Calibration of volume and component biomass equations for Douglas-fir and lodgepole pine in Western Oregon forests

by Krishna P. Poudel[1] and Temesgen Hailemariam[1*]

ABSTRACT

Using data from destructively sampled Douglas-fir and lodgepole pine trees, we evaluated the performance of regional volume and component biomass equations in terms of bias and RMSE. The volume and component biomass equations were calibrated using three different adjustment methods that used: (a) a correction factor based on ordinary least square regression through origin (OLS-RTO method); (b) a correction factor based on OLS with intercept (OLS-WI method); and, (c) an inverse approach. The regional volume equations performed fairly well and produced similar results as the locally fitted volume equations of the same form but the regional predicted component biomass estimates were highly biased. All adjustment methods improved the performance of regional equations for the calibration dataset. Based on leave-one-out cross validation, the calibration based on OLS-RTO and OLS-WI methods reduced the RMSE for all species-component combinations. The inverse approach improved the performance of the regional equations for Douglas-fir but it did not improve lodgepole pine regional biomass equations. The decreasing trend of RMSE in component biomass estimation by using randomly selected trees to calibrate regional equations slowed down considerably after five trees.

Key words: volume equations, component biomass equations, Douglas-fir, lodgepole pine

RÉSUMÉ

Au moyen de données acquises par des tests destructifs sur des tiges-échantillons de sapin Douglas et de pin lodgepole, nous avons évalué la performance d'équations régionales de volume et des composantes de la biomasse quant au biais et à l'erreur moyenne quadratique. Les équations de volume et de composantes de la biomasse ont été calibrées selon trois méthodes d'ajustement différentes en utilisant : (a) un facteur de correction tiré de la méthode des moindres carrés ordinaires (méthode MCO-RTO); (b) un facteur de correction tiré d'une MCO avec étalonnage (méthode MCO-WI); et (c) une approche inverse. Les équations régionales de volume ont fonctionné de façon acceptable et ont généré des résultats similaires à ceux obtenus au moyen d'équations de volume de forme semblable ajustées aux conditions locales, mais l'estimation des composantes de biomasse régionales était fortement biaisée. Toutes les méthodes d'ajustement ont permis d'améliorer la performance des équations régionales après la calibration des données de base. Selon une validation croisée utilisant une méthode d'exclusion d'une donnée (leave-one-out cross-validation), la calibration basée sur les méthodes MCO-RTO et MCO-WI a réduit l'erreur moyenne quadratique de toutes les combinaisons espèces-composantes. L'approche inverse a amélioré la performance des équations régionales dans le cas du sapin de Douglas, mais n'a pas permis d'améliorer les équations régionales de biomasse du pin lodgepole. La tendance à la baisse de l'erreur moyenne quadratique dans l'estimation des composantes de biomasse en utilisant m arbres choisis au hasard pour calibrer les équations régionales, a considérablement diminué après cinq arbres.

Mots clés : équations de volumes, équations des composantes de biomasse, sapin de Douglas, pin lodgepole

Krishna P. Poudel
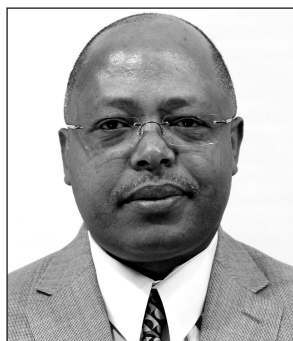
Temesgen Hailemariam

Introduction

The estimation of total and merchantable stem volume is crucial for land managers to ascertain continuous production of timber and other forest products. These estimates are obtained either by using direct volume equations or by the integration of stem profile equations. Volume equations are typically species-specific regression models that range from single entry simple linear regression models that relate tree volume with diameter at breast height (DBH), to multiple entry nonlinear regression models that use multiple explanatory variables such as DBH, height, and crown ratio.

Several published volume equations are available and are typically used at different scales of forest management. Generally these equations are developed for national or regional

[1]College of Forestry, Oregon State University, Corvallis, Oregon 97333; * corresponding author: hailemariam.temesgen@oregonstate.edu

scale applications but are also regularly used to predict tree volume at local scales. However, the differences in stand conditions affect tree form and consequently tree volume (Bluhm *et al.* 2007, Temesgen *et al.* 2015). The Forest Inventory and Analysis program in the Pacific Northwest Research Station (FIA-PNW) uses direct volume equations for tree volume estimation based on Brackett (1973), particularly for Douglas-fir and lodgepole pine. The tarif system is a set of previously constructed local volume tables that are applicable to a specific stand (Zhou and Hemstrom 2010). The accuracy of these equations is critical, not only because the estimate of tree volume is directly related to the economic aspect of forest management, but also because the calculation of other important variables such as biomass is dependent on the estimate of volume. However, the evaluation of existing volume equations is a rare opportunity (Omule *et al.* 1987) because it requires destructively sampled data.

Tree biomass or its components is defined as its oven-dry weight. The amount of biomass present in a forest indicates the amount of carbon dioxide that is stored and sequestered by that forest because the carbon content in biomass is approximately 50%. Various climate change agreements and action plans at scales ranging from local to international, such as the United Nations Framework Convention on Climate Change and in particular the Kyoto Protocol, recognize the importance of the forest carbon sink and the need to monitor, preserve and enhance terrestrial carbon stocks (Zianis *et al.* 2005). Additionally, a number of voluntary and regulated carbon markets have provisions for carbon sequestration credits from forest management projects that meet certain criteria and verification.

In addition to a total above-ground biomass estimate, knowledge of biomass in different components is useful for different purposes. Stem wood is important in timber sales; crown biomass is useful in fuel load assessment, formulating fire management strategies, and developing wildfire models. The biomass in small branches and leaves/foliage determines the possibility of installing bioenergy plants by providing information on the amount of available feedstock. The FIA-PNW calculates biomass of the main stem, whether merchantable or total, from the cubic volume estimates and previously compiled set of wood density factors. The biomass of other components is calculated from published models derived from local tree studies as functions of DBH and total tree height (Zhou and Hemstrom 2009).

Currently above-ground biomass in the United States is calculated using one of three methods: 1) the component ratio method (CRM) of the FIA; 2) equations developed by Jenkins *et al.* (2003) based on meta-analysis of previously published models; and, 3) models developed for regional applications. U.S. official carbon inventories are based on tree biomass estimates from the component ratio method described by Woodall *et al.* (2011), whereas understory vegetation biomass is from Jenkins *et al.* (2003) equations (U.S. EPA 2015). The literature on above-ground biomass estimation is voluminous. Early biomass studies were focused on fitting small-scale site- and species-specific allometric equations, while recent studies have focused more on fitting techniques and application of remote sensing techniques. Large-scale biomass studies based on remotely sensed data use estimates from one of these methods (FIA-CRM; Jenkins; regional equations) as the observed true value of biomass.

The accuracy of these methods is seldom tested against the observed biomass data obtained from destructive sampling. Instead, the estimates obtained from these methods have been compared to one another. For example, Zhou and Hemstrom (2009) found similar estimates of total above-ground biomass from these methods but substantially different estimates for the merchantable biomass for softwood species in Oregon. Jenkins equations produced estimates that were 17% higher than the estimates from regional equations. Similarly, Poudel (2015) found that the Jenkins method for Douglas-fir produced total above-ground biomass that was 18.4% and 23.7% higher than estimates provided by the FIA-PNW and FIA-CRM methods respectively.

Accurate estimates of above-ground biomass are needed to reduce uncertainties in global and regional terrestrial carbon fluxes (Houghton *et al.* 2009). Direct measurement of biomass on the ground is time-consuming and expensive (Houghton 2005) and the component biomass estimates obtained as percentages of total or stem biomass from published information might be very simplistic (Hansen 2002). Similar to volume equations, the evaluation of biomass models is also time-consuming and expensive because it requires destructively sampled data.

Zeng *et al.* (2011) used a dummy variable model approach to construct compatible single-tree biomass equations at different scales. They found these models with local parameters to perform better than a population-average model. Räty and Kangas (2008) noted that regional models may not be unbiased at the local scale if there is spatial variation in tree form due to one or more unknown predictors, and this regional bias could be reduced or removed if the models are localized to each sub-region or subarea.

Model calibration is needed when the predictive validity of a model is in question or when the data are inadequate to estimate model inputs. It is the process of systematic adjustment of model parameters such that the adjusted model predicts the observed outputs more accurately (Taylor *et al.* 2012). Two options are available when sample volume and biomass data are available at a local scale: 1) fit new volume or biomass equation; or, 2) use the available data to calibrate existing regional models. However, the sample size required for fitting new models is relatively larger than that required for calibration of existing equations (Garber *et al.* 2009). Calibration itself is sometimes criticized because it requires destructive sampling if the original model formation is based on destructive sampling (Ketterings *et al.* 2001).

Calibration of large-scale prediction models is not new to forestry. Sadiq and Smith (1983) calibrated volume equations; Eerikäinen *et al.* (2002) the site index model; Lappi (1991), Jayaraman and Zakrzewski (2001), Temesgen *et al.* (2008), and Garber *et al.* (2009) height-diameter equations. The calibration of biomass models by de-Miguel *et al.* (2014) used simulated data. These studies employed mixed effects models in which the mean effect is described by a general functional form and the variability due to stand or plot effects is accounted by the random effect parameter. The calibration process in this approach is based on the best linear unbiased predictor of the random effects parameter. Temesgen *et al.* (2008) used this approach to calibrate nonlinear height-diameter equations while Garber *et al.* (2009) used it in evaluating the effects of height imputation strategies on stand volume estimation. Another commonly used method for calibrating

regional models is based on an ordinary least square (OLS) correction factor. A multiplicative correction factor based on an OLS regression is obtained through measured and regional predicted values of variable of interest (e.g., Hanus *et al.* 1999, Hann 2006, Temesgen *et al.* 2008, Garber *et al.* 2009). Temesgen *et al.* (2008) obtained results from OLS correction factors that were comparable with the results from mixed effects model when four or more trees were subsampled for height while Garber *et al.* (2009) obtained 95% decrease in root mean squared error (RMSE) when the regional models were calibrated with a subsample of three tree heights.

Development of site- and species-specific biomass equations is expensive. Therefore, there is a need for research on calibration of volume and component biomass equations. The objectives of this study are to evaluate the performance of (1) existing volume and biomass equations in terms of their prediction errors; and (2) three different calibration methods that use a correction factor based on (a) OLS regression through origin (OLS-RTO method), (b) OLS with intercept (OLS-WI method), and (c) an inverse approach.

Materials and Methods
Data
A detailed biomass data collection was carried out by destructively sampling twenty-two Douglas-fir (*Pseudotsuga menziesii* (Mirbel) Franco) and thirty-three lodgepole pine (*Pinus contorta* Dougl. ex Loud.) trees in different forests within the state of Oregon. Efforts were made to select trees that gave an approximately equal representation across the range of size class while avoiding ones with severe defects and/or close to stand boundaries. Trees that were forked below breast height and trees with damaged tops were also not included in sampling (Poudel *et al.* 2015). The field work was carried out between the first week of July and third week of September 2012 and 2013. Tree-level attributes such as DBH, total height, height to the base of first live branch, crown width, and stem diameter outside-bark (DOB) at 0.15 m, 0.76 m, 1.37 m, 2.44 m above ground, and every 1.22 m afterwards were recorded. For all first-order branches (branches directly attached to the main bole), their height above ground and diameter at their base were recorded for a total of 2 975 Douglas-fir and 4 458 lodgepole pine branches. Average DBH was 54.9 cm (range 19.3 cm–114.0 cm) and 24.6 cm (range 13.5 cm–42.9 cm) and average height was 33 m (range 16.6 m–48.8 m) and 17 m (range 9.2 m–31.9 m) for Douglas-fir and lodgepole pine respectively.

Data compilation
The observed inside-bark cubic volume, including top and stump (CVTS), was calculated as follows. After the subject tree was felled, it was sectioned into 5.18 m long sections. Three to five centimetre thick disks were removed from the top of the stump and at every 5.18 metres; diameter-inside bark at both ends of the 5.18 m sections were also obtained. The diameter-inside bark at stem heights where DOBs were measured were obtained by linear interpolation of bark thickness ratios at section ends. The inside-bark cubic volume of these sections was calculated using Smalian's formula below:

$$(1) \quad V = \frac{A_1 + A_2}{2} \times L$$

Where V is the volume of the section of length L in cubic metres, $A_i's$ ($i=1, 2$) are the cross-sectional areas at the small and large end of the sections in m^2 with $A_i = \frac{\pi D_i^2}{4}$ and $D_i =$ diameter inside-bark (m) at the ith end of the section.

Inside-bark volume of the stump was calculated as a cylinder whereas the volume of the top section was calculated as a cone.

$$(2) \quad V_s = \frac{\pi D_s^2 H_s}{4}$$

$$(3) \quad V_t = \frac{\pi D_t^2 H_t}{12}$$

Where V_s and V_t are inside-bark volume of stump and top, respectively, in cubic metres, D_s and D_t are the inside-bark diameter (m) at the stump top and bottom of the top section, respectively, and H_s and H_t are the stump height and length of top section in metres, respectively. Total inside-bark CVTS was calculated by summing the inside-bark volumes of all sections, stump, and top.

In the field, green weight of the disks and their diameter inside- and outside-bark as well as four measurements 90° apart along the disk edge for thickness were recorded. Green volume was calculated as the volume of a cylinder. Once in the lab, the whole disks were dried in a kiln at 105°C until constant weight was recorded. Disk densities were then obtained as the ratio between dry-weight to green volume. The volume of each 5.18 m section was converted into biomass by multiplying it by the average density of the disks taken from two ends. Biomass in stump and top section was calculated by converting respective volumes using the densities of the disks taken from top of the stump and top of the last 5.18 m log respectively. Total bole biomass was the sum of the section masses. The outside-bark CVTS was calculated in similar fashion as the inside-bark CVTS using diameter outside-bark instead of diameter inside-bark. Bark volume for each of the 5.18 m sections was obtained by subtracting inside-bark volume from outside-bark volume. Section bark biomass was calculated by multiplying section bark volume by bark density obtained as the ratio between oven dry-weight to green volume of 5- to 10-cm long rectangular bark samples taken from each disk. Bark samples were dried at 70°C. Total stem bark biomass was obtained by summation.

The crowns of the sample trees were divided into three equal length strata. Four, three, and two first-order branches respectively from lower, middle, and upper stratum were then randomly selected for weighing with and without foliage. The needles from those branches were removed in the field to obtain separate green weights of bare branch with bark and with foliage. The branches were brought to lab, keeping branch wood and foliage in separate paper bags for drying. The branches were chipped into small pieces and dried in a kiln at 105°C while the foliage samples were dried without chipping at 70°C. Oven dry-weight was recorded by tracking the weight lost by each sample until no further weight was lost. Individual branch wood and foliage biomass was obtained by fitting species-specific log linear model of the following form:

$$(4) \quad \ln(y_{ij}) = \beta_{0i} + \beta_{1i}\ln(BD_{ij}) + \varepsilon_{ij}$$

Where y_{ij}, and BD_{ij} are oven dry-weight (kg) of branch wood or foliage and branch diameter (cm) at base of the jth branch on i^{th} tree respectively; $\beta_{ij}'s$ are regression parameters to be estimated from the data; $\ln(\cdot)$ is the natural logarithm; $\varepsilon_{ij}'s$ are the random error. Similar model form has been used by Temesgen *et al.* (2011) to estimate foliar biomass in Douglas-fir and ponderosa pine and by Poudel *et al.* (2015) to estimate crown biomass in Douglas-fir. Total branch wood and foliage biomasses were obtained by summing individual branch wood and foliage biomass.

Regional equations

Table 1 summarizes the sets of volume and biomass equations used by FIA-PNW to obtain CVTS and component biomass. The FIA-PNW does not have an equation for foliage biomass (Melson *et al.* 2011) and our FIA-PNW equivalent foliage biomass estimates were based on the foliage biomass equation of Standish *et al.* (1985). The FIA-PNW volume and biomass equations were also refitted with our dataset to compare the differences in published coefficients against the coefficients of refitted models of the same form.

Calibration of regional models

Model calibration is needed when the predictive validity of a model is in question or when the data are inadequate to estimate model inputs. We first graphically compared the regression lines produced by regional volume equations with locally refitted volumes equations of the same form. The empirical cumulative distribution function (ECDF) of the component biomass estimates obtained from the FIA-PNW equations and the ECDFs of the measured volume and biomass estimates in different above-ground components were also graphically compared. The ECDF is the cumulative distribution function associated with the empirical measure of the sample observations. In our case, it shows the percentage of trees in the sample that have less than or equal to specific values of biomass. For independent and identically distributed random variables $x_1, ..., x_n$, the ECDF is defined as:

$$(15) \quad \hat{F}(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}\{X_i \le x\}$$

Where, $\mathbf{1}\{A\}$ is the indicator of an event A.

Three different methods of calibrating regional equations were then applied to adjust the regional prediction obtained from the FIA-PNW equations:

OLS-RTO Method

If a sample of volume and biomass estimates for n new trees is available, the regional equations can be calibrated using a

Table 1. Volume and biomass equations used by FIA-PNW to estimate cubic volume including top and stump (CVTS,m³) and component of above-ground biomass (kg) for Douglas-fir and lodgepole pine.

Species	Volume/Biomass Equation

Douglas-fir

$$CVTSL = \begin{matrix}-3.21809 + 0.04948 \times \log(HT) \times \log(DBH) - 0.15664 \times (\log(DBH))^2 \\ +2.02132 \times \log(DBH) + 1.63408 \times \log(HT) - 0.1685 \times (\log(HT))^2\end{matrix} \quad (5)$$
$$CVTS = 10^{CVTSL}$$

$$Bole = CVTS \times WD \quad (6)$$
$$Foliage = 10.3 + 3.9 \times DBH^2 \times HT \quad (7)$$
$$Bark = 3.1 + 15.6 \times DBH^2 \times HT \quad (8)$$
$$Branch = 12.6 + 23.5 \times DBH^2 \times HT \quad (9)$$

Lodgepole Pine

$$CVTSL = -2.615591 + 1.847504 \times \log(DBH) + 1.085772 \times \log(HT) \quad (10)$$
$$CVTS = 10^{CVTSL}$$
$$Bole = CVTS \times WD \quad (11)$$
$$Foliage = 5.5 + 4.0 \times DBH^2 \times HT \quad (12)$$
$$Bark = 3.2 + 9.1 \times DBH^2 \times HT \quad (13)$$
$$Branch = 7.8 + 12.3 \times DBH^2 \times HT \quad (14)$$

(CVTSL= logarithm of CVTS, base of the logarithm is 10, HT = total tree height (m), and DBH = diameter at breast height (cm); WD is wood density which is 459.7 and 370.0 kilogram per cubic meter for Douglas-fir and lodgepole pine respectively)

correction factor based on an ordinary least squares regression through origin. The technique is described in Draper and Smith (1998) and has been implemented by Temesgen *et al.* (2008) to calibrate nonlinear height-diameter models and by Garber *et al.* (2009) to estimate stand volume. When n sample trees from a new stand are available, the correction factor based on OLS-RTO can be calculated as:

$$(16) \quad k = \frac{\sum_{i=1}^{n} RP_i M_i}{\sum_{i=1}^{n} RP_i^2}$$

Where, M_i and RP_i are measured and regional predicted values of volume or component biomass for the ith tree respectively. Then the corrected (calibrated) value of volume or component biomass for a new tree can be calculated as $\widehat{M_i} = k \times RP_i$. This OLS correction factor can be obtained based on the measurement of a single tree. The correction factor is just a ratio of measured and regional predicted volume or biomass of that tree.

OLS-WI Method

With M_i and RP_i defined as in the OLS-RTO method, the relation between M_i and RP_i can be explained by a simple linear regression model with the following form:

$$(17) \quad M_i = \alpha + \beta RP_i + \varepsilon_i$$

Where α and β are the parameters in the linear regression and ε_i the random error. Then with the least squares estimators a and b of α and β respectively, the corresponding calibration equation is:

$$(18) \quad \widehat{M_i} = a + b * RP_i$$

Thus, if one obtains RP_i as the regional predicted value of volume or biomass, the unknown value (M_i) of volume or biomass estimates for the new tree can be estimated using equation 18. Note that this method violates the assumption of simple linear regression that the independent variable is measured without error because the regional predicted estimate comes with its associated error.

Inverse Approach

The inverse approach of model calibration has been described by Kutner *et al* (2004). By switching the dependent and independent variables in equation (17), the simple linear regression model can be written in the following form:

$$(19) \quad RP_i = \gamma + \delta M_i + \varepsilon_i'$$

Then with the least squares estimators c and d of γ and δ respectively, the fitted values for can be calculated as

$\widehat{RP}_i = c + d * M_i$. Thus, if now one obtains as the regional predicted value of volume or biomass, the unknown value (M_i) volume or biomass for the new tree can be estimated as:

$$(20) \quad \widehat{M}_\iota = \frac{RP_i - c}{d}$$

The performance of all the methods was evaluated based on the bias, bias percent, RMSE, and RMSE percent (collectively called "evaluation statistics") produced by each method. In statistics, bias is defined as the difference between the true value of an unknown parameter and the expected value of its estimator. Bias in this study is defined as the mean difference between the measured and model predicted value of the vari-able of interest. Evaluation statistics were obtained using the leave-one-out cross validation approach (Stone 1974). One tree was removed from the data, a correction factor obtained based on the remaining $n-1$ observations and the process repeated until all the observations have been removed and the evaluation statistics were calculated.

Results and Discussion
CVTS

The volume estimates obtained from FIA-PNW equations with their original coefficients and refitted FIA-PNW equations were similar for Douglas-fir trees that were smaller than approximately 90 cm DBH. For trees larger than 90 cm DBH, the original FIA-PNW equation slightly underestimated the cubic volume (Fig. 1). On the other hand, for lodgepole pine, the original FIA-PNW equation consistently over-estimated CVTS even for trees that were 25 cm DBH (Fig. 1), compared to the refitted PNW equation. The evaluation statistics obtained from these methods are presented in Table 2.

The FIA-PNW volume equation for Douglas-fir is a six-parameter multiple linear regression model (Table 1, equation 5). When refitting this equation to our dataset, only the intercept term was statistically significant at 5% level of significance and the DBH was significant at 10% level of significance. Three out of the six coefficients also had different signs than that of the original coefficient. Therefore, we also fitted a simple volume equation (local volume equation) for Douglas-fir as a function of DBH and height. The refitted PNW equation reduced RMSE in Douglas-fir CVTS estimates by 3% compared to the PNW equation with original coefficients. The simple three-parameter local volume equation produced similar RMSE compared to the six-parameter PNW equation
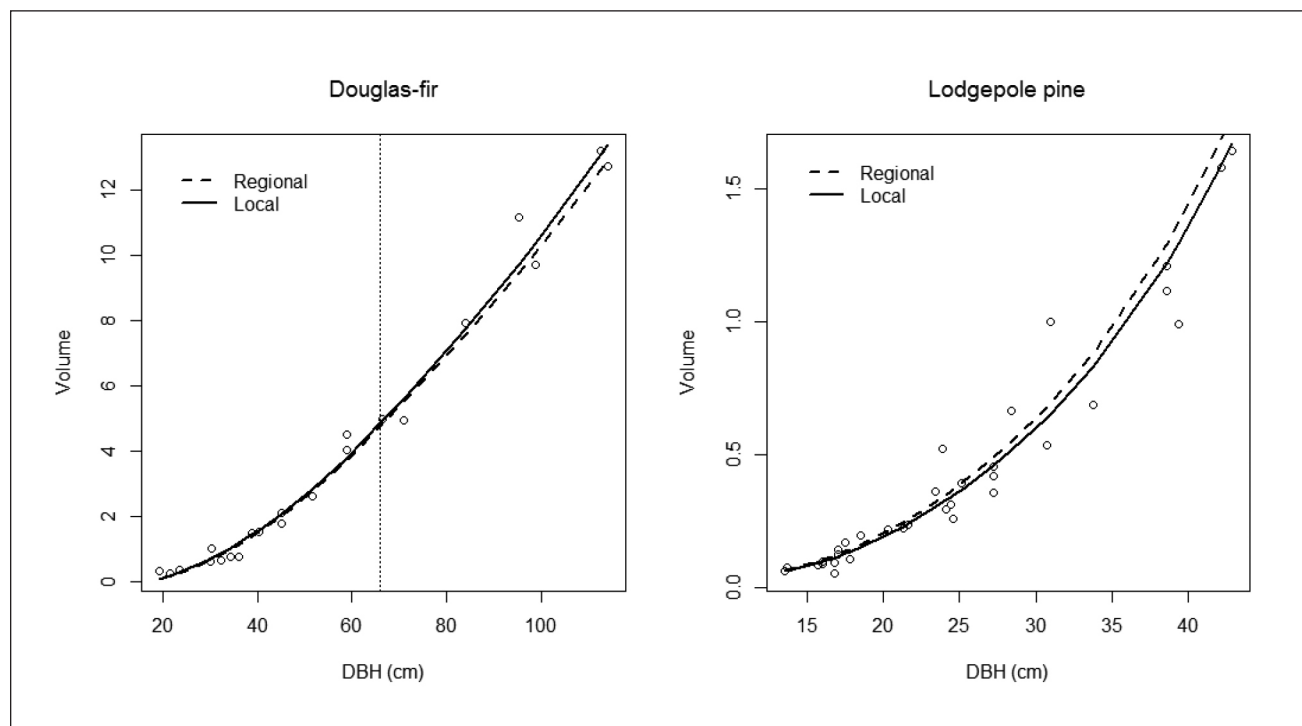
Fig. 1. Graphical comparison of CVTS estimates obtained by using regional model coefficients and locally refitted volume equations of the same form.

Table 2. Evaluation statistics produced by regional, refitted regional, and calibrated equations in estimating inside bark CVTS predictions for Douglas-fir and lodgepole pine.

| Species	Method	Bias (m^3)	Bias Percent	RMSE (m^3)	RMSE Percent
Douglas fir	PNW-Refit	0.0093	0.2373	0.5335	13.6148
	PNW	0.0625	1.5950	0.6489	16.5598
	Local	-0.0091	-0.2322	0.6383	16.2893
	OLS-RTO	-0.0321	-0.8192	0.6344	16.1897
	OLS-WI	0.0000	0.0000	0.6328	16.1489
	Inverse	0.0000	0.0000	0.6401	16.3352
Lodgepole pine	PNW-Refit	-0.0011	-0.2555	0.0555	12.8898
	PNW	-0.0471	-10.9389	0.1004	23.3178
	OLS-RTO	0.0112	2.6012	0.0570	13.2382
	OLS-WI	0.0000	0.0000	0.0547	12.7040
	Inverse	0.0000	0.0000	0.0552	12.8201

with original coefficients (Table 2). For lodgepole pine, the refitted PNW equation reduced RMSE by 10.4%.

The regional volume equations were calibrated using three different methods. In Douglas-fir, the calibrated volume equations also produced similar RMSE percent as the PNW volume equation with original coefficients. This was expected due to the fact that the CVTS estimates obtained from PNW equations with original coefficients were similar to those obtained from the refitted PNW equation for Douglas-fir, hence the calibration would not result in any improvement. However, in case of lodgepole pine, the calibrated PNW equation performed as well as the refitted PNW equation, reducing RMSE by up to 10.6%. The performance of the different calibration methods was similar for both species (Table 2).
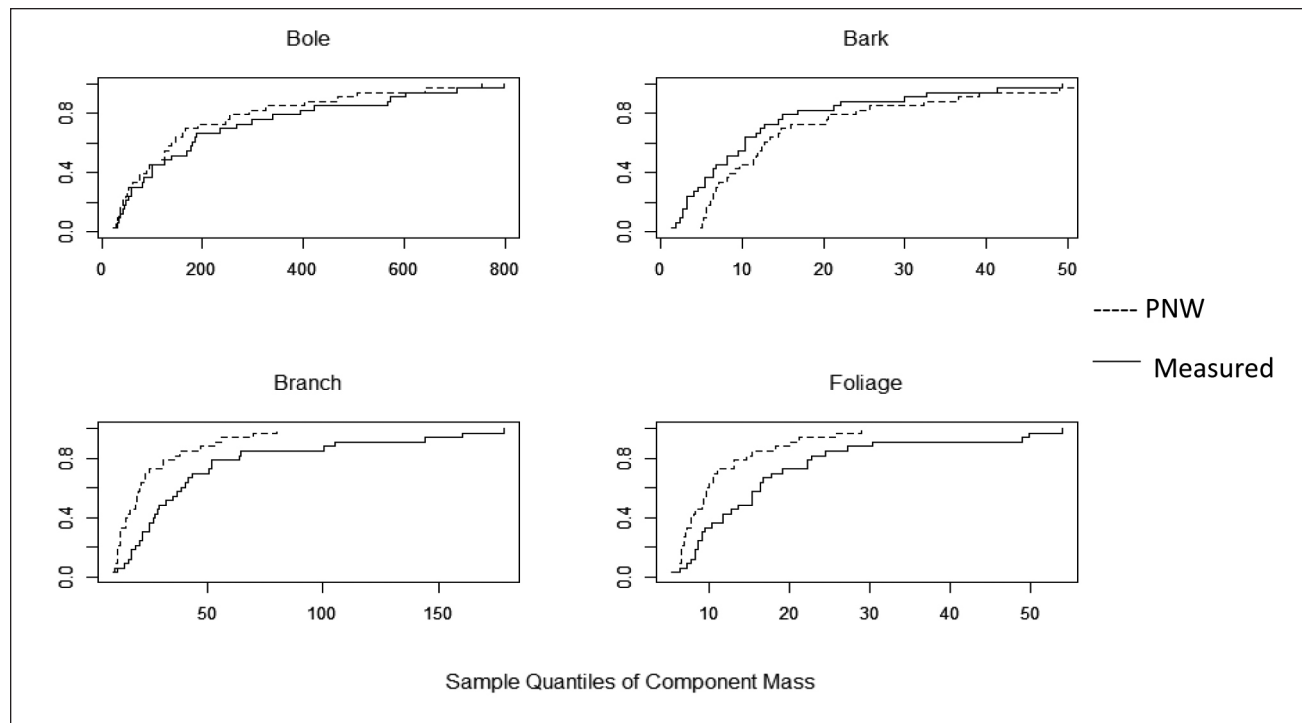
Biomass

The FIA-PNW bole biomass estimates are products of the inside-bark CVTS and the compiled set of species specific wood density values. Thus the curves of inside-bark bole biomass estimates against DBH are similar to inside bark CVTS curves but scaled by wood density. In other words, they have the same shape but different scales. The ECDFs of component biomasses obtained from regional models and measured component biomasses differed for both species, implying that the probability of observing a tree with less than or equal to certain amounts of component biomass based on FIA-PNW equations is different than the actual probability, based on the sample data measured in this study. These differences were more pronounced in components other than in the bole (Fig. 2 and Fig. 3).

Since the FIA-PNW does not have a specific equation for inside-bark bole biomass, we fitted a local bole biomass equation as a function of DBH and total tree height. The evaluation statistics obtained from regional and local or locally refitted component biomass equations are given in Table 3. The FIA-PNW Douglas-fir bark biomass equation does a very poor job and over estimates it by 221.5%. This difference could possibly be attributed to the fact that the FIA-PNW bark biomass equation for Douglas-fir was based on data with maximum DBH of 66 cm. Hence, a careful consideration might be necessary in applying the regional models to estimate bark biomass in Douglas-fir.
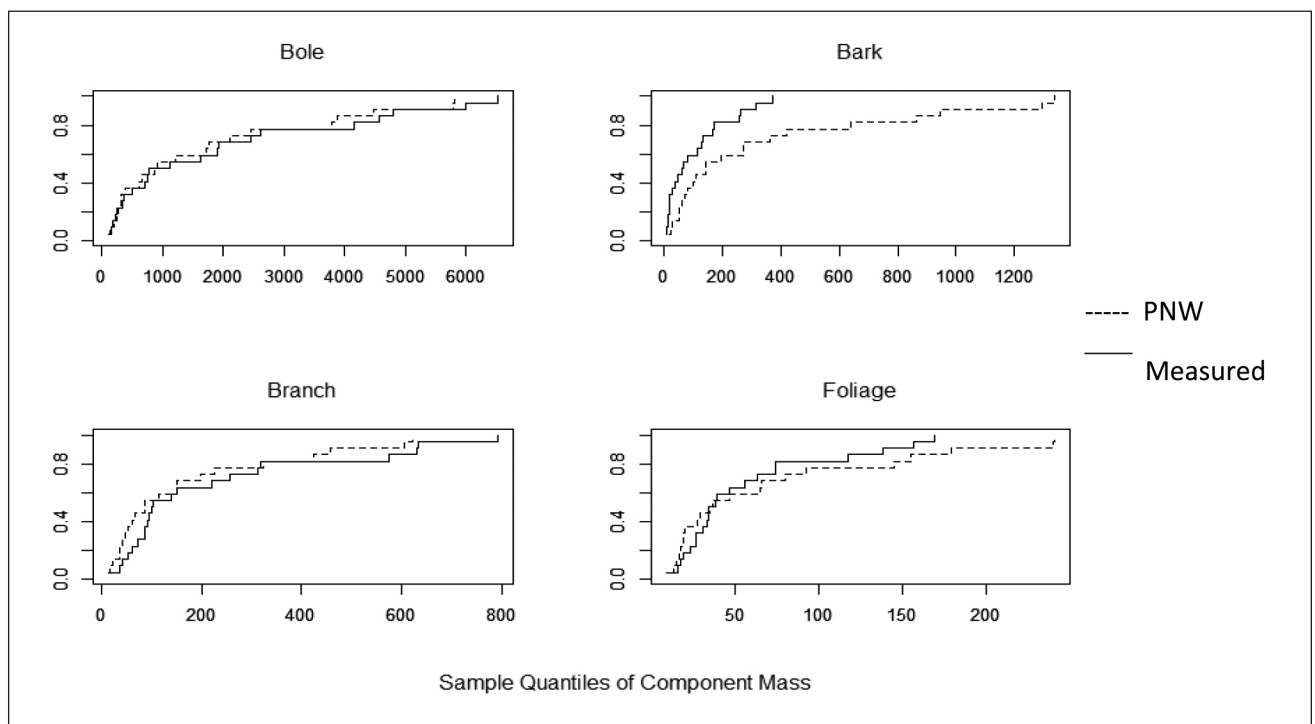
Fig. 2. Comparison of ECDFs of Douglas-fir component biomass estimates obtained from regional equations (dotted line) and measured (solid line) values. The ECDF shows the percentage of trees in the sample that have less than or equal to certain amount of biomass.

Fig. 3. Comparison of ECDFs of lodgepole pine component biomass estimates obtained from regional equations (dotted line) and measured (solid line) values. The ECDF shows the percentage of trees in the sample that have less than or equal to certain amount of biomass.

Table 3. Evaluation statistics in volume and component biomass estimation obtained from regional and locally refitted volume and component biomass equations.

Species	Variable	Local/PNW-Refit				PNW			
		Bias (kg)	Bias Percent	RMSE (kg)	RMSE Percent	Bias (kg)	Bias Percent	RMSE (kg)	RMSE Percent
Douglas fir									
	Bole	-1.52	-0.08	229.94	12.00	182.83	9.54	311.68	16.26
	Bark	-2.49	-2.36	27.82	26.32	-234.16	-221.52	382.25	361.60
	Foliage	0.01	0.02	18.41	32.38	-14.89	-26.18	37.66	66.22
	Branch	0.94	0.42	66.11	29.82	47.05	21.23	85.86	38.73
Lodgepole pine									
	Bole	1.99	0.90	68.56	31.14	38.83	17.64	78.71	35.75
	Bark	0.00	-0.01	7.38	61.47	-3.87	-32.21	9.39	78.14
	Foliage	0.00	0.01	9.96	56.40	6.59	37.31	12.02	68.07
	Branch	-0.01	-0.02	32.42	68.10	22.67	47.61	40.65	85.38

OLS-RTO Method

The evaluation statistics obtained by using a correction factor based on the OLS-RTO method in calibrating regional component biomass equations obtained from the leave-one-out cross validation method are given in Table 4.

The values of OLS-RTO correction factor ranged from 0.0277 (for Douglas-fir bark biomass) to 1.7849 (for lodgepole pine branch biomass). Results from the leave-one-out cross validation showed that this correction factor reduced RMSE in Douglas-fir bole, bark, foliage, and branch biomass by 4.4%, 325.4%, 24.1%, and 2.5%, respectively (see cautionary note about bark biomass estimation in Summary and Conclusion). For lodgepole pine, RMSE was reduced by 2.1%,

7.4%, 6.2%, and 10.1% for bole, bark, foliage, and branch biomass respectively. The performance of the calibration based on an OLS-RTO correction factor calculated using randomly selected Douglas-fir and lodgepole pine trees is shown in Fig. 4 and Fig. 5 respectively.

The RMSE decreases considerably by using two trees compared to just one and the rate of decrease in RMSE slows down after five trees for both species. It should be noted that measuring biomass in a large number of trees to calibrate regional models is impractical (Temesgen *et al.* 2008) because one would rather fit a local model than to calibrate a regional model with the available data set.

Table 4. Evaluation statistics from leave one out cross validation in component biomass estimates obtained from OLS-RTO adjustment of the regional component biomass equations. Values of k are obtained as the average of k values obtained from the leave one out cross validation approach.

Species	Variable	k	Bias (kg)	Bias Percent	RMSE (kg)	RMSE Percent
Douglas-fir						
	Bole	1.0955	15.15	0.79	227.18	11.85
	Bark	0.2770	10.78	10.20	38.30	36.24
	Foliage	0.6920	6.85	12.04	23.97	42.15
	Branch	1.2061	10.26	4.63	80.37	36.25
Lodgepole pine						
	Bole	1.1598	8.88	4.03	74.05	33.63
	Bark	0.7215	0.47	3.92	8.49	70.71
	Foliage	1.5212	0.74	4.17	10.94	61.91
	Branch	1.7849	2.75	5.77	35.86	75.32

OLS-WI Method

The evaluation statistics obtained with OLS-WI adjustment based on leave-one-out cross validation are presented in Table 5. The absolute bias obtained from the leave-one-out cross validation was less than 2% for all component biomass and as expected, it produced unbiased estimates of component biomasses for the dataset on which the calibration equation is based. Compared to the unadjusted PNW component biomass equations, the RMSE in bole, bark, foliage, and branch biomass was reduced by 4.25%, 328.24%, 27.89%, and 2.25%

for Douglas-fir and by 1.88%, 3.45%, 4.91%, and 9.5% for lodgepole pine respectively.

Inverse Approach

The inverse approach of the calibration of regional equations also provided unbiased estimates of component biomass for the data on which the calibrating equation was based. Additionally, the absolute bias obtained from the leave-one-out cross validation was less than 2% for all component biomass from this approach (Table 6). Calibrating regional equations with inverse regression improved the performance of these equations for Douglas-fir but not for lodgepole pine.

It is expected that the calibrated models perform better than the unadjusted regional models. However, from a practical perspective, it is desirable that the calibrated models perform as well as locally fitted models. The OLS-RTO adjustment performed as well as the locally refitted models for CVTS and component biomass models except for Douglas-fir foliage biomass, when all data is used to calculate the correction factor. However, the results from leave-one-out cross validation showed that the locally refitted models performed better than the calibrated models.

Even though having an intercept in the calibration equation provides more flexibility in adjustment, the performance of OLS-WI was similar to the OLS-RTO for most of the
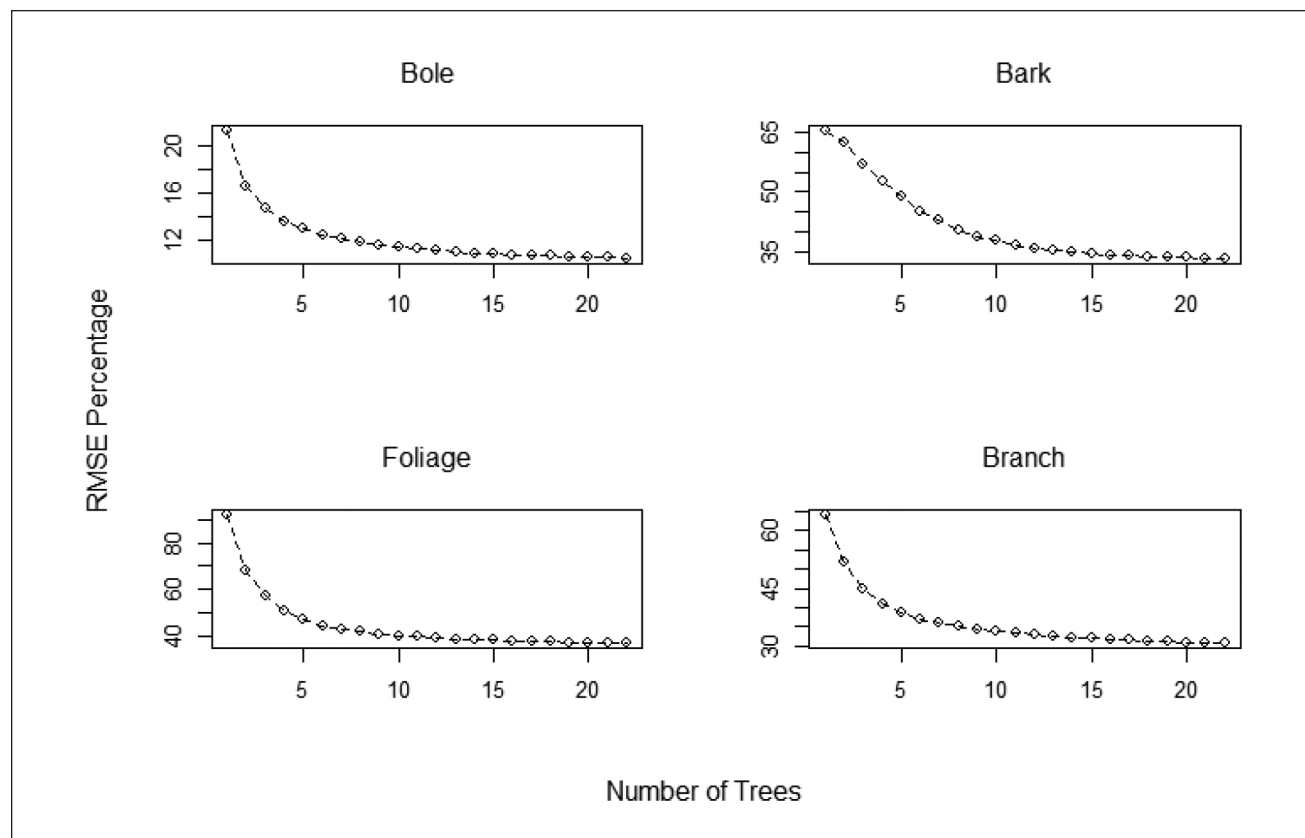
Fig. 4. Trend in RMSE percent in estimating component biomass from regional models adjusted with OLS-RTO method calculated from randomly selected m Douglas-fir trees. The process of selecting m trees was repeated for 5000 times. The RMSE percent from unadjusted regional equations were 16.3%, 361.6%, 66.2%, and 38.7% for bole, bark, foliage, and branch biomass respectively.
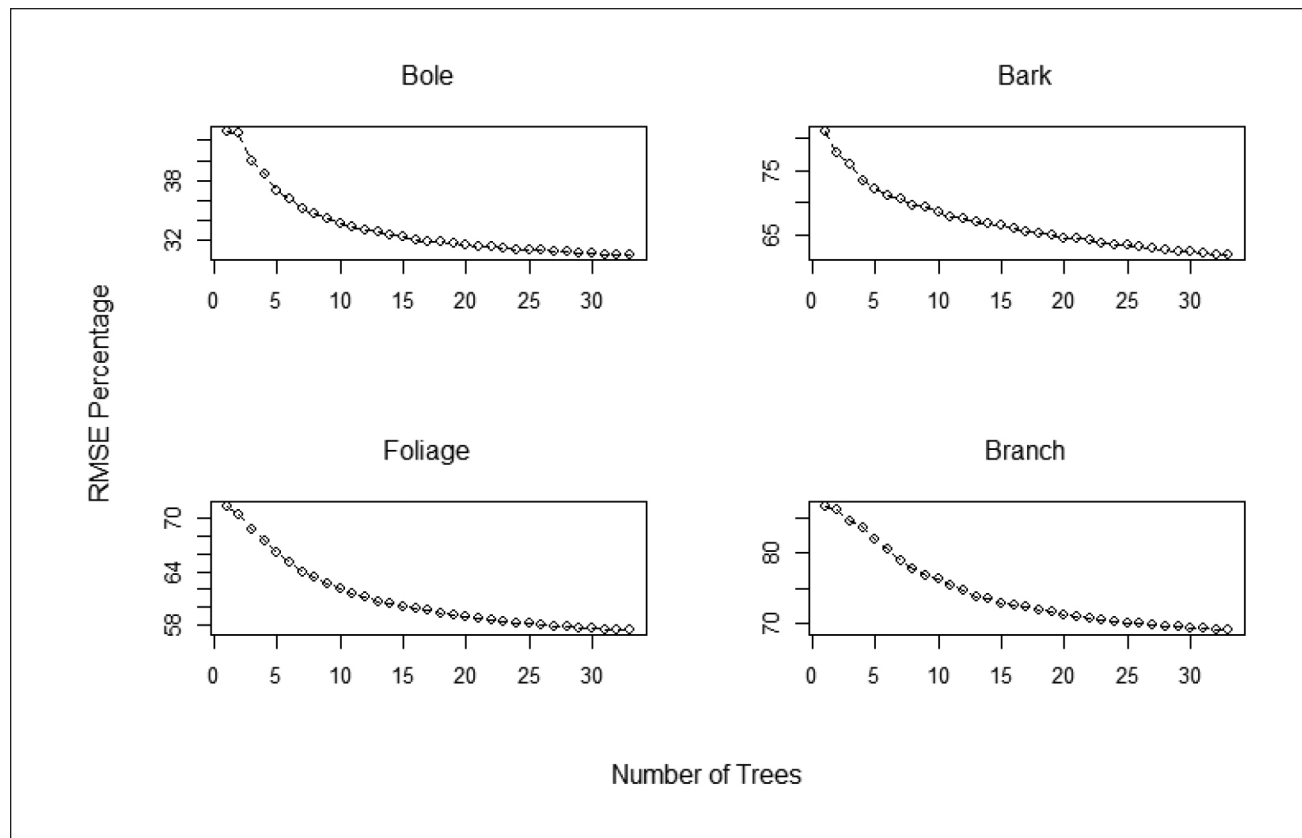
Fig. 5. Trend in RMSE percent in estimating component biomass from regional models adjusted with OLS-RTO method calculated from randomly selected m lodgepole pine trees. The process of selecting m trees was repeated 5000 times. The RMSE percent from unadjusted regional equations were 35.8%, 78.1%, 68.1%, and 85.4% for bole, bark, foliage, and branch biomass respectively.

Table 5. Evaluation statistics from leave one out cross validation in component biomass estimates obtained from **OLS-WI** adjustment of the regional component biomass equations. Values of a and b are the average of a and b values obtained from the leave-one-cross validation.

Species	Variable	*a*	*b*	Bias (kg)	Bias Percent	RMSE (kg)	RMSE Percent
Douglas-fir							
	Bole	33.3182	1.0863	-2.98	-0.16	230.22	12.01
	Bark	19.9250	0.2525	-0.69	-0.65	35.26	33.36
	Foliage	14.6011	0.5892	-0.20	-0.35	21.80	38.33
	Branch	20.9354	1.1498	-0.83	-0.37	80.86	36.48
Lodgepole pine							
	Bole	19.7394	1.1055	-1.45	-0.66	74.58	33.87
	Bark	1.3806	0.6698	-0.18	-1.51	8.97	74.69
	Foliage	3.8887	1.2446	-0.26	-1.47	11.16	63.16
	Branch	9.3362	1.5354	-0.74	-1.56	36.12	75.88

species component combinations. The OLS-WI for the calibration data performed as well as local models for bole and foliage biomass in Douglas-fir but its performance was intermediate between FIA-PNW estimates and locally refitted model for other species component combinations. The results from leave-one out cross validation showed that the inverse regression produced higher RMSE compared to the locally refitted regional models.

Summary and Conclusions

We evaluated the performance of regional volume and component biomass equations in terms of bias and RMSE using data from destructively sampled Douglas-fir and lodgepole pine trees. Three adjustment methods based on ordinary least squares regression were applied. The regional volume equations performed fairly well and produced similar results as the locally refitted volume equations of the same form.

The regional component biomass estimates were highly biased, producing up to 360% RMSE for bark biomass for Douglas-fir. We would like to make a cautionary note because our calculation of observed bark biomass is based on a 3 cm – 5 cm long sample obtained from the disks at every 5.18 m intervals. The density of these bark samples is based on their volume calculated as rectangles. This could have been an erroneous assumption since the bark sample lacks regular shape. To check this, we calculated total bark volume by subtracting inside-bark CVTS from outside-bark CVTS and used the compiled sets of density factors (same as the ones used by the FIA-PNW). The total bark biomass estimates obtained from this approach was similar to what we had obtained before. A better way to estimate bark volume

Table 6. Evaluation statistics from leave one out cross validation in component biomass estimates obtained from inverse regression approach of the regional component biomass equations. Values of c and d are the average of c and d values obtained from the leave-one-cross validation.

Species	Variable	c	d	Bias (kg)	Bias Percent	RMSE (kg)	RMSE Percent
Douglas-fir							
	Bole	-12.8586	0.9113	-3.07	-0.16	232.28	12.12
	Bark	-41.4628	3.6063	-1.22	-1.15	37.46	35.44
	Foliage	-9.2455	1.4245	-0.26	-0.47	23.28	40.93
	Branch	-1.374	0.794	-0.66	-0.30	81.94	36.96
Lodgepole pine							
	Bole	1.3692	0.8173	-1.96	-0.89	79.16	35.95
	Bark	5.4361	0.8699	-0.10	-0.84	10.55	87.81
	Foliage	6.281	0.2715	-0.11	-0.63	18.76	106.20
	Branch	12.2349	0.2671	-0.65	-1.37	56.82	119.36

Table 7. RMSE percent obtained by using OLS-RTO method based on a single tree with specific characteristics in calibrating regional volume and component biomass equations for Douglas-fir.

Variable	Dominant tree	Tree with largest DBH
CVTS	18.78	15.34
Bole	10.45	11.08
Bark	72.97	33.33
Foliage	48.44	51.74
Branch	40.14	44.85

would be to use a water displacement method but that is both time-consuming and expensive compared to the method we used.

All the adjustment methods used in this study improved the performance of regional equations for the calibration data. The calibration of equations for Douglas-fir did not reduce the RMSE substantially. Based on the leave-one-out cross validation, the RMSEs in component biomass estimation were similar for the adjustments from a correction factor based on OLS-RTO and OLS-WI approach except for Douglas-fir foliage and bark biomass where the OLS-WI performed better than the OLS-RTO. The inverse approach performed somewhat poorly compared to these approaches.

The OLS-RTO correction factor can be calculated using only one tree with desired characteristics such as a dominant tree in the stand or one with maximum DBH and therefore might be considered better than the OLS-WI and inverse approach. The choice should be governed by which component model one desires to calibrate. We calculated for such correction factor, assuming that our Douglas-fir data came from a single stand and represent the population of interest, and found it useful to use the tree with maximum DBH to calibrate regional CVTS and bark biomass equations, and the dominant tree to calibrate bole, foliage, and branch biomass (Table 7). The reason behind this could be that there is more volume and bark biomass in the lower sections of the trees

which are more affected by DBH than height, while on the other hand, biomass in branch and foliage is more dependent on tree height. For our lodgepole pine, however, the tree with the largest DBH was also the dominant tree, and the performance of OLS-RTO calibration was poor when the tree with the second largest DBH was used. Selecting trees of such characteristics to calibrate regional models was more useful than randomly selecting one tree. The decreasing trend in RMSE by using randomly selected trees slowed down significantly after five trees.

Accurate estimation of volume and biomass is required to quantify the role of forest management in carbon sequestration and climate change mitigation. However, the differences in scale of model development and application can result in biased estimates of volume and biomass. On the other hand, developing site- and species-specific volume and biomass models is not feasible, given the high cost and destructive nature of data collection. In this scenario, the use of calibration equations developed from limited data can serve as a balance between cost and accuracy. Sometimes the calibrated models may perform even better than locally fitted models using limited sample size (e.g., de-Miguel *et al.* 2014).

References

Bluhm, A. A., S. M. Garber and D.E. Hibbs. 2007. Taper equation and volume tables for plantation-grown red alder. Gen. Tech. Rep. PNW-GTR-735, Portland, OR. USDA For. Serv. Pacific Northwest Research Station. 74 p

Brackett, M. 1973. Notes on TARIF tree-volume computation. DNR Rep. 24. Olympia, WA. State of Washington, Dept. Natur. Res. 26 p

de-Miguel, S., L. Mehtatalo and A. Durkaya. 2014. Developing generalized calibratable mixed-effects meta-models for large-scale biomass prediction. Can. J. For. Res. 44: 648–656.

Draper, N.R. and H. Smith. 1998. Applied Regression Analysis (3rd ed.). John Wiley & Sons, New York.

Eerikäinen, K., D. Mabvurira, L. Nshubemuki and J. Saramäki. 2002. A calibratable site index model for *Pinus kesiya* plantations in southeastern Africa. Can. J. For. Res. 32: 1916–1928.

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

Hann, D.W. 2006. ORGANON user's manual edition 8.2. Oregon State University, Dept. For. Res., Corvallis, OR.

Hansen, M. 2002. Volume and biomass estimation in FIA: National consistency vs. regional accuracy. *In:* McRoberts, R. E., G. A. Reams, P. C. Van Deusen and J. W. Moser (eds.). Proceedings of the Third Annual Forest Inventory and Analysis Symposium. US For. Serv. Gen. Tech. Rep. NC-230. pp. 109–120.

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

Houghton, R.A. 2005. Aboveground forest biomass and the global carbon balance. Glob. Change Biol. 11(6): 945–958.

Houghton, R. A., F. Hall and S. J. Goetz. 2009. Importance of biomass in the global carbon cycle. J. Geophys. Res. 114 G00E03.

Jayaraman, K. and W. T. Zakrzewski. 2001. Practical approaches to calibrating height-diameter relationships for natural sugar maple stands in Ontario. For. Ecol. Manage. 148: 169–177.

Jenkins, C. J., D. C. Chojnacky, L. S. Heath and R. A. Birdsey. 2003. National-scale biomass estimators for United States tree species. For. Sci. 49(1): 12–35.

Ketterings, Q. M., R. Coe, R., M. van Noordwijk, Y. Ambagau and C. A. Palm. 2001. Reducing uncertainty in the use of allometric biomass equations for predicting above-ground tree biomass in mixed secondary forests. For. Ecol. Manage. 146: 199–209.

Kutner, M.H, C. J. Nachtsheim and J. Neter. 2004. Applied linear regression models. 4th ed. McGraw-Hill Irwin, Boston.

Lappi, J. 1991. Calibration of height and volume equations with random parameters. For. Sci. 37(3): 781–801.

Melson, S.L., M. E. Harmon, J. S. Fried and J. B. Domingo. 2011. Estimates of live-tree carbon stores in the Pacific Northwest are sensitive to model selection. Carbon Balance Manage. 6(2): 1–16.

Omule, S. A. Y., V. E. Fletcherand and K. R. Polsson. 1987. Total and merchantable volume equations for small coastal Douglas-fir. B.C. Ministry For. FRDA Rep. No. 010.

Poudel, K., 2015. Strategies for Sampling and Estimation of Aboveground Tree Biomass. PhD dissertation. Oregon State University, OR. 144 p

Poudel, K. P., H. Temesgen and A. N. Gray. 2015. Evaluation of sampling strategies to estimate crown biomass. For. Ecosys. 2: 1

Räty, M. and A. Kangas. 2008. Localizing general models with classification and regression trees. Scand. J. Forest Res. 23(5): 419–430.

Sadiq, R. A and V. G. Smith. 1983. Estimation of individual tree volumes with age and diameter. Can. J. For. Res. 13: 32–39.

Standish, J. T., G. H. Manning and J. P. Demaerschalk. 1985. Development of biomass equations for British Columbia tree species. Can. For. Serv., Info. Rep. BC-X-264, Pacific Forest Resource Centre, Victoria, BC. 47 p

Stone, M. 1974. Cross-validatory choice and assessment of statistical predictions. J. R. Stat. Soc. Series B. 36: 111–147.

Taylor, D. C.A., V. Pawar, D. T. Kruzikas, K. E. Gilmore, M. Sanon and M. C. Weinstein. 2012. Incorporating calibrated model parameters into sensitivity analyses deterministic and probabilistic approaches. Pharmacoeconomics. 30 (2): 119–126.

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

Temesgen, H., V. J. Monleon, A. R. Weiskittel and D. S. Wilson. 2011. Sampling strategies for efficient estimation of tree foliage biomass. For. Sci. 57: 153–163.

Temesgen, H., D. Affleck, K. P. Poudel, A. Gray and J. Sessions. 2015. A review of the challenges and opportunities in estimating above ground forest biomass using tree-level models. Scand. J. For. Res. 30(4): 326–335.

U.S. Environmental Protection Agency. 2015. Inventory of U.S. greenhouse gas emissions and sinks: 1990–2013. EPA 430-R-15-004., Washington, DC

Woodall, C., L. S. Heath, G. M. Domke and M. C. Nichols. 2011. Methods and equations for estimating above-round volume, biomass, and carbon for trees in the US forest inventory, 2010. USDA For. Serv. Nor. Res. Station Gen. Tech. Rep. NRS–88. Newtown Square, PA. 30 p

Zeng, W. S., H.R. Zhang and S. Z. Tang. 2011. Using the dummy variable model approach to construct compatible single-tree biomass equations at different scales: A case study for Masson pine (*Pinus massoniana*) in southern China. Can. J. For. Res. 41(7): 1547–1554.

Zhou, X. and M. A. Hemstrom. 2009. Estimating above-ground tree biomass on forest land in the Pacific Northwest: A comparison of approaches. USDA For. Ser. Res. Pap. PNW-RP-584. Portland, OR. 18 p

Zhou, X. and M. A. Hemstrom. 2010. Timber volume and aboveground live tree biomass estimations for landscape analyses in the Pacific Northwest. USDA For. Serv. Gen. Tech. Rep. PNW-GTR-819. Portland, OR. 31 p

Zianis, D., P. Muukkonen, R. Mäkipää and M. Mencuccini. 2005. Biomass and stem volume equations for tree species in Europe. Silva Fennica Monographs 4. 63 p