

RESEARCH

Open Access



Estimating upper stem diameters and volume of Douglas-fir and Western hemlock trees in the Pacific northwest

Krishna P. Poudel¹, Hailemariam Temesgen^{1*} and Andrew N. Gray²

Abstract

Background: Volume and taper equations are essential for obtaining estimates of total and merchantable stem volume. Taper functions provide advantages to merchantable volume equations because they estimate diameter inside or outside bark at specific heights on the stem, enabling the estimation of total and merchantable stem volume, volume of individual logs, and a height at a given diameter.

Methods: Using data collected from 1218 trees (1093 Douglas-fir (*Pseudotsuga menziesii* (Mirbel) Franco) and 125 western hemlock (*Tsuga heterophylla*)), we evaluated the performance of one simple polynomial function and four variable-exponent taper functions in predicting upper stem diameter. Sample trees were collected from different parts of the states of Oregon, Washington, and California. We compared inside-bark volume estimates obtained from the selected taper equation with estimates obtained from a simple logarithmic volume equation for the data obtained in this study and the equations used by the Forest Inventory and Analysis program in the Pacific Northwest (FIA-PNW) in the state of California and western half of the states of Oregon and Washington.

Results: Variable exponent taper equations were generally better than the simple polynomial taper equations. The FIA-PNW volume equations performed fairly well but volume equations with fewer parameters fitted in this study provided comparable results. The RMSE obtained from taper-based volume estimates were also comparable with the RMSE of the FIA-PNW volume equations for Douglas-fir and western hemlock trees respectively.

Conclusions: The taper equations fitted in this study provide added benefit to the users over the FIA-PNW volume equations by enabling the users to predict diameter at any height, height to a given diameter, and merchantable volume in addition to cubic volume including top and stump (CVTS) of Douglas-fir and western hemlock trees in the Pacific Northwest. The findings of this study also give more confidence to the users of FIA-PNW volume equations.

Keywords: Pacific northwest, Biomass equations, Utilization standards, Diameter inside bark

Background

Volume and taper equations are essential for obtaining estimates of total and merchantable stem volume. Volume equations relate total or merchantable stem volume with easily-measurable variables such as diameter at breast height (DBH; 1.37 m), total tree height, and other variables (e.g. height to crown base or crown ratio) through regression. Estimates of stem volume to certain diameter limits are critical to meet different timber utilization standards.

However, such utilization standards change in response to local market and economic conditions making the use of a fixed merchantable volume equation less attractive (Czaplewski et al. 1989). Taper functions provide advantages to merchantable volume equations because they estimate diameter inside or outside bark (dib or dob) at specific heights on the stem, enabling the estimation of total and merchantable stem volume, volume of individual logs (Kozak 1988), and a height at a given diameter (Li et al. 2012).

Numerous volume and taper equations have been published and are being used at different scales of

* Correspondence: Hailemariam.Temesgen@oregonstate.edu

¹Department of Forest Engineering, Resources, and Management, College of Forestry, Oregon State University, 280 Peavy Hall, Corvallis, OR 97331, USA
Full list of author information is available at the end of the article

forest management. There are also many more unpublished and proprietary equations developed and used by forest companies and agencies. The application of volume equations is not only crucial for economic valuation of timber resources but is also vital to the assessment of biomass availability and carbon sequestration (Poudel and Temesgen 2016). For example, the official U.S. forest carbon report to the United Nation Framework Convention on Climate Change (UNFCCC) is based on a component ratio method that converts the sound wood volume obtained from regional volume equations to stem biomass using wood density and bark and branch scaling factors.

Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco) and western hemlock (*Tsuga heterophylla* (Raf.) Sarg.) are two major tree species in the Pacific Northwest (PNW – States of Oregon, Washington, and California) and account for a substantial portion of the live volume and biomass in the region. A variety of approaches to obtain total and merchantable volume in the PNW are in common use. The Forest Inventory and Analysis (FIA) program of the U.S. Forest Service in the PNW (FIA-PNW) estimates Douglas-fir cubic volume including top and stump (CVTS) using the Brackett (1977) equation in western Oregon and western Washington, the Summerfield (1980) equation in eastern Oregon and eastern Washington, and the MacLean and Berger (1976) equation in California. However, for western hemlock, it uses the Chambers and Foltz (1979) volume equation for all three states (OR, WA, and CA). The Oregon Department of Forestry uses taper functions associated with the Forest Projection and Planning System (Arney et al. 2004). The Washington Department of Natural Resources uses taper functions developed by Flewelling and Ernst (1996), Flewelling (1994), Kozak (1994), or the Brackett (1977) volume equation depending on species and location (east-side or west-side) to estimate volume.

Taper equations have been used in forestry for a long time and can be divided into two major groups. The first group of equations expresses tree form as a single continuous function (Newnham 1988, 1992; Kozak 1988, 2004). The second group of equations (segmented taper equations) uses different models for various parts of the stem and joins these models in such a way that their first derivatives are equal at the point of intersection (Max and Burkhart 1976; Clark et al. 1991).

Differences in stand conditions affect tree form and thus tree volume (Bluhm et al. 2007). Accordingly, different model forms and fitting techniques have been used in developing volume and taper models in the past in different kinds of stands. These models range from simple polynomial to nonlinear and multivariate regression models (Kublin et al. 2008). Traditionally, attempts to improve the predictive ability of taper equations were made by the addition of auxiliary variables such as crown

dimensions, stand and site variables, and upper stem diameter measurements. Recent studies, however, have focused on approaches to account for the observed between-tree variability in stem form (Trincado and Burkhart 2006), included stand density as explanatory variable (Sharma and Parton 2009), and calibration of taper equations using upper stem diameter measurements (e.g. Cao 2009; Aria-Rodil et al. 2014).

Selecting the best taper equation to predict upper stem diameters and consequently total tree volume or volume to a specific diameter or height is crucial for forest managers. Thus, the evaluation of different taper equations is critical. The objectives of this study were to: 1) fit taper equations for Douglas-fir and western hemlock trees; 2) examine the accuracy of these equations in predicting diameter and volume inside bark; 3) develop a simple volume equation based on DBH and tree height; and 4) compare the accuracy of taper-based volume estimates with the volume estimates obtained from the FIA-PNW equations and the simple volume equation fitted in this study.

Methods

Data

Data for this study came from three different sources and consisted of 1218 trees – 1093 Douglas-fir and 125 western hemlock. The first set of data (DATASET I) consisted of measurements on 716 trees (615 Douglas-fir and 101 western hemlock) sampled in 1993 from the western side of the states of Oregon and Washington. Average DBH of these trees was 37.1 cm (range 8.8–92.5 cm) and 36.5 cm (range 16.3–102.1 cm) for Douglas-fir and western hemlock trees, respectively. Average height of these trees was 30.8 m (range 10.2–61.9 m) and 27.5 m (range 12.7–40.7 m) for Douglas-fir and western hemlock trees, respectively. Diameter inside bark in these trees were measured at 0.3, 0.6, 0.9 and 1.37 m, and at each 1/10th of the height above breast height afterward.

The second set of data (DATASET II) consisted of measurements on 399 Douglas-fir trees collected by the Stand Management Cooperative (SMC) from the western half of the states of Oregon and Washington. Average DBH and heights of these trees were 18.2 cm (range 4.7–43.8 cm) and 15.2 m (range 5.1–27.6 m), respectively. Diameter inside bark in these trees were measured at 0.1, 1.0, 1.37 m, and at every 1 m afterward.

The third set of data (DATASET III) consisted of measurements on 103 trees (79 Douglas-fir and 24 western hemlock) sampled in 2012–2015 from the states of Oregon, Washington, and California as a part of an FIA biomass sampling and estimation project at Oregon State University. Average DBH of these trees was 48.6 cm (range 16.5–114.0 cm) and 41.7 cm (range 18.0–69.9 cm) for Douglas-fir and western hemlock

trees, respectively. Average height of these trees was 32.2 m (range 16.5–53.3 m) and 29.5 m (range 14.3–43.7 m) for Douglas-fir and western hemlock trees, respectively. Diameter inside bark in these trees was measured at stump height (approximately 0.3 m) and at every 5.18 m afterward. Numbers of trees by species and diameter class in each dataset are presented in Table 1.

Actual volume computation

Sampling protocols for three datasets differed so greatly that it was necessary to harmonize the method for actual volume computation for all datasets. We considered two linear interpolation approaches to accomplish this. In the first approach, the dibs at specified intervals were

Table 1 Number of trees by diameter class in three datasets used in this study. A total of 1218 trees (1093 Douglas-fir and 125 western hemlock) were used

DBH Class (cm)	DATASET I		DATASET II	DATASET III	
	Douglas-fir	Western hemlock	Douglas-fir	Douglas-fir	Western hemlock
4–8			14		
8–12	5		58		
12–16	6		104		
16–20	42	9	89	2	1
20–24	62	5	62	4	4
24–28	68	14	35	3	1
28–32	72	13	22	2	1
32–36	75	12	8	7	3
36–40	62	15	4	9	3
40–44	54	10	3	7	2
44–48	47	8		13	1
48–52	30	4		6	1
52–56	25	4		10	2
56–60	19	1		4	1
60–64	13	2		1	
64–68	7	1		3	2
68–72	9	1		1	2
72–76	5	1			
76–80	6				
80–84	5			1	
84–88				1	
88–92	2				
92–96	1			1	
96–100				1	
100–104		1			
108–112				1	
112–116				2	
Total	615	101	399	79	24

obtained based on linear interpolation of measured dibs. In the second approach the interpolated dibs were obtained based on linear interpolation of fractional error in predicted dib based on the selected taper equation. The second approach was selected based on graphical comparison and its plausibility, particularly at the butt log. The following steps were used to get the actual volume:

- 1) Fit Kozak (2004) taper equation using measured DIBs and obtain predicted diameter inside bark (PDIB) at each measurement heights;
- 2) Obtain fractional error (FE) in predicted DIB as:

$$FE = \frac{DIB^2 - PDIB^2}{PDIB^2} \tag{1}$$

- 3) Set a consistent stump height (minimum of all stump heights – 0.03 m);
- 4) Divide tree height into 100 equal parts and obtain FE at each percentile of tree height by linear interpolation

$$FE_{int} = \frac{FE_0(h_1 - h) + FE_1(h - h_0)}{h_1 - h_0} \tag{2}$$

where, FE_{int} is the linear interpolated fractional error at height h ; h_0 and h_1 are measurement heights immediately below and above h ; and FE_0 and FE_1 are fractional errors immediately below and above h ;

- 5) Obtain interpolated DIBs at each h by back solving Eq. (1) for DIB;

$$DIB_{FE} = \sqrt{PDIB^2(FE_{int} + 1)} \tag{3}$$

DIB_{FE} is same as the measured dib at heights where actual measurements were made.

- 6) Compute volume below 0.03 m and the top section as cylinder and cone respectively;
- 7) Compute volume of other sections using Smalian’s formula with DIB_{FE} i.e. using numerical integration with step size equal to 1% of total tree height and diameters at two ends obtained from interpolation using Kozak (2004) taper equation.

Observed inside bark cubic volume including top and stump (CVTS) for each tree was then computed by summing volume of all sections, stump, and top.

Methods

Diameter inside bark

Rojo et al. (2005) evaluated 31 different taper functions that belonged to three model groups – simple, segmented, and variable-form taper functions. These models differ in how they describe stem profile as well as their model forms and number of parameters to be estimated. In this study, we used one simple and four variable exponent taper models that Rojo et al. (2005) selected for final evaluation. These models differed in how they describe tree taper along the bole. Except for the Cervera (1973) taper equation, which is a simple polynomial equation, all the models are variable-exponent taper models that describe bole shape with a changing exponent. Geometric properties of these taper functions are displayed by plotting predicted diameter inside bark vs. height on tree bole of a Douglas-fir tree (Fig. 1). At first, the leave-one-out validation was carried out to determine the best model for estimating upper stem diameter. Parameters of the five taper models (Table 2) were obtained based on two datasets (e.g. dataset I and II). To obtain evaluation statistics, fitted models were then applied to the third dataset (e.g. dataset III).

At this point, all models were fitted as fixed effects model and assumed homogenous error variance. Accuracy of these models were compared using mean prediction bias and root mean squared error (RMSE) produced by these models in estimating diameter inside bark along the bole. Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC) values were also obtained for each of these models.

$$\text{Bias} = \frac{1}{n} \sum_{i=1}^n \left(\frac{\sum_{j=1}^{m_i} (y_{ij} - \hat{y}_{ij})}{m_i} \right) \tag{4}$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{\sum_{j=1}^{m_i} (y_{ij} - \hat{y}_{ij})^2}{m_i} \right)} \tag{5}$$

where, y_{ij} and \hat{y}_{ij} are observed and predicted diameter inside bark of j^{th} observation ($j = 1, 2, \dots, m_i$) along the bole on i^{th} tree ($i = 1, 2, \dots, n$).

The Kozak (2004) taper model was selected for further analysis because it produced the smallest root mean squared error among the five models used to predict upper stem diameter of Douglas-fir trees in all datasets and comparable RMSEs for western hemlock trees. Because the taper data consists of multiple measurements taken on a single tree, there is an inherent autocorrelation among the observations obtained from the same tree which can be accounted for by adding individual tree random effects and by specifying a correlation structure in the model (Li et al. 2012).

An error model of the following form was fitted:

$$e^2 = \exp(a_0 + a_1 \ln(\text{DBH}) + a_2 \ln(h) + a_3 \ln(H)) \tag{6}$$

where e is the residual (observed dib – predicted dib) obtained from fitting the Kozak (2004) model as a non-linear fixed effect model; $a_0, a_1, a_2,$ and a_3 are regression

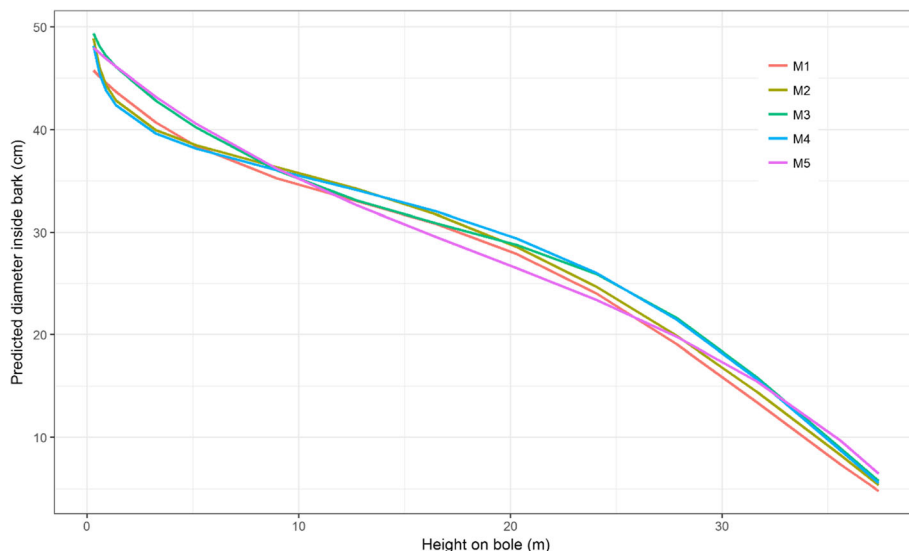


Fig. 1 Geometric properties of five taper functions displayed by plotting predicted diameter inside bark along the bole for a Douglas-fir tree with 46.2 cm DBH and 40.7 m height

Table 2 Taper equations evaluated for predicting upper stem diameter in this study

Model	Equation
M1: Cervera 1973	$\frac{d}{D} = (b_1 + b_2X + b_3X^2 + b_4X^3 + b_5X^4)$
M2: Kozak 1988	$d = b_1 D^{b_2} b_3^D \left(\frac{1-\sqrt{t}}{1-\sqrt{p}} \right)^{(b_4 t^2 + b_5 \ln(t+0.001) + b_6 \sqrt{t} + b_7 \exp(t) + b_8 (\frac{t}{p}))}$
M3: Bi 2000	$\frac{d}{D} = \left(\frac{\ln(\sin(\frac{p}{H}))}{\ln(\sin(\frac{h}{H}))} \right)^{b_1 + b_2 \sin(\frac{p}{H}) + b_3 \cos(\frac{p}{H})} + b_4 \left(\frac{\sin(\frac{p}{H})}{t} \right) + b_5 D + b_6 t \sqrt{D} + b_7 t \sqrt{H}$
M4: Kozak 2004	$d = b_1 D^{b_2} H^{b_3} \left[\frac{1-t}{1-k^3} \right]^{b_4 t^4 + b_5 \left(\frac{1-t}{\exp(p)} \right) + b_6 \left(\frac{1-t}{1-k^3} \right)^{0.1} + b_7 (\frac{t}{H}) + b_8 (H^1 - (\frac{t}{H})^{\frac{1}{k}}) + b_9 \left(\frac{1-t}{1-k^3} \right)}$
M5: Arias-Rodil et al. 2014	$d = 2 \left\{ \frac{b_1 D}{1 - \exp(b_2(1.3-H))} + \left(\frac{D}{2} - b_1 D \right) \times \left[1 - \left(\frac{1}{1 - \exp(b_2(1.3-H))} \right) \right] + \exp(-b_2 h) \times \left[\frac{(\frac{D}{2} - b_1 D) \exp(1.3b_2)}{1 - \exp(b_2(1.3-H))} \right] - \exp(b_3 h) \left[\frac{b_1 D \exp(-b_2 H)}{1 - \exp(b_2(1.3-H))} \right] \right\}$

where, D is the diameter at breast height outside bark (cm); h is the height above ground level (m); d is the diameter inside bark (cm) at height h ; H is the total tree height (m); $b_1 - b_9$ are regression coefficients to be estimated from data; $X = \frac{H-h}{H-1.3}$; $t = h/H$, $k = 1.3/H$; p is the inflection point and was set to 0.25 and 0.12 for Douglas-fir and western hemlock, respectively

parameters; and all other variables are the same as defined previously. The reason for fitting an error model was to approximate the best weighting factor to use in fitting weighted taper equation. A power variance function defined as $\text{Var}(\varepsilon_i) = \sigma^2 |v_i|^{2t}$ was specified in the model to account for the heterogeneity of error variance. Here, ε_i is the model residual, σ^2 is the residual sum of squares, $v_i = \frac{1}{\hat{e}^2}$ is the weighting variable where \hat{e}^2 is the predicted values of e^2 obtained from fitting Eq. 6, and t is the variance function coefficient.

Volume estimation

Taper-based volume At first, we obtained the parameter estimates of the weighted nonlinear mixed effects model with first-order autoregressive error structure. After investigating different combination of random effects, parameters b_4 and b_7 of the selected taper equation (Table 2, M4) were associated with tree specific random effects. Diameter inside bark along the bole for each tree in the held-out dataset was obtained using the fixed effects parameters of the fitted taper equation. Taper-based CVTS was then obtained using numerical integration with step size equal to 1% of total tree height. Evaluation statistics (i.e. bias and RMSE) for predicting diameter inside bark CVTS were also calculated.

Local volume equation Even though taper equations have advantages over direct volume equations in that they have flexibility in estimating volume to different merchantability standards, it is desired that the taper equations provide CVTS with the same accuracy as the direct volume equations. We fitted a simple nonlinear volume equation that predicts the CVTS as a function of DBH and total tree height. Evaluation statistics for the simple volume equation (Eq. 7) were obtained for this model as well based on leave-one-out approach.

$$\text{CVTS} = \exp(a_0 + a_1 \ln(\text{DBH}) + a_2 \ln(H)) \quad (7)$$

where, $\ln(\cdot)$ is the natural logarithm, a_0 , a_1 and a_2 are the parameters to be estimated from data and other variables are as defined previously.

FIA-PNW volume equations The Pacific Northwest unit of the FIA uses different sets of species-specific volume equations depending upon geographic location (See Background section for details). The equation forms along with their coefficients for Douglas-fir and western hemlock trees are given in Table 3.

All statistical analysis was performed in R3.3.2 (R Core Team 2016) and the mixed effects models were fitted with the nlme function in library nlme (Pinheiro et al. 2016). Evaluation statistics for estimating CVTS were obtained as follows:

$$\text{Bias} = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)}{n} \quad (8)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (9)$$

where, y_i and \hat{y}_i are observed and predicted CVTS of the i^{th} ($i = 1, 2, 3, \dots, n$) tree. Bias and RMSE percentages were obtained by dividing these values by the average observed CVTS of n trees.

Results and discussions

Diameter inside bark

The parameter estimates and their corresponding standard errors for the five taper equations are given in Table 4. Plots of standardized residuals vs. fitted values obtained from the final model fitted for both species (Fig. 2) did not show any problems with heteroscedasticity. For Douglas-fir, all the parameters in all taper models were statistically significant at the 0.05 level of significance but for western hemlock some of the parameters namely b_2 and b_3 (p -values 0.8 and 0.9, respectively) for Bi (2000) model and coefficient b_3 (p -value = 0.3) for Kozak (2004) models were not statistically significant. Mean prediction bias and

Table 3 Volume equations used by FIA-PNW to estimate cubic volume including top and stump (CVTS)

Species	State	Volume equation
Douglas-fir	WOR & WWA (Brackett 1977)	$CVTSL = -3.21809 + 0.04948 \log(H) \log(D) - 0.15664 (\log(D))^2$ $+ 2.02132 \log(D) + 1.63408 \log(H) - 0.1685 (\log(H))^2$ $CVTS = 10^{CVTSL}$
	EOR & EWA (Summerfield 1980)	$CVTSL = -6.110493 + 1.81306 \ln(D) + 1.083884 \ln(H)$ $CVTS = \exp(CVTSL)$
	CA ^a (MacLean and Berger 1976)	$BA = D^2 \cdot 0.005454154$ $CF4 = -0.248569 + 0.0253524 \frac{H}{D} - 0.0000560175 \left(\frac{H^2}{D}\right)$ <p>If CF4 < 0.3, then CF4 = 0.3; If CF4 > 0.4 then CF4 = 0.4</p> $CV4 = CF4 \cdot BA \cdot H$ $TERM = \left((1.033 (1.0 + 1.382937 \exp(-4.015292 (\frac{DBH}{10.0})))) (BA + 0.087266) - 0.174533 \right)$ $CVTS = \frac{CV4 \cdot TERM}{BA - 0.087266}$
Western hemlock	OR, WA, & CA (Chambers and Foltz 1979)	$CVTS = 10^{-2.72170 + 2.00857 \log(D) + 1.08620 \log(H) - 0.00568 D}$

WOR Western Oregon, WWA Western Washington, EOR Eastern Oregon, EWA Eastern Washington, CA California, CVTSL log (base 10) of the cubic volume including top and stump; log(.) log base 10, ln(.) natural logarithm, H total tree height, and D diameter at breast height, BA Basal Area

^a CVTS for California Douglas-fir trees in FIA-PNW is calculated differently for trees less than 15 cm in DBH however, no Douglas-fir tree in this study was less than 15 cm DBH

root mean squared error produced by these models in predicting upper stem inside bark diameters obtained based on leave-one-out validation are given in Table 5.

For Douglas-fir trees, the Cervera (1973), Kozak (1988), and Kozak (2004) models produced positive mean prediction biases for dataset I; i.e., these model under-predicted the diameter inside bark along the bole for dataset I. These models for dataset II and III and all other models for Douglas-fir, had negative biases; i.e., these models over-predicted the diameter inside bark along the bole. The Kozak (2004) model had the smallest values for RMSE but in terms mean prediction bias (absolute value), the Arias-Rodil et al. (2014) model

was best for dataset I and II and the Kozak (1988) model was best for dataset III (Table 5). Absolute values of mean prediction bias in predicting dib ranged from 0.09 cm to 1.76 cm and RMSE ranged from 1.04 cm to 4.70 cm. Note that the mean prediction bias and root mean squared errors are based on leave-one(dataset)-out validation.

In the case of western hemlock trees, all taper models produced negative biases for dataset I and positive biases for dataset III in predicting diameter inside bark. Absolute values of biases ranged from 0.37 to 1.83 cm and RMSE ranged from 3.32 to 4.2 cm. The Arias-Rodil et al. (2014) model had the smallest bias (absolute value)

Table 4 Parameter estimates their standard errors (in parenthesis) for the taper equations used in this study

Species	Model	b ₁	b ₂	b ₃	b ₄	b ₅	b ₆	b ₇	b ₈	b ₉
Douglas-fir	M1	0.02435 (0.00349)	0.85814 (0.04419)	2.57379 (0.16020)	-5.55862 (0.21594)	3.05029 (0.09678)	-	-	-	-
	M2	0.79314 (0.01036)	1.02189 (0.00475)	0.99754 (0.00011)	1.10118 (0.0293)	-0.18022 (0.00679)	1.19183 (0.07506)	-0.61055 (0.03841)	0.15539 (0.00388)	-
	M3	0.14030 (0.04284)	0.72960 (0.02170)	0.15440 (0.00481)	-0.12140 (0.02365)	0.00124 (0.00008)	0.02678 (0.00244)	-0.09579 (0.00215)	-	-
	M4	1.04208 (0.00750)	0.99771 (0.00344)	-0.03111 (0.00423)	0.53788 (0.00776)	-1.01291 (0.01796)	0.56813 (0.00622)	4.96019 (0.12262)	0.04124 (0.00125)	-0.34417 (0.01478)
	M5	0.30853 (0.00476)	0.08620 (0.00295)	0.07768 (0.00184)	-	-	-	-	-	-
Western hemlock	M1	0.03080 (0.01520)	0.42992 (0.19493)	4.12143 (0.71334)	-7.85114 (0.96822)	4.27438 (0.43611)	-	-	-	-
	M2	0.92336 (0.06729)	1.00547 (0.02555)	0.99796 (0.00053)	1.05050 (0.18665)	-0.16240 (0.04297)	0.24448 (0.42759)	-0.31429 (0.22831)	0.30181 (0.01235)	-
	M3	1.84835 (0.14518)	0.01014* (0.0756)	0.00175* (0.01611)	-1.04746 (0.08274)	0.0029 (0.00033)	0.03483 (0.00922)	-0.13071 (0.00935)	-	-
	M4	1.05981 (0.04743)	0.99433 (0.0101)	-0.01684* (0.0174)	0.64632 (0.03945)	-1.56599 (0.08468)	0.74293 (0.0238)	4.75618 (0.89389)	0.0389 (0.00764)	-0.19425 (0.0826)
	M5	0.04561 (0.00445)	-0.02362 (0.00204)	-0.86616 (0.06912)	-	-	-	-	-	-

*Coefficient not significant at α = 0.05

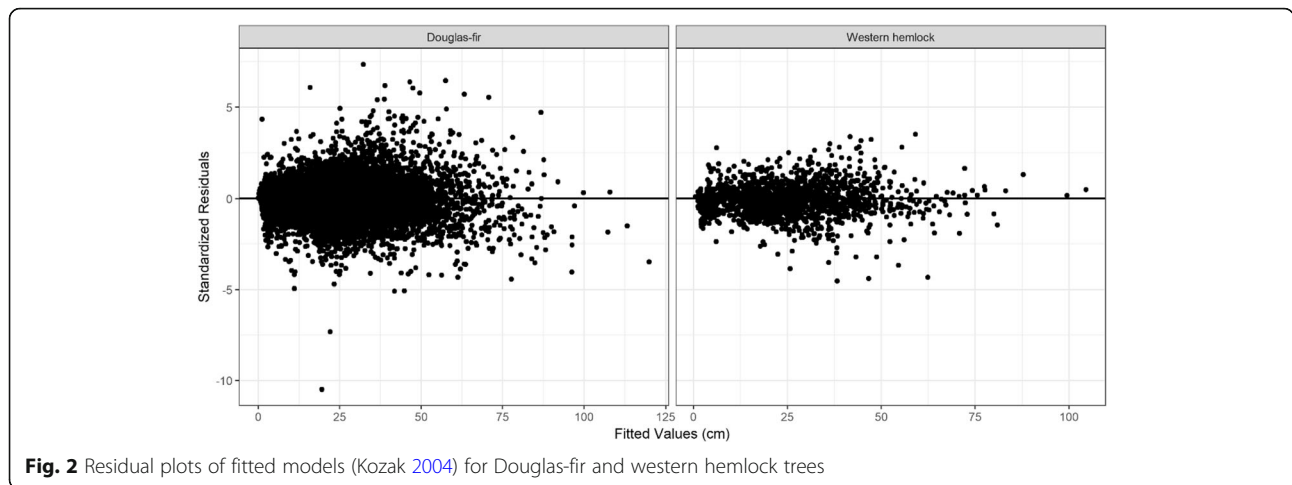


Fig. 2 Residual plots of fitted models (Kozak 2004) for Douglas-fir and western hemlock trees

and RMSE values for dataset III but the Kozak (1988) had the smallest RMSE for dataset I. Once again, the mean prediction bias and root mean squared errors are based on leave-one(dataset)-out validation.

The relative height at inflection points, based on Kozak 1988 model, were set to 0.25 and 0.12 for Douglas-fir and western hemlock, respectively. In general, the Kozak models (1988 and 2004) performed better than other models in predicting diameter inside bark along the bole for both species. Additionally, the Kozak 2004 model had lower RMSE for Douglas-fir and very similar prediction

bias (absolute value) for western hemlock trees compared to the 1988 model. This model also has much lower multicollinearity if the model were to be log transformed to fit as a linear model (Kozak 2004) and has been successfully used by many researchers (e.g. Rojo et al. 2005; Li et al. 2012).

The selected Kozak (2004) taper model was further improved by fitting it as a weighted nonlinear mixed effects model with first-order autoregressive correlation structure (CAR(1)) and provided the smaller BIC compared to the same model with second-order autoregressive (CAR(2)) structure. Thus, our taper-based volume calculations were based on the diameter inside bark predicted from this model. Li et al. (2012) also found the CAR(1) model-fitting approach best for eleven conifer species in the Acadian region of North America. Rojo et al. (2005), however, found the CAR(2) model to perform better for maritime pine in Galicia, Northwestern Spain.

Parameter estimates and their standard errors obtained by fitting the Kozak 2004 model as a weighted nonlinear mixed effects model with first-order autoregressive error structure are given in Table 6. All parameters were statistically significant at the 0.05 level of significance for western hemlock but for Douglas-fir, parameter b_3 was not significant (p -value = 0.4). Note that these parameter estimates are obtained by fitting the model using the entire dataset. We did not see any pattern in bias in predicting upper stem diameter by the Kozak (2004) taper function but the RMSE slightly increased with increasing DBH for both species (Fig. 3). However, there was no obvious pattern in relative RMSE calculated as $100 \times \frac{RMSE}{DBH}$.

Table 5 Mean prediction bias, root mean squared error (RMSE) in estimating diameter inside bark using different models. Statistics were obtained by using leave-one-out cross validation method e.g. bias and RMSE for dataset I were obtained by applying the models fitted using datasets II and III

Model	Dataset	Douglas-fir		Western hemlock	
		Bias (cm)	RMSE (cm)	Bias (cm)	RMSE (cm)
M1	I	0.85	2.67	-1.07	4.02
	II	-0.34	1.30		
	III	-1.10	3.84	1.83	3.64
M2	I	0.61	2.46	-1.06	3.32
	II	-0.36	1.28		
	III	-0.17	3.54	1.35	4.25
M3	I	-0.59	2.69	-0.94	3.34
	II	-0.62	1.30		
	III	-0.55	4.05	1.38	4.25
M4	I	0.35	2.21	-1.09	3.44
	II	-0.28	1.04		
	III	-0.35	3.35	1.49	4.27
M5	I	-0.09	3.04	-0.37	3.83
	II	-0.16	1.26		
	III	-1.76	4.70	1.16	3.61

Volume estimation

Leave-one-out validation statistics in predicting CVTS based on all the taper equations fitted as fixed effects models were also obtained (Table 7). There was no

Table 6 Parameter estimates and their standard error for Kozak (2004) model fitted for Douglas-fir and western hemlock trees. Parameters b_4 and b_7 were associated with tree specific random effects

Parameter	Douglas-fir		Western Hemlock	
	Estimate	SE	Estimate	SE
b_1	1.00403	0.01382	0.80993	0.06838
b_2	0.98460	0.00667	0.97170	0.02351
b_3	-0.00670*	0.00840	0.08714	0.03534
b_4	0.49856	0.00678	0.52283	0.02270
b_5	-0.85726	0.03014	-1.09100	0.10853
b_6	0.54275	0.00916	0.72762	0.02775
b_7	4.47089	0.17168	2.17953	0.92582
b_8	0.04682	0.00116	0.03593	0.00637
b_9	-0.41454	0.01195	-0.32003	0.05524
φ	0.77593		0.74885	
δ	-0.38306		-0.71222	

*Coefficient statistically not significant at $\alpha = 0.05$ (p -value = 0.4); φ = Estimate of first-order continuous autoregressive parameter; δ = Power of the variance weighting factor

model that was consistently better than the other in predicting CVTS for both species and all datasets. For Douglas-fir trees, Bi (2000), Arias-Rodil et al. (2014), and Kozak (1988) taper models produced the smallest absolute bias for datasets I, II, and III, respectively. However, the Kozak (2004) produced smallest RMSE for datasets I and II and Kozak (1988) produced the smallest RMSE for dataset III. For western hemlock trees, Cervera (1973) and Kozak (1988) produced the smallest bias and RMSE, respectively, for dataset I and

Arias-Rodil et al. (2014) model produced smallest bias and RMSE for dataset III.

Table 8 shows the parameter estimates and their standard error of a simple CVTS model (Eq. 7) fitted to the data obtained in this study. All the coefficient in this model were statistically significant at 5% level of significance (p -value < 0.05). Evaluation statistics of CVTS (bias and RMSE) obtained from the fitted taper equations, local volume equation, and the volume equations used by FIA-PNW are given in Table 9. All three methods (local volume equation, FIA-PNW equation, and taper-based volume prediction) over-estimated the logarithmic volume in Douglas-fir trees as indicated by negative bias (Table 9). The local volume equation for Douglas-fir performed marginally less biased than the volume equation used by the FIA-PNW (-0.0103 m³ vs. -0.0185 m³) and Kozak (2004) taper equation (-0.0103 m³ vs. -0.0656 m³). However, the FIA-PNW equation had smaller RMSE than the local and taper-based volume equation. We saw some increase in absolute bias with CVTS prediction using Kozak (2004) equation (M4) for trees larger than 65 cm DBH (Fig. 4). This could be due to the smaller number of larger trees available in the model fitting dataset.

Similar to the Douglas-fir trees, M4 and FIA-PNW equations over-predicted volume for western hemlock trees. The local volume equation, however, underpredicted western hemlock CVTS. M4 produced the smallest bias while the error was highest with the FIA-PNW equation (Table 9). Once again, the FIA-PNW equations produced the smallest root mean squared error.

It is important to note that the simple CVTS equation had smaller bias and comparable RMSE compared with the FIA-PNW equation even though the simple CVTS

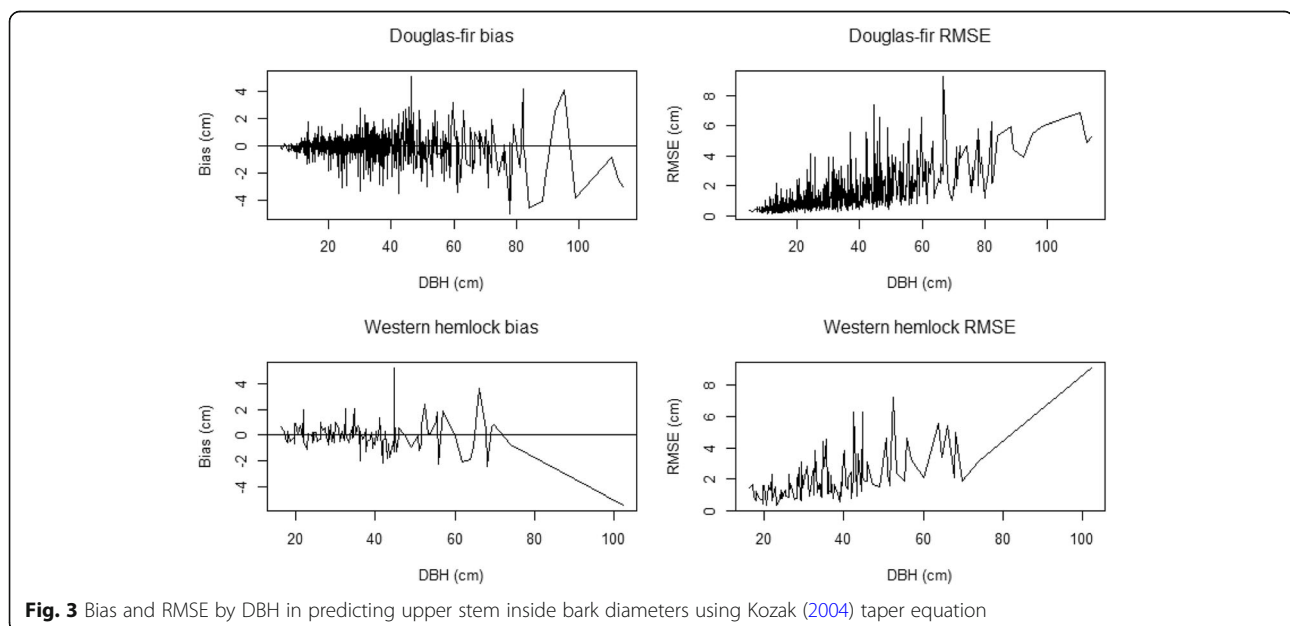


Fig. 3 Bias and RMSE by DBH in predicting upper stem inside bark diameters using Kozak (2004) taper equation

Table 7 Mean prediction bias, root mean squared error (RMSE) in estimating inside bark cubic volume using different models. Statistics were obtained by using leave-one-out cross validation method i.e. bias and RMSE for dataset I were obtained by applying the models fitted using datasets II and III

Model	Dataset	Douglas-fir				Western hemlock			
		Bias (cm)	Bias percent	RMSE (cm)	RMSE percent	Bias (cm)	Bias percent	RMSE (cm)	RMSE percent
M1	I	0.22	12.56	0.35	20.36	-0.06	-4.29	0.58	42.51
	II	0.02	11.03	0.04	20.51				
	III	-0.16	-6.09	0.85	31.55	0.34	15.98	0.45	21.19
M2	I	0.13	7.21	0.31	17.77	-0.16	-11.63	0.35	25.99
	II	-0.01	-5.01	0.03	16.27				
	III	-0.06	-2.07	0.36	13.52	0.23	10.80	0.42	19.67
M3	I	-0.04	-2.31	0.29	16.79	-0.17	-12.26	0.43	31.66
	II	-0.02	-9.20	0.04	18.82				
	III	-0.15	-5.51	0.43	15.96	0.23	10.70	0.42	19.75
M4	I	0.05	2.75	0.24	13.95	-0.18	-13.16	0.46	33.65
	II	-0.01	-4.28	0.03	14.64				
	III	-0.13	-4.89	0.47	17.52	0.22	10.40	0.40	18.83
M5	I	0.07	4.24	0.29	16.45	-0.13	-9.41	0.61	45.14
	II	0.00	-2.07	0.03	15.94				
	III	-0.44	-16.13	1.18	43.91	0.16	7.41	0.28	13.33

equation has only three coefficients compared to six coefficients in FIA-PNW Douglas-fir equation for western Oregon and Washington. FIA-PNW volume equation for western hemlock uses DBH, log DBH, and log height to predict logarithm of CVTS. Our model fitted in this study has one less parameter and had smaller bias (absolute value) compared to the equation used by the FIA-PNW (Table 9).

Bias and RMSE in estimating CVTS based on taper equation also differed by tree diameter for both species (Fig. 4). For Douglas-fir, bias in taper-based CVTS ranged from -12.66 to 0.84%. For western hemlock, bias ranged from -28.64 to 2.29% and performance of taper equation was poor for trees larger than 65 cm DBH, as noted previously. This could be because there were only a few larger trees available for model development. Predicted stem profiles i.e. plot of relative diameter (d/D)

vs. relative height (h/H), of small, medium, and large trees, are shown in Figs. 5 and 6 and show that there is less taper in smaller trees compared with the medium and large sized trees for both species.

Summary and conclusion

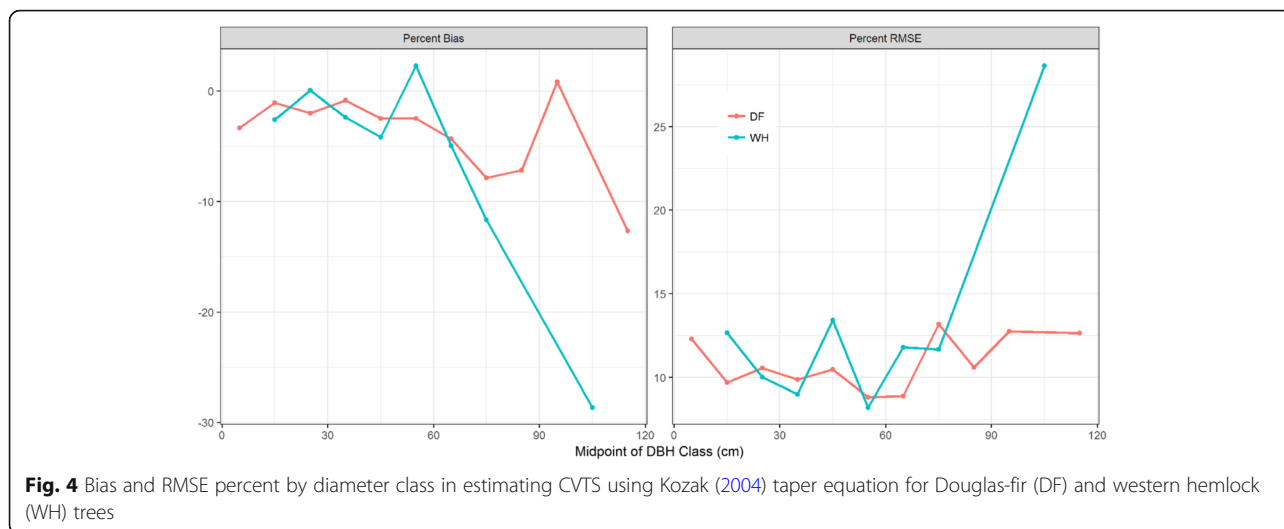
We evaluated the performance of five different taper equations in estimating upper stem diameters and cubic volume including top and stump (CVTS) of Douglas-fir and western hemlock trees based on mean prediction bias and RMSE they produced. Both Kozak (1988) and

Table 8 Parameter estimates and their standard error for a CVTS equation, in the form of $CVTS = \exp(a_0 + a_1 \ln(D) + a_2 \ln(H))$, fitted for Douglas-fir and western hemlock trees. All regression parameters were statistically significant at 5% level of significance

Parameter	Douglas-fir		Western Hemlock	
	Estimate	SE	Estimate	SE
a_0	-9.70405	0.04822	-9.98200	0.23210
a_1	1.61812	0.01195	1.37228	0.03673
a_2	1.21071	0.01771	1.57319	0.08084

Table 9 Mean bias and RMSE for estimating cubic volume including top and stump obtained from fitted taper equations (M4), volume equations fitted in this study (Local), and the volume equations used by FIA-PNW. Evaluation statistics for each dataset were obtained by using leave-one-out cross validation method i.e. bias and RMSE for dataset I were obtained by applying the models fitted using datasets II and III

Species	Model	Bias (m ³)	RMSE (m ³)
Douglas-fir	M4	-0.0656	0.2721
	Local	-0.0103	0.2525
	FIA-PNW	-0.0185	0.2235
Western hemlock	M4	-0.0137	0.4292
	Local	0.0287	0.4234
	FIA-PNW	-0.0866	0.2990



Kozak (2004) variable-exponent taper equations performed better, in terms of RMSE, than the simple polynomial taper equation of Cervera (1973) for Douglas-fir trees. However, Bi (2000) and Arias-Rodil et al. (2014) taper equations, both of which are also variable exponent taper equations, produced higher RMSE compared to the simple taper equation of Cervera (1973). For western hemlock, all the variable-exponent models performed better than the simple polynomial taper function for dataset I. For dataset II, the Cervera (1973) taper equation produced smaller cross validation RMSE than all but Arias-Rodil et al. (2014) equation. This finding is consistent with the findings of Rojo et al. (2005) who compared the performance of 31 taper functions in predicting upper stem diameters for maritime pine in Northwestern Spain. Among the variable-exponent taper equations, Kozak (1988) and Kozak (2004) models performed better than the Bi (2000) and Arias-Rodil (2014) taper equations. Kozak (2004) was chosen as the final model because it produced smaller RMSE values for Douglas-fir, comparable

RMSEs for western hemlock trees, and has been used in the past in several studies to predict upper stem diameters as well as merchantable and total volume. We report the final model parameters based on the taper and local volume equation fitted using the combined dataset. Therefore, we also obtained separate evaluation statistics for each dataset by applying the fitted model to these datasets. These results are presented in Table 10. We observed that the RMSE values were smaller for the datasets with smaller trees compared with the dataset with larger trees for both species. For example, Douglas-fir trees in dataset I and II were both sampled from the western half of states of Oregon and Washington but RMSEs for dataset II, which had small average DBH (37.1 cm vs. 18.2 cm), were less than the RMSE percent for dataset I for all methods. However, the biases were slightly higher for second dataset (Table 10) for all methods.

The volume equations used by FIA-PNW performed fairly well but a volume equation with fewer parameters fitted in this study using DBH and height provided similar

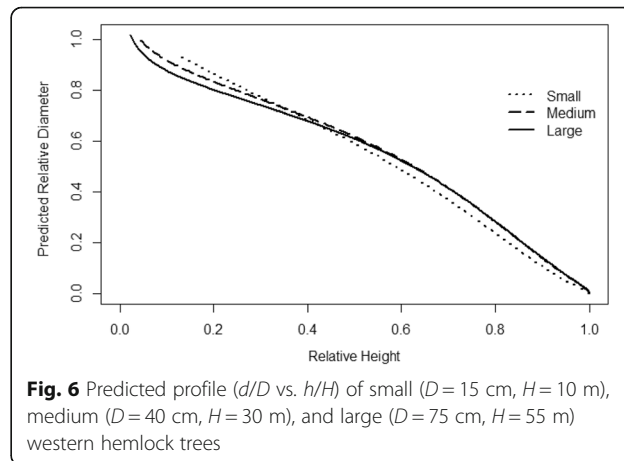
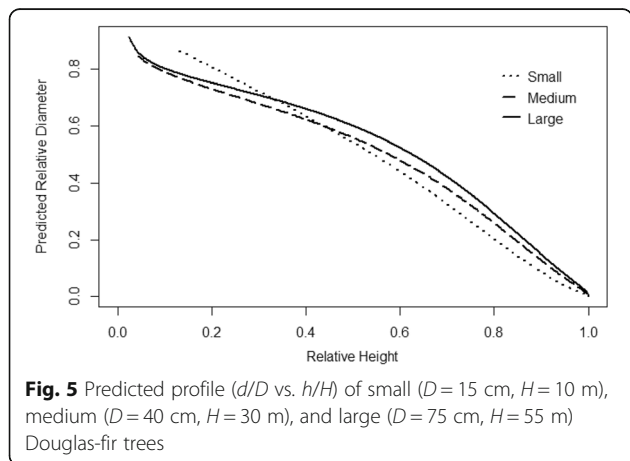


Table 10 Bias and RMSE for estimating cubic volume including top and stump obtained from fitted taper equations (M4), volume equations fitted in this study (Local), and the volume equations used by FIA-PNW for dataset I, II, and III. Final models were fitted using all datasets and evaluation statistics were obtained by applying them to individual dataset

Species	Dataset	Method	Bias (m ³)	RMSE (m ³)
Douglas-fir	I	M4	-0.0277	0.2502
		Local	0.0003	0.2354
		FIA-PNW	0.0688	0.2830
	II	M4	-0.0042	0.0288
		Local	-0.0195	0.0347
		FIA-PNW	-0.0098	0.0310
	III	M4	-0.1713	0.5258
		Local	-0.0563	0.3852
		FIA-PNW	-0.0035	0.3564
Western hemlock	I	M4	-0.0676	0.2850
		Local	-0.0492	0.1773
		FIA-PNW	-0.1504	0.3763
	III	M4	0.1024	0.2706
		Local	0.1411	0.3454
		FIA-PNW	-0.0227	0.2217

RMSE compared with the volume equation used by FIA-PNW. We found that the RMSEs of FIA-PNW volume equations were slightly smaller than the ones produced by the Kozak (2004) taper equation. This result is consistent with the results reported by Poudel (2015) who found CVTS obtained from FIA-PNW volume equations to be more accurate (3% less RMSE) than the estimates obtained from the refitted Kozak (2004) equation for Douglas-fir. In addition, it is expected that the volume equations produced smaller error compared to the taper equations because of the function they minimize.

Taper equations may be preferred over direct volume equations because they account for the variability in stem form and can be used for multiple purposes such as estimating merchantable volume or height. However, the modest gain in precision, as observed in this study, is attained by the increased complexity of the taper equations. Similar studies using the data collected across the region for other species would further help in making decisions on whether the taper-based volume estimates are more accurate than the estimate obtained from existing equations.

Most of the volume equations used by the PNW-FIA program for trees on the west coast use models built in the 1970s and 1980s, many of which employ the TARIF system to estimate the volume of different pre-determined merchantable portions of the tree. It is difficult to use these equations for different merchantability standards

and to incorporate information on trees that have broken tops. More recently, equations have been developed for many species that specifically model the volume of tree boles, and can readily estimate gross volume. The Kozak (2004) variable exponent taper equation produced smaller bias and root mean squared error in predicting upper stem inside bark diameters compared to other taper equations used in this study. CVTS obtained based on this taper equation was very similar to the CVTS estimates obtained from a locally fitted simple volume equation and the volume equations used by the FIA-PNW in the states of Oregon, Washington, and California for Douglas-fir and western hemlock trees. Additionally, the accuracy of a simple volume equation that predicts CVTS as a function of DBH and total tree height was comparable with the FIA-PNW volume equations that has more parameters in the model.

The taper equations fitted in this study provide added benefit to the users over the FIA-PNW volume equations by enabling the users to predict diameter at any height, height to a given diameter, and merchantable volume in addition to CVTS of Douglas-fir and western hemlock trees in the Pacific Northwest. The findings of this study also give more confidence to the users of FIA-PNW volume equations because the volume estimates obtained based on taper equation and the local volume equation were very similar to the ones obtained from the existing FIA-PNW volume equations. However, similar studies performed at the large scale are needed to verify the appropriateness of replacing existing volume equation with the newer sets of taper equations.

Abbreviations

CVTS: Cubic volume including top and stump; DBH: Diameter outside bark at 1.3 m above ground; DIB: Diameter inside bark; FIA-PNW: Forest Inventory and Analysis program in the Pacific Northwest; PDIB: Predicted diameter inside bark

Acknowledgements

We would like to thank Dr. James W. Flewelling for providing insight on interpolation approaches and for providing feedback on manuscript draft.

Funding

The project is funded by the USDA Forest Service, Forest Inventory Analysis Unit.

Availability of data and materials

The data can be available upon request to the Corresponding Author.

Authors' contributions

KP conceived and designed the comparison of the taper functions, wrote the manuscript and R codes, conducted the analyses, THM formulated the idea and contributed significant input into the manuscript through many reviews and supervised the study, ANG provided significant input throughout the manuscript preparation and supervised the study. All authors read and approved the final manuscript.

Authors' information

KP. Poudel is Postdoc Scholar at Oregon State University; H. Temesgen is Professor of Forest Biometrics and Measurements at Oregon State University, and A.N. Gray is Research Ecologist, Team Leader, USDA Forest Service, PNW Research Station.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Forest Engineering, Resources, and Management, College of Forestry, Oregon State University, 280 Peavy Hall, Corvallis, OR 97331, USA.

²USDA Forest Service, PNW Research Station, 3200 SW Jefferson Way, Corvallis, OR 97331, USA.

Received: 4 October 2017 Accepted: 30 January 2018

Published online: 12 March 2018

References

- Arias-Rodil M, Diéguez-Aranda U, Puerta FR, López-Sánchez CA, Libano EL, Obregón AC, Castedo-Dorado F (2014) Modelling and localizing a stem taper function for *Pinus radiata* D. Don in Spain. *Can J For Res* 45:647–658
- Arney JD, Milner KS, Jafvert MA, Vopicka CE (2004) Forest projection and planning system users guide and tutorial. Technical report no. 11. College of Forestry and Conservation, University of Montana, Missoula, MT, Forest Biometrics Research Institute, p 306
- Bi H (2000) Trigonometric variable-form taper equations for Australian eucalypts. *For Sci* 46(3):397–409
- Bluhm AA, Garber SM, Hibbs DE (2007) Taper equation and volume tables for plantation-grown red alder. Gen tech rep PNW-GTR-735. U.S. Department of Agriculture, Forest Service, Portland, OR, p 74
- Brackett M (1977) Notes on TARIF tree-volume computation. Resources management report no. 24. Washington Department of Natural Resources, Olympia, WA, p 26
- Cao QV (2009) Calibrating a segmented taper equation with two diameter measurements. *South J Appl For* 33(2):58–61
- Cervera J (1973) El área basimétrica reducida, el volumen reducido y el perfil. *Montes* 174:415–418
- Chambers CJ, Foltz B (1979) The TARIF system - revisions and additions. Resource management report no. 27. WA Dept of Nat Resources, Olympia, WA, p 8
- Clark AC, Souter RA, Schlaegel BE (1991) Stem profile equations for southern tree species. Res pap SE – 282. U.S. Department of Agriculture, Forest Service, Asheville, NC, p 113
- Czaplewski RL, Brown AS, Walker RC (1989) Profile models for estimating log end diameters in the Rocky Mountain region. Res pap RM-284. U.S. Department of Agriculture, Forest Service, Fort Collins, CO, p 7
- Flewelling JW (1994) Stem form equation development notes. Unpublished report to the northwest taper cooperative. U.S. Department of Agriculture, Forest Service, Forest Management Service Center, Fort Collins, CO
- Flewelling JW, Ernst RL (1996) Stem profile estimation - east-side species. Unpublished report to the inland northwest growth and yield cooperative. U.S. Department of Agriculture, Forest Service, Forest Management Service Center, Fort Collins, CO
- Kozak A (1988) A variable-exponent taper equation. *Can J For Res* 18:1363–1368
- Kozak A (1994) Development of taper equations by BEC zones and species. Unpublished Report. B.C. Ministry of Forests
- Kozak A (2004) My last words on taper equations. *For Chron* 80(4):507–514
- Kublin E, Augustin NH, Lappi J (2008) A flexible regression model for diameter prediction. *Eur J Forest Res* 127:415–428
- Li R, Weiskittel A, Dick AR, Kershaw Jr JA, Seymour RS (2012) Regional stem taper equations for eleven conifer species in the Acadian region of North America: development and assessment. *North J Appl For* 29(1):5–14
- MacLean C, Berger JM (1976) Softwood tree-volume equations for major California species. Res. Note. PNW-266. U.S. Department of Agriculture, Forest Service, Portland, OR, p 34
- Max TA, Burkhart HE (1976) Segmented polynomial regression applied to taper equations. *For Sci* 25:283–289
- Newnham RM (1988) A variable-form taper function. For can Petawawa Natl for inst Inf rep, PI-X-83
- Newnham RM (1992) Variable-form taper function for four Alberta tree species. *Can J For Res* 22:210–223

Pinheiro J, Bates D, DebRoy S, Sarkar D, R Core Team (2016) nlme: linear and nonlinear mixed effects models. R package version 3.1-128. <http://CRAN.R-project.org/package=nlme>. Accessed 20 Sep 2017

Poudel KP (2015) Strategies for sampling and estimation of aboveground tree biomass. Oregon State University, PhD Dissertation, p 144

Poudel KP, Temesgen H (2016) Methods for estimating aboveground biomass and its components for Douglas-fir and lodgepole pine trees. *Can J For Res* 46(1):77–87

Rojo A, Perales X, Sanchez-Rodriguez F, Alvarez-Gonzalez JG, von Gadow K (2005) Stem taper functions for maritime pine (*Pinus pinaster* Ait.) in Galicia (northwestern Spain). *Eur J For Res* 124(3):177–186

Sharma M, Parton J (2009) Modeling stand density effects on taper for jack pine and black spruce plantations using dimensional analysis. *For Sci* 55(3):268–282

Summerfield E (1980) In-house memo describing equations for Douglas-fir and ponderosa pine. State of Washington, Department of Natural Resources, Pacific Northwest Research Station, Portland, OR

Trincado G, Burkhart HE (2006) A generalized approach for modeling and localizing stem profile curves. *For Sci* 52:670–682

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com