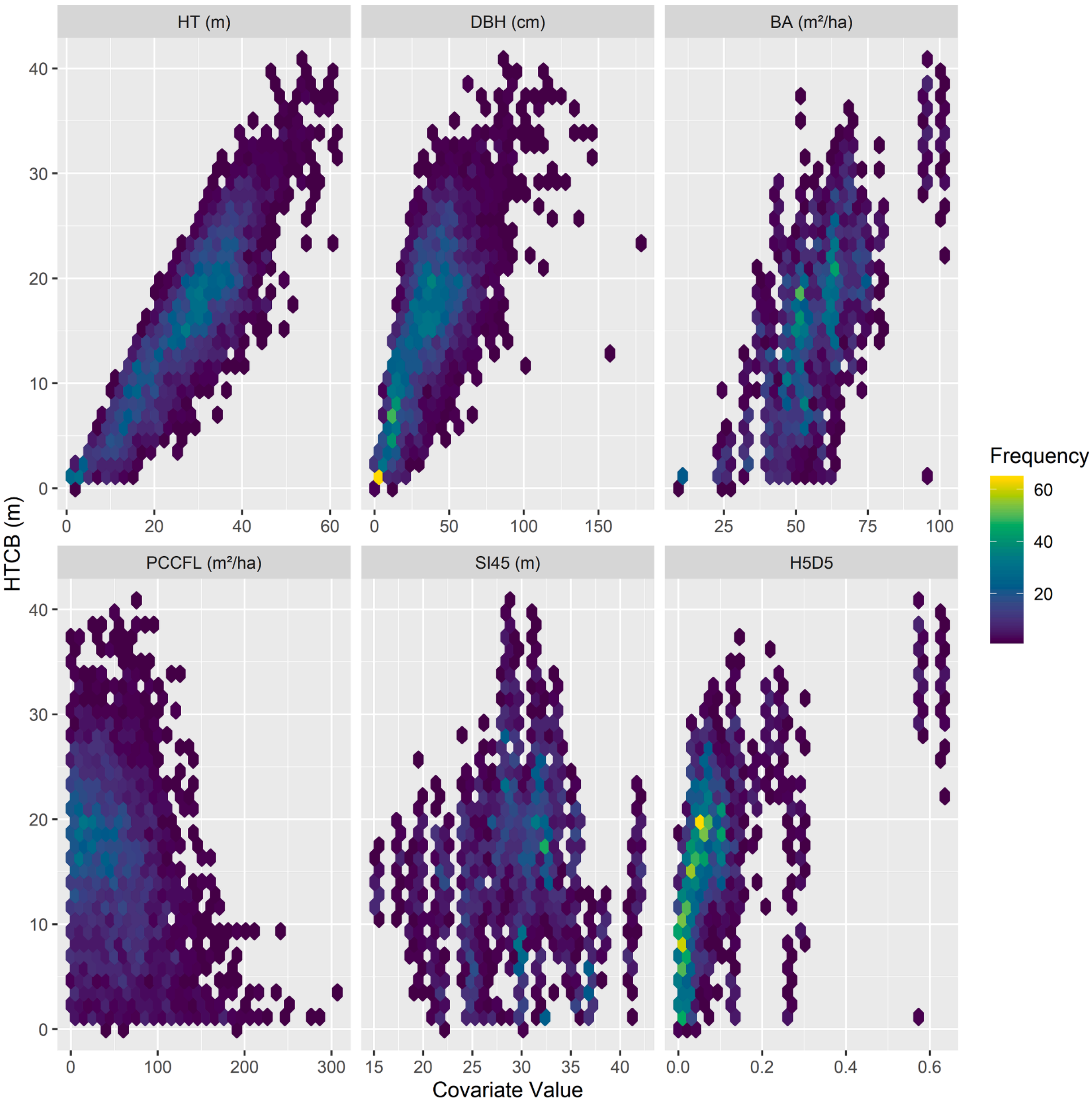

Fig. 2. Covariates used in the HTCB models plotted against HTCB, cells are colored by the number of trees in the discrete hexagonal bin.

where $\widehat{HT}_{1,ijk}$ is the predicted height using the NFEM (Eq. (2)) for tree k in plot j in stand i and HT_{ijk} is the measured height. The correction factor is then applied to all predicted heights in stand i for trees without HT measurements. We obtain:

$$\widehat{\mathrm{HT}}_{2,ijk} = 1.37 + \widehat{k}_{ij}\left(\widehat{\alpha}_1 + \widehat{\alpha}_2\mathrm{CCFL}_{ijk} + \widehat{\alpha}_3\mathrm{BA}_i\right)\left(1 - \exp\left(\widehat{\alpha}_4\mathrm{DBH}_{ijk}\right)\right)^{\widehat{\alpha}_5} \quad (4)$$

Non-linear Mixed Effects Model (NMEM-HT): HT imputations can be calibrated with a non-linear mixed effects model. Under the NMEM-HT model, let a_{1i} represent the random effect for stand i, and let a_{1ij} represent the random effect for plot j in stand i. Both are assumed to be independently and identically distributed $N(0, \sigma_{a_s}^2)$ and $N(0, \sigma_{a_p}^2)$ for the stand and plot random effects, respectively. Furthermore the stand and plot random effects are independent of each other. The random effect is predicted using a best linear unbiased predictor (BLUP), as implemented by Lynch et al. (2005), Temesgen et al. (2008), and Garber et al. (2009), that incorporates height measurements from subsampled trees in the stand. We modified the method to obtain stand and plot-level random effects. We obtain the prediction of the random effect for stand i:

$$\widehat{\boldsymbol{a}}_{1ij} = \widehat{\mathbf{A}}\mathbf{Z}_{ij}^T\left(\widehat{\boldsymbol{R}}_{ij} + \mathbf{Z}_{ij}\widehat{\mathbf{A}}\mathbf{Z}_{ij}^T\right)^{-1}\begin{bmatrix}\mathrm{HT}_{ij1} - \widehat{HT}_{1,ij1} \\ \vdots \\ \mathrm{HT}_{ijn_{ij}} - \widehat{HT}_{1,ijn_{ij}}\end{bmatrix} \quad (5)$$

where $\widehat{\boldsymbol{a}}_{1ij} = \left[\widehat{a}_{1i}, \widehat{a}_{1ij}\right]^T$ is the vector of predicted random effects for plot j in stand i, $\widehat{\mathbf{A}} = diag(\sigma_{a_s}^2, \sigma_{a_p}^2)$ is the estimated variance-covariance matrix of the random effects, $\widehat{\boldsymbol{R}}$ is the estimated variance–covariance matrix for the residuals, and$\mathbf{Z}_{ij} = [z_{ij}, z_{ij}]$ is the design matrix for the random effects, where each j^{th} element is defined:

$$[z_{ij}]_k = \left(1 - \exp\left(\widehat{\alpha}_4\mathrm{DBH}_{ijk}\right)\right)^{\widehat{\alpha}_5} \quad (6)$$

The height of a tree not included in the subsample can then be calculated using Eq. (7):

$$\widehat{\mathrm{HT}}_{3,ijk} = 1.37 + \left(\widehat{\alpha}_1 + \widehat{a}_{1i} + \widehat{a}_{1ij} + \widehat{\alpha}_2 \mathrm{CCFL}_{ijk} + \widehat{\alpha}_3 \mathrm{BA}_i\right)\left(1 - \exp\left(\widehat{\alpha}_4 \mathrm{DBH}_{ijk}\right)\right)^{\widehat{\alpha}_5} \quad (7)$$

2.3. Height to crown base imputation methods

Like Section 2.2, we compare three different methods for imputing HTCB. However, the HTCB imputation methods we consider rely on some measure of total height for each tree. A general case will consider either using measured HT or imputed HT for heights that were not selected in the subsample. Let $\check{HT}_{g,ijk} \in \{HT_{ijk}, \widehat{HT}_{g,ijk}\}$ represent either the measured height of the tree or the imputed height using the g th height imputation method from Section 2.2. Doing so we obtain six cases, three HTCB imputation methods where each either uses imputed tree heights or uses all measured heights. We describe each HTCB imputation method below using $\check{HT}_{g,ijk}$ as a stand-in for either a measured or imputed height using the gth height imputation method in Section 2.2.

Non-linear Mixed Effects Model (NFEM-HTCB): We used a logistic model developed by Hanus et al. (2000), who used the 1981–1983 Douglas-fir data sets to model the relationship between the covariates and HTCB. We assume height to crown base follows the model:

$$\mathrm{HTCB}_{ijk} = \frac{\mathrm{HT}_{ijk}}{1 + \exp(\beta_1 + \theta_{ijk} + \eta_i)} + \zeta_{ijk} \quad (8)$$

$$\theta_{ijk} = \beta_2 \mathrm{HT}_{ijk} + \beta_3 \mathrm{PCCFL}_{ijk} + \beta_4 \left(\frac{\mathrm{DBH}_{ijk}}{\mathrm{HT}_{ijk}}\right) \quad (9)$$

$$\eta_i = \beta_5 \ln(\mathrm{BA}_i) + \beta_6(\mathrm{SI}_i - 1.37) + \beta_7 \left(\frac{\mathrm{HT5}_i \times \mathrm{DBH5}_i}{10,000}\right)^2 \quad (10)$$

where HTCB_{ijk} is the height to the base of the compacted crown for tree k in plot j stand i; HT_{ijk} is the total height of the stem for tree k in plot j in stand i; PCCFL_{ijk} is the crown competition factor of trees that are located in the same plot and are larger by diameter than tree k in plot j stand i; SI_i is the site index of stand i; $\mathrm{HT5}_i$ is the average height of the five trees with the largest diameter in stand i; $\mathrm{DBH5}_i$ is the average diameter at breast height of the five trees with the largest diameter in stand i; $\beta_1, \ldots, \beta_7$ are slope parameters; and ζ_{ijk} is the residual error term, assumed to be independently distributed $N(0, \sigma^2_{HTCB}|HT_{ijk}|^{2\delta})$, where δ is a parameter controlling heteroskedasticity. We obtain the prediction for the height to crown base of tree k in plot j in stand i using the non-linear mixed effects model:

$$\widehat{\mathrm{HTCB}}_{1,ijk} = \frac{\check{HT}_{1,ijk}}{1 + \exp\left(\widehat{\beta}_1 + \check{\theta}_{1,ijk} + \check{\eta}_{1,i}\right)} \quad (11)$$

where $\check{\theta}_{1,ijk}$ and $\check{\eta}_{1,i}$ are obtained using either measured or imputed heights from $\widehat{HT}_{1,ijk}$ for covariates involving HT.

Non-linear Fixed Effects Model with Plot-level Correction (NFEMwC – HTCB): The second strategy uses Eq. (11) but proportionally adjusts the predicted heights to crown base by using n_{ij} measured heights from crown base from subsampled trees in plot j in stand i. We obtain:

$$\widehat{h}_{ij} = \frac{\sum_{k=1}^{n_{ij}} \left[\left(\widehat{HTCB}_{1,ijk} \times HTCB_{ijk}\right)\right]}{\sum_{j=1}^{n_i} \left[\left(\widehat{HTCB}_{1,ijk}\right)^2\right]} \quad (12)$$

where $\widehat{HTCB}_{1g,ijk}$ is the predicted height to crown base from Eq. (11), $HTCB_{ijk}$ is the observed height to crown base and $\widehat{h}_{ij}$ is the correction factor (analogous to $\widehat{k}_{ij}$ for the NFEMwC-HT method). The correction factor is then applied to HTCB predictions for unobserved trees:

$$\widehat{HTCB}_{2,ijk} = \widehat{h}_{ij}\left(\frac{\check{HT}_{2,ijk}}{1 + \exp\left(\widehat{\beta}_1 + \check{\theta}_{2,ijk} + \check{\eta}_i\right)}\right) \quad (13)$$

Non-linear Mixed Effects Model (NMEM-HTCB): The third strategy modifies the NFEM-HTCB method with the inclusion of random effects. For some stand i, the random effect is predicted using a best linear unbiased predictor (BLUP) that incorporates height to crown base measurements from subsampled trees in the stand. We obtain:

$$\widehat{\boldsymbol{b}}_{ij} = \widehat{\boldsymbol{B}}\boldsymbol{Q}_{ij}^T\left(\widehat{\boldsymbol{S}}_{ij} + \boldsymbol{Q}_{ij}\widehat{\boldsymbol{B}}\boldsymbol{Q}_{ij}^T\right)^{-1}\begin{bmatrix} \mathrm{HTCB}_{ij1} - \widehat{HTCB}_{1,ij1} \\ \vdots \\ \mathrm{HTCB}_{ijn_{ij}} - \widehat{HTCB}_{1,ijn_{ij}} \end{bmatrix} \quad (14)$$

where $\widehat{\boldsymbol{b}}_{ij} = \left[\widehat{b}_i, \widehat{b}_{ij}\right]^T$ is the vector predicted random effects for plot j in stand i, $\widehat{\boldsymbol{B}} = diag\left(\widehat{\sigma}^2_{b_s}, \widehat{\sigma}^2_{b_p}\right)$ is the estimated variance–covariance matrix of the random effects, $\widehat{\boldsymbol{S}}_{ij}$ is the estimated variance–covariance matrix of the residuals, and $\boldsymbol{Q}_{ij} = [\boldsymbol{q}_{ij}, \boldsymbol{q}_{ij}]$ is the design matrix for the random effects, which in this case is a n_{ij}x2 matrix containing ones. The height to crown base of an unobserved tree can then be predicted:

$$\widehat{\mathrm{HTCB}}_{3,ijk} = \left(\frac{\check{HT}_{3,ijk}}{1 + \exp\left(\widehat{\beta}_1 + \check{\theta}_{3,ijk} + \check{\eta}_{3,i}\right)}\right) + \widehat{b}_i + \widehat{b}_{ij} \quad (15)$$

2.4. Simulation procedure

To compare different HTCB imputation methods we used a cross-validation procedure meant to assess the performance of the models for visiting a new, unobserved stand and plot at different subsampling levels. For each stand in the dataset, we obtain model coefficient estimates using all trees not in the stand. Then, we obtain different sets of cross-validated errors by predicting HTCB for all trees in all plots in the hold-out stand. Within each plot we subsample with sizes ranging between zero and six trees. Cross-validated errors are obtained by subtracting the predicted HTCB from the observed HTCB for each imputation method. For subsample sizes greater than zero, calibration methods discussed in Sections 2.2 and 2.3 are used to adjust predictions from the fixed effects. For each stand $i \in \{1, \cdots, 87\}$:

Obtain fixed parameter estimates for the NFEM and NMEM HT and HTCB models using trees from all stands $\neq i$
For each subsample size $p \in \{0, 1, 2, 3, 4, 5, 6\}$ select a sample of trees within stand i of size p using simple random sampling without replacement.
For each HTCB imputation method $g \in \{1, 2, 3\}$ obtain the cross-validated errors for all unsampled trees in the plot, $e^{(l)}_{ijk} = \widehat{HTCB}_{g,ijk} - HTCB_{ijk}$. Note, for subsample size $p = 0$, $\widehat{h}_{ij}$, $\widehat{k}_{ij}$ are set to one and $\widehat{a}_{1i}$, $\widehat{a}_{1ij}$, $\widehat{b}_{1i}$, $\widehat{b}_{1ij}$ are set to zero, representing the case of no calibration.

2.5. Performance measures

In the simulation procedure, a set of errors is obtained for every tree in the data set using each method defined in the simulation. *RMSE* and *Bias* are obtained by taking the mean of the squared cross-validated errors and the mean of the errors, respectively. Each set of cross-validated errors can be indexed by l. We produce the root mean squared error for one iteration of the simulation:

$$RMSE^{(l)}\left(\widehat{HTCB}_g\right) = \sqrt{\overline{e_g^{(l)^2}}} \tag{16}$$

where $\overline{e_g^{(l)^2}}$ is the mean of the squared cross-validated errors for all errors produced in the lth iteration, and the bias for the lth iteration:

$$Bias^{(l)}\left(\widehat{HTCB}_g\right) = \overline{e_g^{(l)}} \tag{17}$$

We obtain the mean *RMSE* across all 30 iterations:

$$\overline{RMSE}\left(\widehat{HTCB}_g\right) = \frac{1}{L}\sum_{l=1}^{L} RMSE^{(l)}\left(\widehat{HTCB}_g\right) \tag{18}$$

and the mean *Bias* across all 30 iterations:

$$\overline{Bias}\left(\widehat{HTCB}_g\right) = \frac{1}{L}\sum_{l=1}^{L} Bias^{(l)}\left(\widehat{HTCB}_g\right) \tag{19}$$

In addition to these simulation procedures, we also fit the NFEM-HT, NFEM-HTCB, NMEM-HT and NMEM-HTCB models using all available data. From these fits we obtain standardized residual plots and R^2 for each model:

$$R^2_{g,HT} = 1 - \sum_{i=1}^{m}\sum_{j=1}^{n_{ij}}\left(HT_{ijk} - \widehat{HT}_{g,ijk}\right)^2 \Bigg/ \sum_{i=1}^{m}\sum_{j=1}^{n_{ij}}\left(HT_{ijk} - \overline{HT}_{g,ijk}\right)^2 \tag{20}$$

and

$$R^2_{g,HTCB} = 1 - \sum_{i=1}^{m}\sum_{j=1}^{n_{ij}}\left(HTCB_{ijk} - \widehat{HTCB}_{g,ijk}\right)^2 \Bigg/ \sum_{i=1}^{m} \times \sum_{j=1}^{n_{ij}}\left(HTCB_{ijk} - \overline{HTCB}_{g,ijk}\right)^2 \tag{21}$$

where $\overline{HT}$ and $\overline{HTCB}$ are the mean HT and HTCB in the data set.

3. Results

3.1. Selected models and baseline performance

In total, four models are fit, a NFEM and NMEM for both HT and HTCB. Tables 2 and 3 report the parameter estimates for each of these models fit to the entire data set. For the HT models, parameter estimates are similar between NFEM and NMEM models, with the larger differences existing between α_1, an intercept parameter, and α_2 and α_3 which are the coefficients for CCFL and BA, respectively. The NMEM-HT model has a stand-level random effect variance of 4.075 and a plot-level random effect variance of 2.141 implying that plot- and stand-level variance comprises a large share of the total variance, with plot-level variance composing a smaller share. Parameter estimates are similar between NFEM-HTCB and NMEM-HTCB models except for the β_1 and β_7 parameters, which are the intercept for the linear predictor and the coefficient for the squared diameter-to-height ratio, respectively. The stand- and plot-level random effect variance estimates for the NMEM-HTCB model are 3.551 and 1.050, respectively, a similar pattern to that of the NMEM-HT model, with stand-level variance composing the largest share of variance, followed by the plot-level and the residual variance. For all models, the parameter δ, which controls the heteroskedastic variance function, was 0.443 and 0.557 for NFEM-HT and NMEM-HT and 0.788 and 0.630 for NMEM-HT and NMEM-HTCB. This indicates that variance increases with increasing diameter and height for HT and HTCB, respectively.

Table 2
Parameter estimates and standard errors for HT NFEM and NMEM models using all available data.

Parameter	NFEM-HT		NMEM-HT	
	Estimate	Std. Error	Estimate	Std. Error
α_1	46.497	1.718	28.232	2.308
α_2	0.182	0.012	0.066	0.009
α_3	0.038	0.017	0.231	0.036
α_4	−0.023	0.001	−0.029	0.001
α_5	1.191	0.016	1.070	0.018
σ_{a_s}	–	–	4.075	–
σ_{a_p}	–	–	2.141	–
σ_{HT}	0.472	–	0.526	–
δ	0.557	–	0.433	–

*Indicates an absolute value <0.0005.

Table 3
Parameter estimates and standard errors for HTCB NFEM and NMEM models using all available data.

Parameter	NFEM-HTCB		NMEM-HTCB	
	Estimate	Std. Error	Estimate	Std. Error
β_1	1.872	0.192	2.744	0.461
β_2	−0.019	0.001	0*	0.001
β_3	0.004	0.000*	−0.00	0.000*
β_4	−0.699	0.049	−0.768	0.115
β_5	0.551	0.038	0.209	0.033
β_6	0.022	0.002	0.013	0.004
β_7	0.415	0.002	−0.776	0.209
σ_{b_s}	–	–	3.551	–
σ_{b_p}	–	–	1.050	–
σ_{HTCB}	0.216	–	0.252	–
δ	0.788	–	0.630	–

* Indicates an absolute value <0.0005.

HT and HTCB models generally describe the response variable well. R^2 values ranged between 0.85 and 0.94 for HTCB and 0.91 and 0.96 for HT. Fig. 3 displays these values along with standardized residuals for each model. The standardized residuals appear homoscedastic for all models. The extreme upper range of HTCB predictions, e.g., predictions exceeding 35 m for the NFEM-HTCB model, indicate some tendency for a positive bias in the prediction. The NMEM-HTCB model does not demonstrate this pattern.

For each subsample size and for each iteration of the simulation, a set of cross-validated errors can be obtained for each imputation method. Fig. 4 displays these cross-validated errors for the first iteration of the simulation as an example. Note that the figure does not explicitly display the NFEM-HTCB errors, they are given instead as a special case of NFEMwC-HTCB at a subsample size of zero (i.e., by setting $k_{ij} = 1$). The trend lines in Fig. 4 indicate a tendency for over-prediction of HTCB for most methods, regardless of measured or imputed heights or subsample size, with NMEM-HTCB using imputed and measured heights as notably egregious cases. However, as subsample size increases, all imputation methods resolve this bias to some extent. NMEM appears to obtain a smaller error variance for subsample sizes greater than zero relative to NMEMwC for both imputed and measured heights.

3.2. Impact of subsampling on model performance

In nearly all field surveys, some level of subsampling for HT and HTCB measurements will occur, which allows for the use of either the OLS correction factor or the predicted random effect. Using the simulation procedure (Section 2.4) we obtained $\overline{RMSE}$ and $\overline{Bias}$ for the three HTCB methods made using any of the matching HT methods as well as using measured heights. Fig. 5 displays the $\overline{RMSE}$ for each subsample

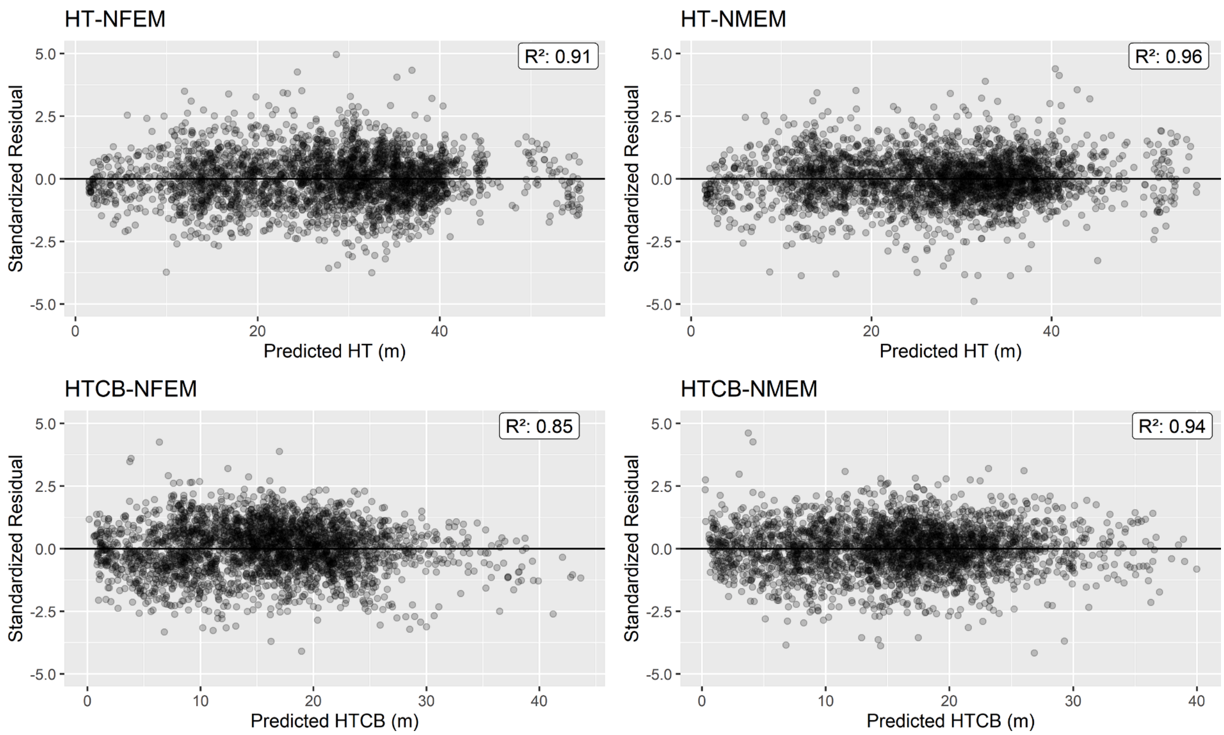

Fig. 3. Standardized residual plots for the four base models. R^2 values are provided in the top right corner for each model.

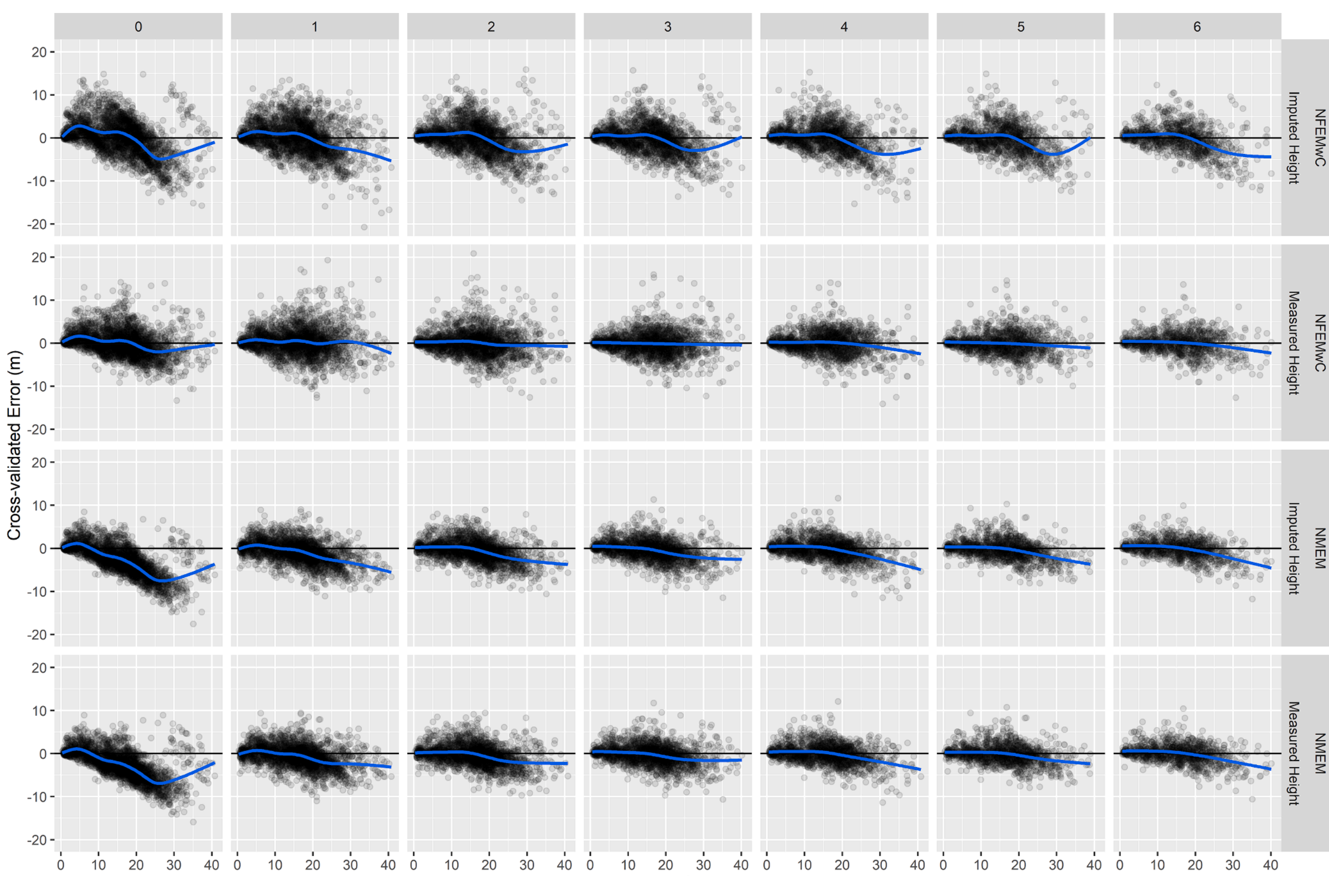

Fig. 4. Cross-validated errors of the HTCB imputation methods for subsample sizes zero through six for the first iteration of the simulation. LOESS smoothing curves are given in blue to indicate trend.

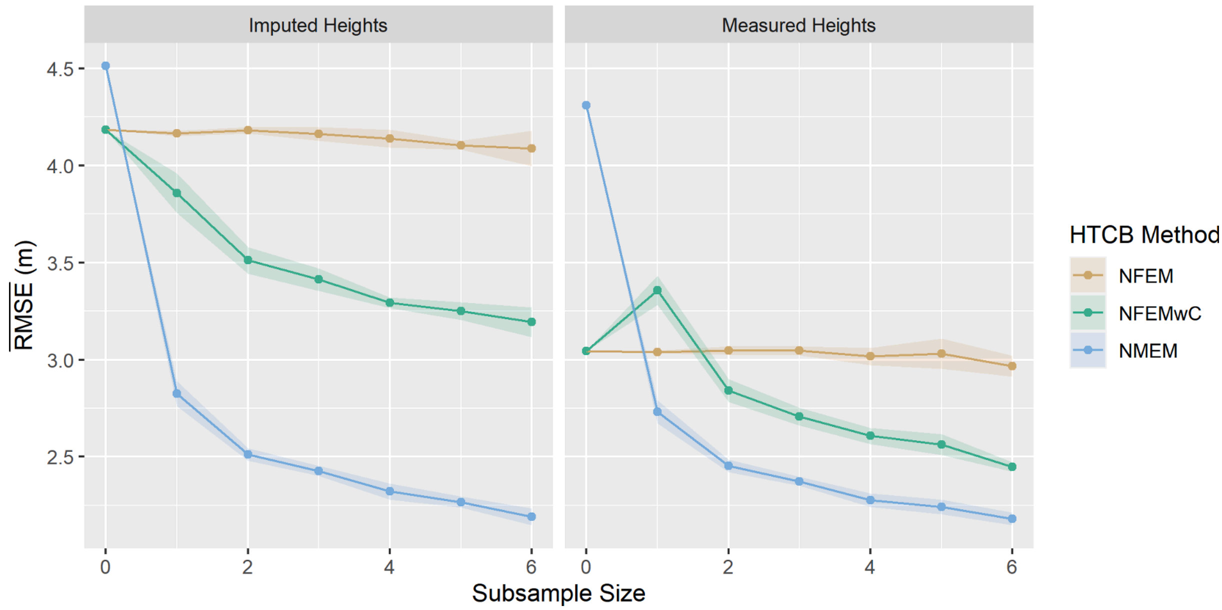

Fig. 5. $\overline{RMSE}$ for each HT imputation method and each HTCB method for each sample size. The colored envelopes of each line indicate the bound created by ±1 standard deviation of the RMSEs across the 500 simulations.

size for each method. All methods benefit with an increase in subsample size, and substantial $\overline{RMSE}$ reductions are apparent for the NMEM model when moving from zero to one subsample size. Uniformly, for each subsample size, using known heights obtains the lowest $\overline{RMSE}$ where the NMEM-HTCB model obtains the lowest $\overline{RMSE}$. Table 4 contains the $\overline{RMSE}$ for each HT-HTCB method combination for sample sizes $\in \{0, 2, 4, 6\}$. When HT imputation is required, the minimum $\overline{RMSE}$ is obtained when NMEMs are used to impute both missing HT and HTCB measurements, except for a subsample size of zero ranging between 2.19 and 2.51 m. When heights are fully measured, imputing HTCB with NMEM obtained the lowest $\overline{RMSE}$ for all sample sizes, except for a subsample size of zero, ranging between 2.18 and 2.84 m. NMEMs demonstrated much higher $\overline{RMSE}$ with a subsample size of zero compared to the NFEMs, with an increase of 0.33 m and 1.27 m for imputed and measured heights, respectively. All methods demonstrated diminishing returns as subsample size increases, with the largest gains in precision achieved with sample sizes zero through three.

Smaller subsample sizes demonstrated larger biases than large subsample sizes, especially for the NMEM-HTCB method. The biases for each subsample size across the 30 simulations are reported in Fig. 6. NMEM demonstrates relatively large negative bias for subsample sizes less than four trees, but is mitigated with an increase in subsample size.

Table 4
$\overline{RMSE}$ and $\overline{Bias}$ for all HTCB imputation methods under HT imputation and measured HT for a subset of the subsample sizes $\in \{0, 2, 4, 6\}$. The lowest RMSE and bias closest to zero for a given subsample size is indicated in bold for each subsample size and imputed vs. measured crossing.

Subsample Size	HTCB Imputation Method	Imputed HT		Measured HT	
		$\overline{RMSE}$	$\overline{Bias}$	$\overline{RMSE}$	$\overline{Bias}$
$n_{ij} = 0$	NFEMwC	**4.18**	0.00*	**3.04**	**0.03**
	NMEM	4.51	−2.86	4.31	−2.67
$n_i = 2$	NFEMwC	3.51	−0.08	**2.84**	0.07
	NMEM	**2.51**	−0.51	2.45	−0.35
$n_i = 4$	NFEMwC	3.29	−0.22	2.61	−0.03
	NMEM	**2.32**	−0.34	**2.28**	−0.19
$n_i = 6$	NFEMwC	3.19	−0.26	2.45	−0.07
	NMEM	**2.19**	−0.23	**2.18**	−0.11

* Indicates a value less than 5e−3.

NFEM and NMEMwC did not demonstrate these bias issues. Importantly, all methods achieve smaller biases with increases in subsample size. Table 4 reports the biases for each combination of imputation method under imputed and measured HT. For measured HT, biases for subsample sizes greater than zero ranged between −0.35 to 0.08 m across all three methods and all sample sizes, with NFEM achieving the lowest absolute bias for all sample sizes.

4. Discussion

Overall, the NMEM-HTCB imputation method performed the best of all three methods under subsampling, providing the smallest $\overline{RMSE}$ at any subsample size for both measured and imputed HT (Fig. 5, Table 4), but demonstrated bias issues using small subsample sizes (Fig. 6). The NFEMwC-HTCB method demonstrated advantages over using only the NFEM as well, but to a lesser extent than the mixed-effects model, due to a relatively larger $\overline{RMSE}$. This result can be framed with respect to the bias-variance tradeoff. For predictions of elements in groups, such as the predicted HTCB in a plot, mixed-effects models will compensate between a "global" part of the model, common to all groups, and the "local" observations within the group via the predicted random effect, which can lead to bias over all predictions (Townsend et al., 2013; Clark and Linzer, 2015). This is structurally different from the OLS correction factor that regresses purely on the local residuals and does not incorporate the relative share of random effect variance as a weight, reducing potential bias for plot-level predictions, and driving up the variance, resulting in larger RMSEs relative to the NMEM.

Despite the superiority of the NMEM-HTCB imputation method under larger subsample sizes, we observed higher $\overline{RMSE}$ values for subsample sizes of zero, i.e., the case that did not leverage subsampling, compared to the NFEM-HTCB imputation method, as well as positive bias issues in the upper range of HTCB predictions (Fig. 3). Predictions of HTCB that do not leverage subsampling may be a common scenario in forest inventories when imputing HTCB for stands that have already been visited and lack HTCB measurements. The implications of this result should be considered for this type of application. Ultimately, the mixed effects model we used for HTCB contains two additional parameters. It is reasonable to expect some disagreement between parameter estimates of the NFEM and NMEM models for the parameters that are shared (Tables 2 and 3), because the fitting algorithms are optimizing for

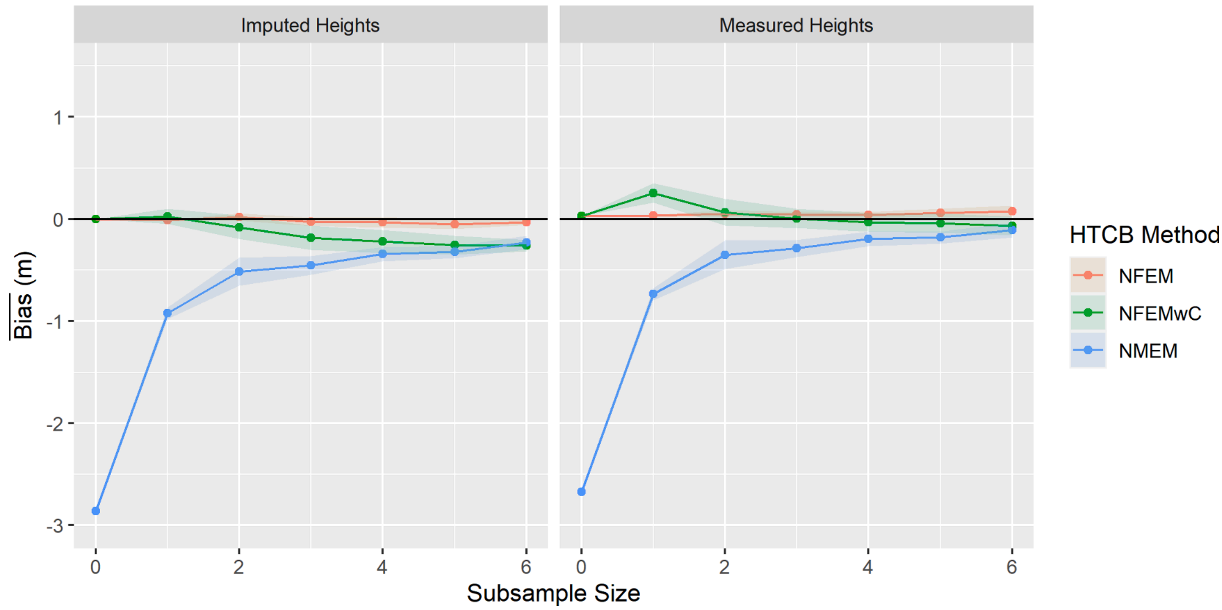

Fig. 6. $\overline{Bias}$ for each HT imputation method and each HTCB method for each sample size. The colored envelopes of each line indicate the bound created by ±1 standard deviation of the biases across the 500 simulations.

different solutions, and can lead to differences in the cross-validated errors in hold-out stands (Fig. 3). We utilized a previously published model for the NFEM-HTCB and NMEM-HTCB models (Hanus et al., 2000), but further research could examine the NMEM-HTCB model in greater detail, or develop alternative modeling systems such as multivariate models (e.g. Eerikäinen, 2009), to potentially alleviate the positive bias problems we observed.

Using imputed HT as inputs into HTCB models leads to error propagation into the final HTCB predictions, leading to potentially increased RMSEs, biases, or both. How the errors in HT imputation interact and propagate will depend on the error structure in the HT model and the HTCB model formulation. Temesgen et al. (2008) found that the NFEMwC-HT method, fit to a similar dataset, tended to have a bias of 0.25 m, which may explain the positive biases we observed that are of similar magnitude in the case of HTCB predictions (Fig. 4). The magnitude of this bias declines with an increase in subsample size.

The use of measured HT to impute HTCB resulted in an overall reduction of $\overline{RMSE}$ when compared to using imputed total heights (Fig. 5), which is driven by the elimination of prediction error for HT. Across all subsample sizes and iterations, measuring total height for all sampled trees reduced the error by 2.04 and 0.20 m for the NFEM and NMEM without subsampling, respectively (Table 4). Under subsampling, reductions in RMSE resulting from using measured rather than imputed HT ranged between 0.67 and 0.74 m and 0.01–0.60 m for NFEMwC and NMEM, respectively. Furthermore, not only was the mean predictive error less within HTCB imputation methods when HT was measured, it was less for all subsample sizes across imputation methods as well, especially for smaller subsample sizes. Operationally, this implies that HTCB imputation will always obtain the lowest RMSE when using measured rather than subsampled and imputed HT. That said, if subsampling is deemed necessary, the greatest gains in efficiency are observed using a subsample size of two or more trees (Fig. 5) for the NFEMwC and NMEM methods, with rapidly diminishing returns thereafter.

Finally, our study is limited to the case of Douglas-fir trees in southwestern Oregon. Practitioners may be interested in the behavior of these methods for other species and using other data sets. We stress that our data set contains many sample trees (n = 2884), enabling a reliable assessment of the methods via the cross-validation simulation procedure. For smaller data sets, that may arise from rare species, limited field data or both, such a simulation procedure may not be tenable and produce measures of RMSE and bias with substantial second-order variances (e.g. Hastie et al., 2009 pp. 248–249). One alternative is to frame the problem of HT and HTCB imputation within the purview of model-based inference. For example, Bayesian hierarchical models for both HT and HTCB can be used to fully propagate the error from HT imputation through the HTCB model for a model-based measure of RMSE and bias that does not rely on cross-validation methods. Error propagation in forest inventories has been explored in many contexts, including uncertainty of carbon stock and biomass assessments (e.g., Chave et al., 2004; Holdaway et al., 2014; Chen et al., 2015). In the context of allometric models, Molto et al. (2013) linked a diameter-height model, a wood-specific gravity model and an above-ground biomass model using a Bayesian approach. Such an approach could be adapted using the HT and HTCB models we employed to assess prediction error without the need for cross-validation methods.

5. Conclusion

We can draw several broad conclusions from the results of this study. First, using a measured HT to predict HTCB will, on average, result in a more accurate prediction than if an imputed value for total height is used instead. If total heights must be imputed to predict HTCB, the use of a NFEM with an OLS corrections appear to suffer from high RMSE. In fact, if analytical capacity exists within an organization, significant predictive gains may be realized by fitting NMEMs for HT and HTCB imputation and localizing them with BLUPs based on a subsample of trees with HT and HTCB measured. Regardless of the method of localization, we recommend that a subsample of two or more trees be measured for each plot in order to maximize gains in predictive accuracy and avoid the significant biases observed at smaller subsample sizes, with lower RMSEs and bias achieved at higher subsampling rates, but at a diminishing rate.

CRediT authorship contribution statement

Elijah Allensworth: Conceptualization, Formal analysis, Writing - original draft, Methodology, Investigation. **Hailemariam Temesgen:** Resources, Supervision, Writing - review & editing, Project administration. **Bryce Frank:** Methodology, Formal analysis, Visualization, Investigation, Writing - review & editing. **Andrew Gray:** Writing - review & editing, Funding acquisition, Resources.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study was supported by a grant provided by USDA Forest Service, PNW Research Station: Fitting Biomass Equations for Pacific Northwest Tree Species: Agreement Number 18-JV-11261959-049.

References

Bianchi, S., Huuskonen, S., Siipilehto, J., Hynynen, J., 2020. Differences in tree growth of Norway spruce under rotation forestry and continuous cover forestry. For. Ecol. Manage. 458, 117689. https://doi.org/10.1016/j.foreco.2019.117689.

Calama, R., Montero, G., 2004. Interregional nonlinear height diameter model with random coefficients for stone pine in Spain. Can. J. For. Res. 34 (1), 150–163.

Chave, J., Condit, R., Aguilar, S., Hernandez, A., Lao, S., Perez, R., 2004. Error propagation and scaling for tropical forest biomass estimates. Philos. Trans. R. Soc. Lond. B Biol. Sci. 359 (1443), 409–420.

Chen, Q., Laurin, G.V., Valentini, R., 2015. Uncertainty of remotely sensed aboveground biomass over an African tropical forest: Propagating errors from trees to plots to pixels. Remote Sens. Environ. 160, 134–143.

Clark, T.S., Linzer, D.A., 2015. Should I use fixed or random effects. Political Sci. Res. Methods 3 (2), 399–408.

Colin, F., Houllier, F., 1992. Branchiness of Norway spruce in northeastern France: predicting the main crown characteristics from usual tree measurements. Annales des sciences forestières. EDP Sciences. 49 (5), 511–538.

D'Orazio, M., 2003. Estimating the variance of the sample mean in two-dimensional systematic sampling. J. Agric. Biol. Environ. Stat. 8 (3), 280.

Ducey, M.J., 2009. Predicting crown size and shape from simple stand variables. J. Sustainable For. 28 (1–2), 5–21.

Dyer, M.E., Burkhart, H.E., 1987. Compatible crown ratio and crown height models. Can. J. For. Res. 17 (6), 572–574 doi:10.1139/x87-096.

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 Sci. 55 (6), 480–493.

Fu, L., Zhang, H., Sharma, R.P., Pang, L., Wang, G., 2017. A generalized nonlinear mixed-effects height to crown base model for Mongolian oak in northeast China. For. Ecol. Manage. 384, 34–43. https://doi.org/10.1016/j.foreco.2016.09.012.

Garber, S.M., Temesgen, H., Monleon, V.J., Hann, D.W., 2009. Effects of height imputation strategies on stand volume estimation. Can. J. For. Res. 39 (3), 681–690.

Hann, D., 1982. Aerial Photo and Field Procedures for Selecting Candidate Stands for Growth Measurements. Oregon State University.

Hann, D.W., 2011. ORGANON User's Manual Edition 9.1. Oregon State University.

Hann, D.W., Scrivani, J.A., 1987. Dominant-height-growth and site-index equations for Douglas-fir and ponderosa pine in southwest Oregon. Oregon State University.

Hanus, M.L., Hann, D.W., Marshall, D.D., 2000. Predicting height to crown base for undamaged and damaged trees in Southwest Oregon. Oregon State University.

Hasenauer, H., Monserud, R.A., 1996. A crown ratio model for Austrian forests. For. Ecol. Manage. 84 (1–3), 49–60. Elsevier.

Hastie, T., Tibshirani, R., Friedman, J., 2009. The Elements of Statistical Learning: Data Mining, Inference, and Prediction, second ed. Springer Science & Business Media.

Holdaway, R.J., McNeill, S.J., Mason, N.W., Carswell, F.E., 2014. Propagating uncertainty in plot-based estimates of forest carbon stock and carbon stock change. Ecosystems 17 (4), 627–640.

Hynynen, J., 1995. Predicting tree crown ratio for unthinned and thinned Scots pine stands. Can. J. Forest Res. 25 (1), 57–62.

Lynch, T.B., Budhathoki, C., Wittwer, R.F., 2012. Relationships between height, diameter, and crown for eastern cottonwood (populus deltoides) in a Great Plains riparian ecosystem. Western J. Appl. Forestry 27 (4), 176–186.

Maguire, D.A., Hann, D.W., 1990. Constructing models for direct prediction of 5-year crown recession in southwestern Oregon Douglas-fir. Can. J. For. Res. 20 (7), 1044–1052, 10.1139/x90-139.

Maguire, D.A., Hann, D.W., Kershaw Jr, J.A., 1988. Prediction of branch diameter and branch distribution for Douglas-fir in southwestern Oregon. USDA Forest Service general technical report NC-North Central Forest Experiment Station (USA).

Molto, Q., Rossi, V., Blanc, L., Freckleton, R., 2013. Error propagation in biomass estimation in tropical forests. Methods Ecol. Evol. 4 (2), 175–183. https://doi.org/10.1111/mee3.2013.4.issue-210.1111/j.2041-210x.2012.00266.x.

Özçelik, R., Cao, Q.V., Trincado, G., Göçer, N., 2018. Predicting tree height from tree diameter and dominant height using mixed-effects and quantile regression models for two species in Turkey. Forest Ecol. Manage. 419, 240–248.

Paine, D.P., Hann, D.W., 1982. Maximum crown-width equations for southwestern Oregon tree species.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., R Core Team. 2021. nlme: Linear and Nonlinear Mixed Effects Models. Available from https://CRAN.R-project.org/package=nlme.

R Core Team, 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available from https://www.R-project.org/.

Raptis, D.I., Kazana, V., Kazaklis, A., Stamatiou, C., 2021. Mixed-effects height–diameter models for black pine (Pinus nigra Arn.) forest management. Trees 1–17.

Ritchie, M.W., Hamann, J.D., 2008. Individual-tree height-, diameter- and crown-width increment equations for young Douglas-fir plantations. New Forests 35 (2), 173–186. https://doi.org/10.1007/s11056-007-9070-7.

Ritchie, M.W., Hann, D.W., 1987. Equations for predicting height to crown base for fourteen tree species in southwest Oregon. Oregon State University.

Robinson, A.P., Wykoff, W.R., 2004. Imputing missing height measures using a mixed-effects modeling strategy. Can. J. For. Res. 34 (12), 2492–2500.

Sharma, R.P., Vacek, Z., Vacek, S., Podrázský, V., Jansa, V., 2017. Modelling individual tree height to crown base of Norway spruce (Picea abies (L.) Karst.) and European beech (Fagus sylvatica L.). PLOS ONE 12 (10), e0186394. https://doi.org/10.1371/journal.pone.0186394.

Soares, P., Tomé, M., 2001. A tree crown ratio prediction equation for eucalypt plantations. Ann. Forest Sci. 58 (2), 193–202.

Temesgen, H., Hann, D.W., Monleon, V.J., 2007. Regional height–diameter equations for major tree species of Southwest Oregon. Western J. Appl. Forestry 22 (3), 213–219. https://doi.org/10.1093/wjaf/22.3.213.

Temesgen, H., LeMay, V., Mitchell, S.J., 2005. Tree crown ratio models for multi-species and multi-layered stands of southeastern British Columbia. The Forestry Chronicle 81 (1), 133–141.

Temesgen, H., Monleon, V.J., Hann, D.W., 2008. Analysis and comparison of nonlinear tree height prediction strategies for Douglas-fir forests. Can. J. For. Res. 38 (3), 553–565.

Townsend, Z., Buckley, J., Harada, M., Scott, M.A., 2013. The choice between fixed and random effects. The SAGE handbook of multilevel modelling. SAGE Publications Ltd London, United Kingdom, pp. 73–88.

Valentine, H.T., Mäkelä, A., 2005. Bridging process-based and empirical approaches to modeling tree growth. Tree Physiol. 25 (7), 769–779.

Yang, Z., Liu, Q., Luo, P., Ye, Q., Sharma, R.P., Duan, G., Zhang, H., Fu, L., 2020. Nonlinear mixed-effects height to crown base model based on both airborne LiDAR and field datasets for Picea crassifolia Kom trees in northwest China. For. Ecol. Manage. 474, 118323.

Zhao, D., Kane, M., Borders, B.E., 2012. Crown ratio and relative spacing relationships for loblolly pine plantations. Open J. Forestry 2 (03), 101.