Development of height to crown base models for thirteen tree species of the North American Acadian Region

by Baburam Rijal¹, Aaron R. Weiskittel¹ and John A. Kershaw, Jr.²

ABSTRACT

Height to live crown base (HCB) is an important input variable for several growth and yield models. Since HCB is rarely measured in the field, it is often predicted using static models. Instead of predicting HCB, the Forest Vegetation Simulator Northeastern Variant (FVS-NE) uses an equation that predicts crown ratio (CR), which has not been well validated. The main goal of the present study was to construct a regional HCB model for thirteen selected tree species of the Acadian Region of North America. The specific objectives were to: 1) evaluate FVS-NE model predictions, 2) compare suitable model forms, and 3) assess influence of various covariates to improve predictions. We evaluated three model forms, namely Holdaway (1986), logistic, and exponential. The findings indicated that FVS-NE models were significantly biased for all species as the overall mean bias and root mean square error (RMSE) were 0.11 m and 1.80 m, respectively. A logistic equation with size (diameter at breast height [DBH], total height [HT] and ratio of DBH to HT), and competition (crown competition factor [CCF] and basal area larger than subject tree [BAL]) gave the best predictions for all species in this analysis. This model had an overall mean bias <0.01 m and an RMSE of 1.59 m, which represents a significant improvement in predictions compared to FVS-NE. Despite the range of species and observed variation in the data, the equations worked well and can be easily calibrated to new stands with a few local observations.

Key words: crown ratio, Maine, Nova Scotia, New Brunswick, Quebec, nonlinear mixed effects, Forest Vegetation Simulator

RÉSUMÉ

La hauteur de la cime (HC) représente une variable importante pour plusieurs modèles de croissance et de rendement. Compte tenu que la HC est rarement mesurée sur le terrain, elle est souvent prédite au moyen de modèles statiques. Au lieu de prédire la HC, le Forest Vegetation Simulator Northeastern Variant (FVS-NE) utilise une équation qui prédit le ratio de la cime (RC), laquelle n'a pas été validée de façon satisfaisante. L'objectif principal de cette étude visait à établir un modèle régional de HC pour treize espèces retenues de la forêt acadienne de l'Amérique du Nord. Les objectifs spécifiques étaient de : 1) évaluer les prédictions du modèle FVS-NE; 2) comparer les versions utilisables du modèle et 3) évaluer l'influence de différentes co-variables pouvant améliorer les prédictions. Nous avons évalué trois versions du modèle, nommément Holdaway (1986), logistique et exponentielle. Les résultats ont indiqué que les modèles FVS-NE sont significativement biaisés pour toutes les espèces compte tenu que le biais global moyen et l'erreur quadratique moyenne (MSE) était respectivement de 0,11 et de 1,80. Une fonction logistique comprenant la taille (diamètre à hauteur de poitrine [dhp], la hauteur totale [HT] et le ratio entre le dhp et la HT) et la compétition (facteur de compétition de la cime [FCC] et la surface terrière supérieure à l'arbre étudié [STS]) a généré les meilleures prédictions pour toutes les espèces de cette étude. Ce modèle affichait un biais global moyen <0,01 m et une MSE de 1,59 m, ce qui représente une amélioration significative des prédictions par rapport au FVS-NE. Malgré la diversité des espèces et la variation observée dans les données, les équations ont bien fonctionné et peuvent être facilement calibrées pour de nouveaux peuplements au moyen de quelques observations locales.

Mots clés : ratio de la cime, Maine, Nouvelle-Écosse, Nouveau-Brunswick, Québec, variance aléatoire non linéaire, Simulateur de végétation forestière

¹University of Maine, School of Forest Resources, 5755 Nutting Hall, Orono, Maine 04469, USA. E-mail: Baburam_Rijal@umit.maine.edu

²University of New Brunswick, Faculty of Forestry and Environmental Management, PO Box 4400, 28 Dineen Drive, Fredericton, New Brunswick E3B 5A3.

Introduction

Individual tree growth and development is closely related to leaf area (LA). Total LA and its distribution within crowns and canopies influence key physiological processes such as light interception, photosynthesis, and stomatal conductance (Kramer and Kozlowski 1979). Hence, tree growth and stand net primary production (NPP) are often a linear function of LA (Landsberg 2003). This implies that some measure of LA should be an important covariate in any static or dynamic growth and yield equation. However, diameter at breast height (DBH) and total height (HT) are much more commonly used in constructing statistical growth and yield equations, particularly for forest management-orientated applications.

Crown length (CL) and width are common static measures of crown size and are often used as surrogates for LA in tree growth modeling (Weiskittel *et al.* 2011b). Height to crown base (HCB), CL, bole length (BL) as well as crown ratio (CR) and bole ratio (BR) are interrelated by simple algebraic relationships. These attributes have several other important uses in forestry. For example, crown size and position have a strong influence on tree bole form (Larson 1963). Silviculturists often measure crown size both as a response to silvicultural treatments (Curtis and Reukema 1970) and as an index of future growth potential (Hamilton 1969). Furthermore, CR commonly reflects cumulative effects of past management (Wykoff 1990). Consequently, growth modellers routinely employ crown size as an explanatory variable in a variety of individual tree equations. For example, crown size is often used as predictor variable in individual tree diameter increment (Wykoff *et al.* 1982, Ritchie and Hann 1985), height increment (Short and Burkhart 1992), mortality (Burkhart *et al.* 1987), volume (Walters *et al.* 1985), and taper equations (Walters and Hann 1986). In order to calculate or estimate CR, we need to measure HCB or CL. However, measurements of HCB and CL are difficult in dense and/or multilayered stands (Temesgen *et al.* 2005), prone to error due to conflicting definitions (McRoberts *et al.* 1994), and time-consuming; therefore, measurements are often limited. Static HCB equations can be used to predict missing HCB measurements (Hann *et al.* 1997), and also to estimate changes in crown ratio (Stage 1973).

There are numerous studies modeling HCB and CR across a variety of regions and species such as Hasenauer and Monsrud (1996) in Austria, Temesgen *et al.* (2005) in British Columbia, Ritchie and Hann (1987) as well as Zumrawi and Hann (1989) in the US Pacific Northwest, and Liu *et al.* (1995) in the US southeast. In the northeastern US, the forests are quite complex in terms of stand history, condition, structure, and management, but very little work has been done on crown size modeling in the region. Bechtold (2003) as well as Russell and Weiskittel (2011) have worked on modeling stand-grown largest crown diameter for a number of tree species in the region. However, there is limited work on estimating HCB or CR. For example, the Forest Vegetation Simulator (FVS) is a forest growth and yield model widely used in the US and parts of Canada (Crookston and Dixon 2005). The FVS-Northeastern variant (FVS-NE) uses the Holdaway (1986) equations for predicting CR. The parameters for the Holdaway (1986) equations are based on data available from the 1960s and 1970s inventories of the US Lake States, but lit-

tle work has focused on validating these equations across the region where they are currently applied. Furthermore, FVS-NE covers a wide geographic region, extending from Ohio east to New Jersey and north to Maine, encompassing an area with vastly different forest types, climatic zones, geology, and forest management strategies.

The overall goal of this study was to develop species-specific regional HCB equations for the major tree species across the Acadian Region. The specific objectives of the analysis were to: (1) assess and evaluate the effectiveness of the equations and model forms currently used in FVS-NE; (2) compare alternative model forms; and (3) evaluate different tree- and stand-level variables for better predictions of HCB. Throughout the text, we use the term regional model for a prediction equation that includes tree- and stand-level covariates in addition to DBH following the terminology of Soares and Tomé (2002).

Methods

Study area

The Acadian Region encompasses three Atlantic Provinces of Canada (New Brunswick, Nova Scotia, and Prince Edward Island), the southern parts of Quebec, and the State of Maine in the US. The Acadian Forest lies in the transition zone between boreal forests to the north and northern broadleaves to the south (Braun 1950). The forest is dominated by naturally regenerated, multi-species, multi-cohort structures, created by a long history of selection harvests and natural disturbances. However, portions of New Brunswick are characterized by single-species, intensively managed plantations. The major forest types consist of mixed conifer (*Picea-Abies*; spruce-fir), mixed broadleaf (*Populus-Betula*; aspen-birch), mixed broadleaf and conifer (*Quercus-Pinus*; oak-pine) and wetland forest (*Acer-Picea*; maple-spruce). More than 60 tree species are documented in the region. Based on availability in the dataset and commercial importance, 13 species were selected for this study: seven broadleaf and six conifer species (Table 1).

Data

As a part of ongoing research at the University of Maine School of Forest Resources and the University of New Brunswick Faculty of Forestry and Environmental Management, an extensive regional database of tree-level measurements has been compiled for the Acadian Forest Region (Weiskittel *et al.* 2010). The database consists of repeated measures of individual tree DBH and height with nearly 2.9 and 1.6 million observations respectively, in about 16 000 stands across the region. However, there were a substantially smaller number of HCB observations. Overall, there were 252 152 observations of HCB measurement for the 13 species included in this study. Details of the species and summary statistics by species are shown in Table 1.

The database consisted of several sources of data that have been collected by various organizations including the US Forest Service, University of Maine Cooperative Forest Research Unit, and the Departments of Natural Resources of several Canadian provinces. Each data source roughly equates to the different geographic areas across the region. Within each data source, there were a number of stands (locations), which were

Table 1. Summary statistics of diameter at breast height (DBH), total height and height to live crown base (HCB) by species. The species codes are AB: American beech (*Fagus grandifolia* Ehrh.); GB: grey birch (*Betula populifolia* Marsh.); PB: paper birch (*Betula papyrifera* Marsh.); RM: red maple (*Acer rubrum* L.); RO: northern red oak (*Quercus rubra* L.); SM: sugar maple (*Acer saccharum* Marsh.); YB: yellow birch (*Betula alleghaniensis* Britt.); BF: balsam fir (*Abies balsamea* [L.] Mill.); BS: black spruce (*Picea mariana* [Mill.] BSP); EH: eastern hemlock (*Tsuga canadensis* [L.] Carrière); RS: red spruce (*Picea rubens* Sarg.); WP: white pine (*Pinus strobus* L.); and WS: white spruce (*Picea glauca* [Moench] Voss).

Species code	Hierarchy (# of observations)				DBH (cm.)				Height (m.)				HCB (m.)			
	Source	Stand	Year	Total	Min.	Mean	Max.	SD	Min.	Mean	Max.	SD	Min.	Mean	Max.	SD
Broadleaf																
AB	4	227	393	2057	9.1	16.9	53.5	7.5	1.4	9.8	24.0	3.2	0.1	6.1	17.0	2.4
GB	5	252	405	1433	1.3	10.9	25.1	3.8	1.6	10.4	18.6	2.8	0.1	8.5	12.5	2.3
PB	9	1272	2144	10677	0.2	14.8	65.1	6.3	0.9	11.7	30.2	4	0.1	7.1	18.9	2.9
RM	9	2200	3846	30716	0.4	16.5	76.0	6.8	1.5	12.4	27.5	3.3	0.1	7.8	19.2	2.5
RO	1	156	278	1247	9.1	18.7	47.7	6.9	2.0	11.8	23.8	3.4	1.5	7.6	15.7	2.5
SM	5	375	709	5752	3.3	18.8	71.3	8.6	2.5	14.0	27.1	3.4	0.1	8.7	21.0	2.5
YB	8	792	1431	5443	1.5	20.8	84.5	10.5	2.4	12.7	31.5	3.5	0.1	7.1	18.0	2.6
Conifers																
BF	10	2764	4974	99489	0.2	13.0	46.9	5.3	0.6	9.4	49.0	3.4	0.1	5.0	18.0	2.5
BS	5	829	1798	25443	0.2	12.1	51.7	5.2	0.2	8.3	26.5	3.1	0.1	4.7	17.5	2.6
EH	7	237	399	3008	0.5	22.6	88.8	11.7	1.5	12.8	27.9	4.7	0.2	6.2	17.6	3.1
RS	6	1943	3344	47269	0.4	15.5	107.0	7.3	1.0	11.3	55.0	3.9	0.1	6.6	20.0	3.2
WP	9	685	1179	5562	1.4	24.3	104.0	13.9	0.6	13.9	33.7	5.1	0.1	7.9	23.5	3.8
WS	8	1054	2275	14056	0.3	16.1	61.0	7.4	1.1	10.3	28.4	3.9	0.1	5.4	21.5	3.0

further divided into plots that contained individual tree measurements. Numerous plots (>75%) were permanent research installations and repeatedly measured. Although the various datasets used slightly different definitions of crown base, preliminary analysis indicated that the differing definitions did not strongly influence the results, likely because of the high measurement error associated with determining crown base. All datasets used in this analysis actually measured HCB.

Model forms

A variety of linear and nonlinear model forms were evaluated in this analysis. Initially, the FVS-NE model form of Holdaway (1986) (eq. 1) was evaluated:

$$[1] \quad CR = 10 * \left[\frac{b_1}{(1 + b_2 * BAPH)} + b_3 * (1 - e^{b_4 * DBH}) \right]$$

where CR is crown ratio (%), DBH is diameter at breast height outside bark (cm), BAPH is basal area ($m^2 ha^{-1}$), and the b_i 's are model parameters. In this analysis, the species parameters were obtained from the most current version of FVS-NE (Dixon *et al.* 2009: 18) and by refitting equations. We evaluated the FVS model for HCB by calculating HCB from the predicted CR and the measured height. During preliminary analysis, it was found that the HCB bias was much smaller and the HCB predictions were more precise than the CR predictions. Consequently, we refitted the model using both CR and HCB and recalculated bias and precision. Our diagnostic analyses indicated that HCB rather than CR had a stronger statistical relationship with DBH and HT (Fig. 1). Since HCB is a dimensional measurement of a tree and CR is a relative measurement, it is not surprising that HCB would

be more strongly correlated with other dimensional measures such as DBH and HT.

Two other model forms were selected for comparison: the logistic equation (eq. 2) and the exponential equation (eq. 3). Both of these model forms are bound by zero from below and cannot be larger than HT. The general forms of the logistic and exponential functions, respectively, are:

$$[2] \quad HCB = \frac{HT}{[1 + c * e^{-k * X}]^{1/m}}$$

where c , k , and m are model parameters;

$$[3] \quad HCB = HT * \left[1 - c * e^{-k * X^w} \right]$$

where c , k , and w are model parameters and X is a linear combination of covariates expressed as: $X = \sum_{i=0}^k \beta_i x_i$ (for $i = 0, x_i = 1$). Initially, it was attempted to fit m and w as model parameters, but model convergence problems were experienced. As suggested by Soares and Tomé (2001), the Richards and Weibull equations are special forms of logistic (eq. 1) and exponential (eq. 2) equations, respectively, where the parameters c and k are set to 1, and parameters m and w are set to 6 and 10. Among the different model forms (logistic, exponential, Richards, and Weibull), our initial results indicated that the Richards form yielded better fits in terms of different model selection criteria (discussed below) when compared to their base model forms (results not reported here) and we did not experience any convergence problems in any of the species. Therefore, the Richards equation was taken as the most appropriate form for this analysis. In order to differentiate and describe this "Richards" form of the equation that has specific values for the constants of the logistic

Fig. 1. Height to live crown base (HCB) and crown ratio (CR) over diameter at breast height (DBH) and total height (HT) with a Lowess regression line and Pearson correlation coefficient (r) in the upper right corner. The top panel is for broadleaves and bottom panel is for conifers.

equation, we present the equation separately in equation 4. For ease, we refer to it as “Richards” in the Results and Discussion sections (below). Hence, the Richards equation is presented as below (eq. 4):

$$[4] \quad HCB = \frac{HT}{[1 + c \cdot e^{-k \cdot X}]^{1/m}}$$

where c and $k=1$, $m = 6$, and X is as above.

Addition of covariates

We added tree- and stand-level covariates to capture within- and between-stand-level variability (Table 2). We categorized the covariates into three groups as used by Hasenauer and Monserud (1996) and Temesgen *et al.* (2005). These were (a) size, (b) competition, and (c) site-specific covariates. Size-related attributes were DBH, HT, and the ratio between the two (DHR). Initially, HCB models were tested with different transformations of DBH including the square, natural logarithm, and a power function, but the linear form was found to work best.

Competition-related covariates included two-sided measures of crown competition factor (CCF) (Krajicek *et al.* 1961) and total basal area (BAPH; $m^2 ha^{-1}$), as well as one-sided

measures such as basal area in larger trees (BAL; $m^2 ha^{-1}$) and crown competition factors in larger trees (CCFL). These covariates were tested in various combinations and preliminary evaluation showed that CCF and BAL were best (Table 3). CCF was estimated using species-specific maximum crown width equations developed for the region (Russell and Weiskittel 2011).

Site-specific covariates included climatic site index (CSI) and geographic factors that contained elevation, aspect, slope, latitude, and longitude (Table 2). Since site index is rarely measured in the Acadian Region and detailed soil maps do not exist for much of the region, CSI was derived based on the geographical location of sample plots. This index was based on 1-km² climate normals from 1960 to 1991 and an empirically derived relationship with observed site index (Weiskittel *et al.* 2011a). We also tried the Hopkins (1938) Index to transform elevation, latitude, and longitude. For slope and aspect, we use the transformations suggested by Stage (1976) and Wyckoff (1990). Addition of these geographic attributes to a model that included CSI did not improve predictions and they were subsequently dropped. When all covariates were added, our full model was as given in eq. [4] and where X was:

Table 2. Model forms and covariates used in the study.

Model code	Model form	Covariates included ^a	Covariates in initial examination
Model 1	FVS	DBH, BAPH	–
Model 2	FVS refitted	DBH, BAPH	–
Model 3	FVS-HCB	DBH	–
Model 4	Richards base	DBH	DBH, DBH ² , DBH ^{0.5} , log(DBH)
Model 5	Richards size covariates	DBH, HT and DHR	DBH and HT, DBH, HT and DHR
Model 6	Richards with size and competition	DBH, CCF and BAL	DBH, CCF and BAL DBH, CCF and CCFL DBH, CCF and BAPH DBH, BAPH and BAL DBH, BAPH and CCFL DBH, CCFL and BAL
Model 7	Richards with size and competition	DBH, HT, DHR, CCF and BAL	DBH, DHR, CCF and BAL DBH, DHR, CCF and CCFL DBH, DHR, CCF and BAPH DBH, DHR, BAPH and BAL DBH, DHR, BAPH and CCFL DBH, DHR, CCFL and BAL
Model 8	Richards with size, competition, and site	DBH, CCF, BAL, ELEV, SLOPE, ASPECT and CSI	DBH, CCF, BAL, ELEV, SLOPE, ASPECT, and CSI
Model 9	Richards with size, competition, and site	DBH, HT, HDR, CCF, BAL and CSI	DBH, HDR, CCF, BAL, ELEV, SLOPE, ASPECT, CSI, LAT and LONG

^a DBH – diameter at breast height (cm), HT – total height (m), BAPH – stand basal area (m² ha⁻¹), DHR – DBH/HT, HDR – HT/DBH, CCF – crown competition factor, CCFL – CCF larger than subject tree, BAL – basal area larger than subject tree (m² ha⁻¹), ELEV – Elevation (m), CSI – Climatic site index, LAT – Latitude (°), LONG – Longitude (°), SLOPE – percent slope (%), and ASPECT – aspect (°).

$$[5] \quad X = \sum_{i=0}^6 b_i x_i = (\beta_0 + \beta_1 * DBH + \beta_2 * HT + \beta_3 * HDR + \beta_4 * \log(CCF) + \beta_5 * \log(BAL) + \beta_6 * CSI)$$

Natural logarithm transformations were used for some covariates because they were found to improve model fits.

Parameter estimation

Initially, general nonlinear least squares (GNLS) techniques were used to estimate and compare between the different base model forms. Because there was evidence of heteroskedastic residual variances, the models were fitted using a variance-weighting function. Among the tested weighting functions, a power function of DBH gave better fits: $\text{Var}(\varepsilon_i) = \sigma^2 \text{DBH}_i^\delta$, where $\text{Var}(\varepsilon_i)$ is residual variance for the i^{th} tree, σ^2 is the mean residual sums of squares, and δ is the variance function coefficient. Likewise, models were tested with different correlation structures ($\Gamma\alpha$) to test for spatial and temporal autocorrelations. Preliminary assessment showed that modeling with only variance power produced unbiased predictions and there was no gain in model performance when different correlation structures were added. Therefore, our GNLS models were fitted with an assumption of independence between observations among repeated measurements.

Because of the repeated and hierarchical nature of the data, we applied a nonlinear hierarchical mixed effects

(NLME) approach to fit the best model selected from the GNLS procedure. The model was fitted with three levels of hierarchy: data source, stand, and year. Although the hierarchy could be comprised of five levels that would include plot and individual tree, we used three levels as the plot and individual tree levels provided no significant model improvement and vastly increased model complexity (results not shown). We assumed our data were independent within and between source and stands. Preliminary analyses also indicated that treating the intercept (β_0) as mixed worked best and it was computationally efficient. This is similar to the model forms derived by Weiskittel *et al.* (2007).

The models were fitted using the “nlme” package (Pinheiro and Bates 2009) of the R statistical package (R Development Core Team 2011) using restricted maximum likelihood estimation (REML). In assessing NLME model performance, both fixed effects only and fixed plus random effects were used since the fixed effects will generally be used for prediction, unless local observations are available for calibration.

Model selection

The entire dataset was used for model construction since data-splitting techniques are rarely effective for relatively large datasets such as used in this analysis (Kozak and Kozak 2003) and parameters estimated from the entire dataset are more precise than those derived from portions of split data (Hirsch

Table 3. Mean bias (MB; Bias = observed–predicted), mean absolute bias (MAB), root mean square error (RMSE), Bayesian Information Criteria (BIC), and coefficients of determination (R^2) for FVS-NE crown ratio model parameter suggested by Dixon *et al.* (2009; p. 18), refitted using CR, and refitted using height to crown base (HCB). Parameters were estimated by general nonlinear least square (GNLS). Bold indicates values significantly different from zero at $p = 0.05$ based on bootstrap sampling.

Species	FVS-NE ^a (original parameters in Dixon <i>et al.</i> 2009) (Model 1)					Refitted with crown ratio (Model 2)					Refitted with HCB (Model 3)				
	MB	MAB	RMSE	MB	MAB	RMSE	BIC	R^2	MB	MAB	RMSE	BIC	R^2	MB	MAB
Broadleaf															
AB	0.2219	1.2405	1.5788	-0.0580	1.1812	1.5374	16968	0.03	-0.0252	1.1742	1.5287	7471	0.58		
GB	-0.0651	0.9250	1.2028	-0.0029	0.9181	1.1987	10798	0.14	-0.0059	0.9324	1.207	4428	0.73		
PB	-0.1708	1.0519	1.4142	0.0314	1.0539	1.4017	78042	0.14	-0.0543	1.06294	1.41201	34287	0.75		
RM	1.0640	1.5199	1.8533	-0.0533	1.1318	1.5024	225543	0.05	-0.033	1.14279	1.51244	104206	0.63		
RO	0.5916	1.3408	1.6279	-0.0047	1.1635	1.4528	9431	0.03	0.0048	1.16085	1.4436	4284	0.67		
SM	1.1008	1.5904	1.9379	-0.1257	1.1643	1.5489	43173	0.07	-0.0276	1.1894	1.5495	20458	0.63		
YB	0.1935	1.3010	1.6683	-0.0848	1.2568	1.6324	42538	0.06	-0.0587	1.25337	1.62816	20045	0.62		
Average	0.4194	1.2814	1.6119	-0.0426	1.1242	1.4677	–	0.08	-0.0286	1.1309	1.4688	–	0.66		
Conifers															
BF	0.2281	1.2420	1.5794	-0.0482	1.3281	1.6585	8333978	0.11	-0.0039	1.20463	1.525	370606	0.58		
BS	-1.0712	2.0930	2.6743	0.0466	1.4412	1.8543	222859	0.01	-0.0452	1.45169	1.89323	99105	0.49		
EH	-0.7576	1.7501	2.3488	-0.0415	1.6737	2.1268	24512	0.19	0.03425	1.62383	2.03571	12381	0.56		
RS	0.0733	1.3560	1.7349	-0.8007	1.4934	1.9035	391099	0.03	-0.1799	1.36761	1.72158	177676	0.72		
WP	-0.6534	1.4743	1.8968	-0.6498	1.4961	1.9313	43149	0.21	-0.1384	1.43663	1.81798	20993	0.78		
WS	0.7705	1.4314	1.8500	0.0884	1.2643	1.6279	116986	0.24	-0.4103	1.3083	1.6757	52977	0.70		
Average	-0.2351	1.5578	2.0140	-0.2342	1.4495	1.8504	–	0.13	-0.1239	1.3988	1.7782	–	0.64		

^aFVS-NE model form: Holdaway (1986) equation: $CR = 10 * \left[\frac{b_1}{(1 + b_2 * BAPH)} + b_3 * (1 - e^{b_4 * DBH}) \right]$, where, CR = crown ratio (%), BAPH = stand basal area ($m^2 ha^{-1}$), DBH = diameter at breast height (cm), and b_1, b_2, b_3 , and b_4 are model parameters. Please note that MB, MAB and RMSE in Models 1 and 2 are calculated based on converted values of predicted CR to HCB, and the Models 1, 2 and 3 are compared based on those statistics calculated from predicted HCB

1991). Prediction mean bias (MB), mean absolute bias (MAB), and root mean square error of prediction (RMSE) were used as major model selection criteria. Bias is defined as the difference between observed and predicted response variables. The MB was calculated as:

$$[6] \quad MB = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)$$

where y_i is observed $\hat{y}_i$ and is predicted response values, respectively and n is number of observations. Likewise, MAB was calculated as:

$$[7] \quad MAB = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

and RMSE was calculated as:

$$[8] \quad RMSE = \frac{1}{n-k-1} \sqrt{\sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

where k is the number of parameters.

The bootstrap method (Effron and Tibshirani 1993) was used to assess the distributions (mean and precision) of MB, MAB and RMSE. The models were compared based on 95% confidence bands of the various model selection criteria. When calculating these statistics in Models 1 and 2, we calculated HCB from predicted CR. Models were compared based on the RMSE, MAB, MB calculated from the predicted HCB. Bayesian Information Criteria (BIC; Schwarz 1978) were also used when the major model selection criteria were not significantly different. Since coefficient of determination (R^2) has been used in the past as a primary model selection criterion, we also computed a generalized R^2 (Kvålseth 1985) to compare our models to those previous results. Finally, residual analyses were also conducted to ensure there were no obvious biases across the range of predicted HCB.

Results

FVS evaluation

The FVS model (Model 1, Table 2) with parameters suggested by Dixon *et al.* (2009) gave significant ($p < .05$) bias for HCB prediction in both broadleaf and conifer species. Species-specific evaluation showed that the model yields biased predictions for all species, except grey birch (*Betula populifolia* Marsh.; GB). The maximum mean bias was found in sugar maple (*Acer saccharum* Marsh.; SM) (1.10 m) and minimum bias was in GB (0.06 m). Mean absolute bias varied from 0.9 m (GB) to 2.09 m (eastern hemlock (*Tsuga canadensis* [L.] Carrière; EH). Of the thirteen species evaluated, the model significantly underestimated HCB for eight of these species (Table 3). The four species where predictions were overestimated were paper birch (*Betula papyrifera* Marsh.; PB), black spruce (*Picea mariana* [Mill.] BSP; BS), EH, and white pine (*Pinus strobus* L.; WP). The nature of the model biases varied across the predicted ranges. In some species, like American beech (*Fagus grandifolia* Ehrh.; AB) and BS, biases were larger at the two extremes, while the magnitude (either positive or negative) increased with predicted HCB in other species, like red maple (RM) and balsam fir (*Abies balsamea* [L.] BF) (Fig. 2).

The data were used to refit the FVS equation and it gave unbiased mean bias estimates for five of the thirteen species (Table 3). However, two parameters (b_3 and b_4) were non-

significant in all species, except AB, GB, and northern red oak (*Quercus rubra* L.; RO). Using HCB rather than CR as the dependent variable produced significant parameters for all species, except AB, GB, and RM. Regardless, the refitted FVS model (Model 2) did not produce unbiased CR predictions for species such as PB, RM, yellow birch (*Betula alleghaniensis* Britt.; YB), BF, red spruce (*Picea rubens* Sarg.; RS), WP, and white spruce (*Picea glauca* [Moench] Voss; WS) (Table 3). HCB refitted models (Model 3) were found to be more precise and produced unbiased predictions in six species (Table 3). Although mean bias was generally improved for some species in the data, the bias magnitude varied significantly across the predicted HCB ranges (Fig 2).

Model form

The Richards equation (Model 4) was the best model form among those tested because it gave unbiased predictions for six of the thirteen species with higher precision. The base model (that included only DBH as covariate) gave biased predictions for all conifer species (except EH) and was unbiased for all broadleaf species, except PB and YB (Table 4). In comparison to the FVS model, it gave lower mean biases in all species except BF; however, the nature of the biases across the predicted ranges varied (Fig. 2). Although RMSE was slightly higher for the Richards model form compared to the FVS model, it is likely due to having a fewer number of parameters. Like Model 3, when BAPH was added to Model 4, it performed superior to Model 3 in terms of all model selection criteria (data not shown). In order to improve the model, we added other size- and competition-related covariates.

Addition of covariates

As explained in the Methods section, we evaluated several covariates with various different combinations (Table 2). The most suitable covariates were found to be HT, DHR, CCF, BAL and CSI. Adding tree size covariates (HT, DHR) yielded unbiased predictions for all species and the improvement in RMSE, MAB and R^2 were significantly higher for conifers (data not shown). When competition-related variables (CCF and BAL) were added to the equation, RMSE and MAB were further significantly improved in all conifer species (Table 4). When the site-specific covariate (CSI) was added, almost the same magnitude of bias and precision was produced as for models that used CCF and BAL. Overall, these results showed that the optimum model would be one that used both tree size and competition factors.

The selected models fitted by GNLS had an R^2 that ranged from 0.47 in EH to 0.79 in WP. The range of R^2 among the seven broadleaf species was much narrower (0.60–0.77) than the six conifers (0.47–0.79). When the models were fitted using NLME, a noticeably higher R^2 was achieved in the conifers (up to 0.88). The NLME procedure with the random effects gave unbiased predictions with much higher precision (Table 5). However, the analysis showed that the fixed-only effect in the mixed effect approach did not produce unbiased predictions and the precision was also much less than the GNLS techniques (Table 5). This finding is consistent with the results of Temesgen *et al.* (2008) and Meng *et al.* (2009). We have presented parameters obtained from both GNLS and NLME in Tables 6 and 7, respectively. GNLS parameters should be used when no observations are available for local calibration.

Fig. 2. Height to live crown base residual plots (observed – predicted; m) of the Forest Vegetation Simulator Northeast variant (FVS-NE) with parameter estimates from Dixon *et al.* (2009), refitted FVS, and Richards base (with only DBH in covariate) and Richards base plus competition related covariates for a few primary species in this analysis. The lines are Lowess regression lines of the biases. The parameters were estimated by generalized nonlinear least square (GNLS) techniques.

Discussion

CR is an important variable for a variety of reasons, and localized equations to accurately predict CR in the Acadian Region are currently lacking. FVS-NE currently predicts CR, but it covers a rather broad geographic region. In addition, it was developed based on data with CR measurements, which were often visually estimated. Finally, the model form used by FVS-NE is exponential for DBH, but linear in BAPH. This might be one reason why the predictions were found to be significantly biased for all species, except GB, using the parameters suggested by Dixon *et al.* (2009). The model form

even remained biased for most species even after being refitted to our dataset. In general, we found that the greatest biases were obtained for larger trees in low-density stands, which was still observed when the FVS-NE equation was refitted to our data. The FVS-NE equation is also unconstrained, which can produce strange behaviour. For example, predicted CR may range from negative values to more than 100% due to being unconstrained. Local adjustment is possible by self-calibration in FVS (Crookston and Dixon 2005), but it may not resolve prediction biases. Consequently, these results suggested that a new model form would be desirable.

Table 4. Mean bias (MB, Bias = observed – predicted), mean absolute bias (MAB), root mean square error (RMSE), Bayesian Information Criteria (BIC), and coefficients of determination (R^2) for the Richards base model^a (DBH only; Model 4) and with added size and competition related covariates (Model 7). Parameters were estimated by general nonlinear least square (GNLS). Bold values in MB are significant at $p = 0.05$. Single star (*) in MAB and RMSE columns show the respective value that are significantly different from Models 4, 5 and 6. Star in BIC in Model 7 show that the values are within 10 units of the BIC in Model 9 (Model 9 not shown). Please see Table 2 for details about model forms for respective model numbers.

Species	Richards base model (Model 4)					Richards size and competition (Model 7)				
	MB	MAB	RMSE	BIC	R^2	MB	MAB	RMSE	BIC	R^2
Broadleaf										
AB	0.0231	1.1779	1.5334	7401	0.58	-0.0031	1.1368	1.4972	7334	0.59
GB	-0.0238	1.0016	1.3062	4578	0.69	0.0052	0.9060	1.1878	4337*	0.74
PB	-0.0783	1.1176	1.4775	34934	0.73	0.0020	1.0231	1.3742	33325*	0.77
RM	-0.0085	1.1533	1.5234	104032	0.62	0.0000	1.1276	1.4954	102862	0.64
RO	-0.0100	1.1712	1.4579	4272	0.66	0.0012	1.1540	1.4388	4269	0.67
SM	-0.0303	1.1679	1.5495	20480	0.63	-0.0009	1.1526	1.5276	20358*	0.64
YB	-0.0684	1.2806	1.6588	20182	0.60	-0.0025	1.2339	1.6069	19830*	0.63
Average	-0.0280	1.1529	1.5010	–	0.64	0.0003	1.1048	1.4468	–	0.67
Conifers										
BF	-0.0391	1.3229	1.6506	366319	0.57	0.0111	1.2426*	1.5676*	353492*	0.61
BS	-0.0462	1.4343	1.8553	95947	0.48	-0.0058	1.3941*	1.7896*	94731*	0.52
EH	0.0047	1.9184	2.3852	12900	0.40	0.0475	1.7523*	2.2506*	12474	0.47
RS	-0.1856	1.4241	1.7862	179788	0.70	-0.0020	1.1775*	1.5164*	163424*	0.78
WP	-0.2258	1.6163	2.0360	21956	0.72	0.0094	1.3715*	1.7790*	7334	0.79
WS	-0.1399	1.4475	1.8049	54264	0.65	0.0039	1.2449*	1.6054*	4337*	0.72
Average	-0.1053	1.5272	1.9197	–	0.59	0.0107	1.3638*	1.7514	–	0.65

^aRichards equation: $HCB = \frac{H^c}{[1 + c * e^{-k * X}]^{1/m}}$, where c and $k = 1$, $m = 6$, and X is linear combination of covariates

Table 5. Mean bias (MB; Bias = observed – predicted), mean absolute bias (MAB), root mean square error (RMSE) and coefficients of determination (R^2) for the final model when parameters were estimated by generalized nonlinear least squares (GNLS) and nonlinear mixed effects with and without the inclusion of the random effects (NLME). Bold letters in MB indicate significant bias at $\alpha = 0.05$.

Species	MB			MAB			RMSE			R^2		
	GNLS	Fixed	Fixed + random	GNLS	Fixed	Fixed + random	GNLS	Fixed	Fixed + random	GNLS	Fixed	Fixed + random
Broadleaf												
AB	-0.003	0.256	-0.021	1.137	1.200	0.831	1.497	1.537	1.108	0.59	0.57	0.73
GB	0.005	1.392	-0.013	0.906	1.620	0.603	1.188	1.940	0.800	0.74	0.64	0.80
PB	0.002	0.809	-0.017	1.023	1.304	0.712	1.374	1.661	0.965	0.77	0.66	0.86
RM	0.000	2.162	-0.012	1.128	2.346	0.764	1.495	2.713	1.034	0.64	0.22	0.80
RO	0.001	0.012	-0.013	1.154	1.156	0.764	1.439	1.444	0.974	0.67	0.67	0.81
SM	-0.001	0.298	-0.012	1.153	1.216	0.842	1.528	1.586	1.142	0.64	0.61	0.77
YB	-0.003	0.127	-0.033	1.234	1.270	0.883	1.607	1.640	1.170	0.63	0.61	0.73
Conifers												
BF	0.011	0.692	-0.012	1.243	1.405	0.838	1.568	1.776	1.100	0.61	0.50	0.80
BS	-0.006	0.762	-0.037	1.394	1.763	0.923	1.790	2.176	1.202	0.52	0.41	0.88
EH	0.048	1.616	-0.020	1.752	2.366	1.005	2.251	2.869	1.333	0.47	0.30	0.70
RS	-0.002	1.997	-0.006	1.178	2.361	0.802	1.516	2.875	1.071	0.78	0.61	0.88
WP	0.009	1.422	-0.024	1.372	2.034	0.905	1.779	2.536	1.193	0.79	0.56	0.88
WS	0.004	0.382	-0.032	1.245	1.330	0.819	1.605	1.713	1.094	0.72	0.68	0.84

Table 6. Estimated parameters for the final model^a using generalized nonlinear least squares (GNLS) with variance power function on diameter at breast height (DBH). Non-significant parameters at $\alpha = 0.05$ have not been displayed. Standard errors of the parameters are shown in parentheses. The standard deviation of the residual is also given.

Species	β_0	β_1 (DBH)	β_2 (HT)	β_3 (DHR)	β_4 (CCF)	β_5 (BAL)	Var. Power	Res. SD
Broadleaf								
AB	-8.7828 (0.7590)	-0.0931 (0.0183)	0.1009 (0.0333)	0.8791 (0.1919)	0.7188 (0.1000)	0.1797 (0.0568)	0.36 –	0.5546 –
GB	-5.0086 (0.6851)	0.1242 (0.0432)	-0.1436 (0.0439)	-1.3013 (0.4327)	0.2849 (0.0961)	0.7285 (0.0675)	0.37 –	0.5008 –
PB	-6.0534 (0.1784)	-0.0322 (0.0102)	0.0778 (0.0135)	– –	0.3194 (0.0259)	0.3031 (0.0239)	0.35 –	0.5440 –
RM	-5.7729 (0.1524)	-0.0185 (0.0054)	-0.0110 (0.0076)	0.1851 (0.0684)	0.4261 (0.0214)	0.2380 (0.0135)	0.29 –	0.6645 –
RO	-6.1177 (0.8234)	-0.0526 (0.0228)	– –	– –	0.6235 (0.1226)	– –	0.30 –	0.6001 –
SM	-5.1140 (0.3894)	-0.0306 (0.0109)	0.0339 (0.0157)	– –	0.3484 (0.0470)	0.1217 (0.0280)	0.12 –	1.0847 –
YB	-7.2102 (0.3936)	-0.0209 (0.0101)	0.0668 (0.0170)	– –	0.4978 (0.0518)	0.1773 0.0300	0.21 –	0.8502 –
Conifers								
BF	-5.2493 (0.0650)	-0.0191 (0.0042)	0.0314 (0.0054)	-0.1048 (0.0408)	-0.0598 (0.0115)	0.7502 (0.0125)	0.47 –	0.4732 –
BS	-6.2261 (0.1802)	-0.3357 (0.0162)	0.5609 (0.0217)	2.0556 (0.1310)	-0.2905 (0.0275)	0.2257 (0.0373)	0.57 –	0.4392 –
EH	-16.3324 (0.8305)	– –	– –	0.8278 (0.2483)	1.1744 (0.1236)	1.1608 (0.0690)	0.55 –	0.4088 –
RS	-2.2605 (0.1093)	0.1424 (0.0036)	-0.0704 (0.0046)	-3.9084 (0.0543)	0.3163 (0.0148)	0.3263 (0.0107)	0.33 –	0.6205 –
WP	-4.5285 (0.3514)	0.0662 (0.0077)	-0.0260 (0.0127)	-1.9907 (0.1291)	0.3989 (0.0419)	0.3599 (0.0221)	0.42 –	0.4756 –
WS	-5.6740 (0.2504)	0.0584 (0.0122)	– –	-1.7613 (0.1456)	0.2620 (0.0338)	0.6033 (0.0269)	0.44 –	0.4803 –

^aThe model form is: $HCB = \frac{HT}{[1 + e^{\exp(-\beta_0 - \beta_1 \cdot DBH - \beta_2 \cdot HT - \beta_3 \cdot DHR + \beta_4 \cdot \log(CCF + 1) - \beta_5 \cdot BAL)]^{1/6}}}$, where HCB is height to live crown base (m), HT is total height (m), DBH is diameter at breast height (cm), CCF is crown competition factor (%), DHR is the diameter to height ratio, and BAL is basal area larger than subject tree ($m^2 ha^{-1}$), and β_i are model parameters (fixed).

Using a logistic form to model CR and HCB ensured that predictions were constrained between 0 and 1, and 0 and HT, respectively. The logistic form also significantly reduced biases and we did not observe any problems of convergence. Adding HT and DHR to the base model further improved HCB predictions for all species, particularly for RS, WP, and WS. Adding CCF and BAL further significantly reduced RMSE and MAB for YB and all conifer species. Hence, the addition of competition-related covariates helped improve model precision more in conifers than in broadleaf species. In contrast, adding CSI as a site-specific covariate slightly reduced RMSE in all conifers, but the improvement was not statistically significant. In an earlier study on individual tree

height and diameter modeling (H-D), CSI significantly improved prediction bias and precision (Rijal *et al.* in review). A reason for this might be that the H-D model was based on more geographically diverse data than the HCB models. CSI may still be an important covariate in the regional HCB model.

There are a number of studies modeling CR and HCB across the US and Europe. Model performance has generally been measured using coefficient of determination (R^2). Ritchie and Hann (1987) constructed a HCB model using a logistic form for 14 tree species in the US Pacific Northwest and the R^2 values ranged 0.03 to 0.63. The model covariates in Ritchie and Hann (1987) included HT, CCFL, BAPH, and site

Table 7. Estimated parameters for the final model^a using nonlinear mixed effects (NLME) with a variance power function on diameter at breast height (DBH). Non-significant parameters at $\alpha = 0.05$ have not been displayed. Standard errors of the parameters are shown in parentheses. The standard deviation of the residual is also given.

Species	β_0	β_1 (DBH)	β_2 (HT)	β_3 (DHR)	β_4 (CCF)	β_5 (BAL)	Var. Power	Res. SD
Broadleaf								
AB	-8.5175 (1.2795)	-0.0654 (0.0184)	– –	0.9843 (0.1824)	0.4806 (0.2061)	0.4481 (0.0851)	0.42 –	0.3621 –
GB	-5.6799 (1.1072)	0.2112 (0.0399)	-0.3148 (0.0439)	-2.3681 (0.4189)	0.6558 (0.1489)	0.3621 (0.1211)	0.18 –	0.5811 –
PB	-7.7631 (0.4894)	-0.0307 (0.0100)	– –	0.1567 (0.1074)	0.5832 (0.0589)	0.2839 (0.0337)	0.30 –	0.4700 –
RM	-6.9749 (0.5396)	– –	-0.0854 (0.0072)	-0.1391 (0.0599)	0.5178 (0.0405)	0.1251 (0.0177)	0.31 –	0.4658 –
RO	-7.8267 (1.1880)	-0.0649 (0.0215)	– –	0.4662 (0.2402)	0.8805 (0.1935)	– –	0.30 –	0.4486 –
SM	-4.2863 (0.6230)	– –	-0.0477 (0.0154)	– (0.0951)	0.2882 (0.0320)	0.1628 –	0.10 –	0.9021 –
YB	-5.4637 (0.4985)	– –	-0.0372 (0.0171)	– –	0.3407 (0.0766)	0.1953 (0.0362)	0.17 –	0.7746 –
Conifers								
BF	-9.2994 (0.6007)	-0.0450 (0.0041)	-0.0411 (0.0054)	0.5525 (0.0364)	0.7053 (0.0187)	0.3344 (0.0131)	0.39 –	0.4169 –
BS	-11.8533 (1.5162)	-0.2493 (0.0143)	0.3524 (0.0192)	2.8216 (0.1159)	0.5537 (0.0484)	0.4554 (0.0377)	0.44 –	0.4165 –
EH	-10.3208 (1.1431)	-0.0706 (0.0143)	0.0911 (0.0237)	0.5472 (0.1874)	0.5030 (0.1362)	0.1776 (0.0546)	0.41 –	0.3878 –
RS	-2.1081 (0.9299)	0.0715 (0.0045)	-0.1559 (0.0060)	-2.0333 (0.0581)	– –	0.1432 (0.0131)	0.54 –	0.2542 –
WP	-3.4407 (0.6103)	0.0774 (0.0073)	-0.1499 (0.0118)	-1.7585 (0.1201)	0.2655 (0.0530)	0.2474 (0.0257)	0.40 –	0.3651 –
WS	-10.9675 (1.0830)	0.0291 (0.0116)	-0.0529 (0.0170)	-1.0755 (0.1270)	1.2757 (0.0716)	0.3220 (0.0360)	0.55 –	0.2522 –

^aThe model form is: $HCB = \frac{HT}{[1 + \exp^{-d(\beta_0 + \beta_1 * DBH + \beta_2 * HT + \beta_3 * DHR + \beta_4 * \log(CCF + 1) + \beta_5 * BAL)}]^{1/\theta}}$, where HCB is height to live crown base, HT is total height (m), DBH is

diameter at breast height (cm), CCF is crown competition factor (%), DHR is the diameter to height ratio, and BAL is basal area larger than subject tree ($m^2 ha^{-1}$) and β_i are model fixed parameters, and b_1 , b_2 and b_3 are source, stand, and year level random effects, respectively.

index (SI). In contrast, Zumrawi and Hann (1989) modeled bole ratio (BR) for six species and obtained R^2 between 0.16 and 0.33 using HT, CCF, BAPH, and DHR as covariates. For modeling *Eucalyptus* CR in Portugal, Soares and Tomé (2001) used DBH, age, number of trees, SI, and stand dominant height, which explained 74% to 77% of the original variation. Likewise, Dyer and Burkhart (1987) developed a CR model using age as a covariate, but this effectively limited the equation to even-aged stands. Leites *et al.* (2009) evaluated FVS-NI (north Idaho) variants for CR. The equations consisted of the logistic and exponential form of Hatch's (1980) model. Both equations gave significantly biased predictions, which had important implications for the accuracy of diameter

growth projections. However, the two equations differed in their covariates, as the logistic form incorporated DBH, HT and BAPH, while the exponential form incorporated DBH, HT, BAPH, CCF, habitat types, and percentile basal area (Leites *et al.* 2009). As indicated by this analysis, prediction of HCB and calculation of CR is likely more effective than trying to predict CR directly.

In multi-cohort, mixed species stands, like those that predominate in the Acadian Region, age is not very meaningful and SI is often difficult to accurately assess. Hasenauer and Monserud (1996) and Temesgen *et al.* (2005) used additional geographical variables in their CR models; however, R^2 in both models were generally less than 0.6, while the site level

Fig. 3. Predicted height to live crown base (HCB; m) over diameter at breast height (DBH; cm) and total height (HT; m) for some of the major broadleaf and conifer species, namely American beech (AB), red maple (RM), balsam fir (BF) and red spruce (RS) with and without competition. The measures of competitions are crown competition factor (CCF) and basal area larger than subject tree (BAL). The left panel is HCB against DBH and HT for Richards model (DBH only), the middle panel is Richards with competition of CCF = 50 and BAL 20, and the right panel is Richards with competition of CCF = 200 and BAL 40. These are based on generalized nonlinear least square (GNLS) parameters.

contributions to the R^2 were very minimal (1% to 5%). In our analysis, a large number of site-related covariates were added and no significant improvement in model performance was achieved. In addition, many of the parameters associated with the site-related covariates were statistically insignificant ($p > 0.05$). In our study, the overall R^2 varied from 0.52 to 0.79 with competition and site contributions generally about 0.05 and 0.01, respectively. Tree size-related covariates explained

most of the variation (40% to 77%), which is generally higher than in the other studies cited.

Due to the stochastic and autonomous nature of individual branches, HCB can vary between trees for a given tree size and stand condition. Generally, a larger DBH and shorter HT are associated with a lower HCB (or higher CR) and vice versa. DHR describes tree slenderness and is often a good surrogate for local competition. Consequently, HCB should gen-

erally decrease as DHR becomes greater. In general, our parameter estimates matched these expectations. However, there were some species (namely GB, RS, WP, and WS) where the parameter estimates were not consistent with expectations (Table 7). There may be a number of reasons for this, including the availability of data across a range of stand conditions and the strong correlations between DBH, HT, and DHR. Overall, the analysis indicated that HCB was most sensitive to HT among the attributes tested (Fig. 3). With increased competition, HCB will generally increase, which was found for all species (Fig. 3). The sensitivity of HCB to competition was similar for all species.

Conclusion

Regional HCB equations were developed for thirteen tree species of Acadian Region using an extensive data set. A Richards model form using HCB as the dependent variable was found to be superior to the CR prediction approach currently employed in FVS-NE. The final model included both tree-size- and competition-related variables, with parameters estimated using GNLS and NLME. Although NLME significantly improved predictions, GNLS parameters are presented and suggested for use unless new observations are available for local calibration (e.g., Temesgen *et al.* 2008, Meng *et al.* 2009). Overall, model performance was quite good given the geographic range and high variation in the original data.

Acknowledgements

This research was funded by the University of Maine School of Forest Resources, Cooperative Forestry Research Unit (CFRU), Northeastern States Research Cooperative and the US Forest Service Agenda 2020 program. Data were provided by the US Forest Service, CFRU, Maine Forest Service, Acadian Forest Ecosystem Research Program, and the Departments of Natural Resources of New Brunswick, Newfoundland, Nova Scotia, Prince Edward Island, and Québec. Thanks to two anonymous reviewers and Associate Editor Dr. Peter Marshall for their help in improving a previous version of this manuscript.

References

Bechtold, W.A. 2003. Crown-diameter prediction models for 87 species of stand-grown trees in the eastern United States. *Southern Journal of Applied Forestry* 27: 269–278.

Braun, E.L. 1950. *Deciduous forests of eastern North America*. Hafner, New York.

Burkhart, H.E., R.L. Farrar, R.L. Amateis and R.F. Daniels. 1987. Simulation of individual tree growth and stand development in loblolly pine plantations on cutover, site-prepared areas. FWS-1-87. Virginia Polytechnic Institute and State University, School of Forestry and Wildlife Resources, Blacksburg, VA. 47 p.

Crookston, N.L. and G.E. Dixon. 2005. The Forest Vegetation Simulator: A review of its applications, structure, and content. *Computers and Electronics in Agriculture* 49: 60–80.

Curtis, R.O. and D.L. Reukema. 1970. Crown development and site estimates in a Douglas-fir plantation spacing test. *Forest Science* 16: 287–301.

Dixon, G.E., R. White and J. Frank. 2009. Northeast (NE) Variant Overview. USDA Forest Service, Forest Management Service Center, Fort Collins, CO. 44 p.

Dyer, M.E. and H.E. Burkhardt. 1987. Compatible crown ratio and crown height models. *Canadian Journal of Forest Research* 17: 572–574.

Effron, B. and R.J. Tibshirani. 1993. *An introduction to the bootstrap*. Chapman & Hall/CRC, Boca Raton, FL.

Hamilton, G.J. 1969. Dependence of volume increment of individual trees on dominance, crown dimensions, and competition. *Forestry* 42: 133–144.

Hann, D.W., A.S. Hester and C.L. Olson. 1997. *ORGANON User's Manual*. Department of Forest Resources, Oregon State University, OR.

Hasenauer, H. and R.A. Monserud. 1996. A crown ratio model for Austrian forests. *Forest Ecology and Management* 84: 49–60.

Hatch, C.R. 1980. Modeling crown size using inventory data. *Mitt. Forstl. Bundes-Versuchsanst.* 130: 93–97.

Hirsch, R.P. 1991. Validation samples. *Biometrics* 47: 1193–1194.

Holdaway, M.R. 1986. Modeling tree crown ratio. *The Forestry Chronicle* 62: 451–455.

Hopkins, A.D. 1938. *Bioclimatics: A science of life and climate relations*. USDA Miscellaneous publication 280, Washington, DC. 188 p.

Kozak, A. and R. Kozak. 2003. Does cross validation provide additional information in the evaluation of regression models? *Canadian Journal of Forest Research* 33: 976–987.

Krajicek, J.E., K.A. Brinkman and S.F. Gingrich. 1961. Crown competition: a measure of density. *Forest Science*: 35–42.

Kramer, J.K. and T.T. Kozlowski. 1979. *Physiology of Woody Plants*. Academic Press, New York.

Kvålseth, T.O. 1985. Cautionary note about R^2 . *American Statistician* 39: 279–284.

Landsberg, J.J. 2003. Physiology in forest models: history and the future. *Forest Biometry, Modeling and Information Sciences* 1: 49–63.

Larson, B.C. 1963. Stem form development of forest trees. *Forest Science Monograph* 5: 1–42.

Leites, L.P., A.P. Robinson and N.L. Crookston. 2009. Accuracy and equivalence testing of crown ratio models and assessment of their impact on diameter growth and basal area increment predictions of two variants of the Forest Vegetation Simulator. *Canadian Journal of Forest Research* 39: 655–665.

Liu, J.P., H.E. Burkhardt and R.L. Amateis. 1995. Projecting crown measures for loblolly pine trees using a generalized thinning response function. *Forest Science* 41: 43–53.

McRoberts, R.E., J.T. Hahn, G.J. Hefty and J.R. Van Cleve. 1994. Variation in forest inventory field measurements. *Canadian Journal of Forest Research* 24: 1766–1770.

Meng, S.X., S.M. Huang, Y.Q. Yang, G. Trincado and C.L. Vander-Schaaf. 2009. Evaluation of population-averaged and subject-specific approaches for modeling the dominant or codominant height of lodgepole pine trees. *Canadian Journal of Forest Research* 39: 1148–1158.

Pinheiro, J.C. and D.M. Bates. 2009. *Mixed-effects models in S and S-PLUS*. Springer Verlag.

R Development Core Team. 2011. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rijal, B., A.R. Weiskittel and J. Kershaw. (In review). Development of regional height to diameter static equations for fifteen tree species in the North American Acadian Region.

Ritchie, M.W. and D.W. Hann. 1985. Equations for predicting basal area increment in Douglas-fir and grand fir. In: *Oregon State University, Forestry Research Laboratory, Corvallis, OR*.

Ritchie, M.W. and D.W. Hann. 1987. Equations for predicting height to crown base for fourteen tree species in southwest Oregon. Oregon State University, Forestry Research Laboratory, Corvallis, OR.

Russell, M.B. and A.R. Weiskittel. 2011. Maximum and largest crown width equations for 15 species in Maine. *Northern Journal of Applied Forestry* 28: 84–91.

Schwarz, G. 1978. Estimating dimension of a model. *Annals of Statistics* 6: 461–464.

- Short, E.A. and H.E. Burkhart. 1992.** Predicting crown-height increment for thinned and unthinned loblolly-pine plantations. *Forest Science* 38: 594–610.
- Soares, P. and M. Tomé. 2001.** A tree crown ratio prediction equation for eucalypt plantations. *Annals of Forest Sciences* 58: 193–202.
- Soares, P. and M. Tomé. 2002.** Height-diameter equation for first rotation eucalypt plantations in Portugal. *Forest Ecology and Management* 166: 99–109.
- Stage, A.R. 1973.** Prognosis model for stand development. USDA Forest Service, Intermountain Forest and Range Experiment Station, Ogden, UT. Research Paper INT-137. 32 p.
- Stage, A.R. 1976.** An expression for the effect of slope, aspect, and habitat type on tree growth. *Forest Science* 22: 457–460.
- Temesgen, H., V. LeMay and S.J. Mitchell. 2005.** Tree crown ratio models for multi-species and multi-layered stand of southeastern British Columbia. *The Forestry Chronicle* 81: 133–141.
- Temesgen, H., V.J. Monleon and D.W. Hann. 2008.** Analysis and comparison of nonlinear tree height prediction strategies for Douglas-fir forests. *Canadian Journal of Forest Research* 38: 553–565.
- Walters, D.K. and D.W. Hann. 1986.** Taper equations for six conifer species in southwest Oregon. Oregon State University, Forest Research Laboratory, Corvallis, OR. Research Bulletin 56. 36 p.
- Walters, D.K., D.W. Hann and M.A. Clyde. 1985.** Equations and tables predicting gross total stem volumes in cubic feet for six major conifers of southwest Oregon. Oregon State University, Forest Research Laboratory, Corvallis, OR. Research Bulletin 50. 36 p.
- Weiskittel, A.R., N.L. Crookston and P.J. Radtke. 2011a.** Linking climate, gross primary productivity, and site index across forests of the western United States. *Canadian Journal of Forest Research* 41: 1710–1721.
- Weiskittel, A.R., S.M. Garber, G.P. Johnson, D.A. Maguire and R.A. Monserud. 2007.** Annualized diameter and height growth equations for Pacific Northwest plantation-grown Douglas-fir, western hemlock, and red alder. *Forest Ecology and Management* 250: 266–278.
- Weiskittel, A.R., D.W. Hann, J. Kershaw and J.K. Vanclay. 2011b.** Forest growth and yield modeling. John Wiley & Sons, West Sussex, UK. 430 p.
- Weiskittel, A.R., R.G. Wagner and R.S. Seymour. 2010.** Refinement of the Forest Vegetation Simulator, Northeastern Variant growth and yield model: Phase 1. *In* S.R. Meyer (ed.). 2009 Annual Report. University of Maine, Cooperative Forestry Research Unit, Orono, ME. pp. 44–48.
- Wykoff, W.R. 1990.** A basal area increment model for individual conifers in the northern Rocky Mountains. *Forest Science* 136: 1077–1104.
- Wykoff, W.R., N.L. Crookston and A.R. Stage. 1982.** User's guide to the Stand Prognosis Model. USDA Forest Service Intermountain Forest and Range Experiment Station, Ogden, UT. 119 p.
- Zumrawi, A.A. and D.W. Hann. 1989.** Equations for predicting the height to crown base of six species in the Central Western Willamette Valley of Oregon. Oregon State University, Forest Research Laboratory, Corvallis, OR. Research Paper 52. 16 p.