Bootstrap Evaluation of a Young Douglas-Fir Height Growth Model for the Pacific Northwest

Nicholas R. Vaughn, Eric C. Turnblom, and Martin W. Ritchie

Abstract: We evaluated the stability of a complex regression model developed to predict the annual height growth of young Douglas-fir. This model is highly nonlinear and is fit in an iterative manner for annual growth coefficients from data with multiple periodic remeasurement intervals. The traditional methods for such a sensitivity analysis either involve laborious math or rely on prior knowledge of parameter behavior. To achieve our goals, we incorporate a bootstrap approach to obtain estimates of the distribution of predicted height growth for any set of input variables. This allows for a sensitivity analysis with knowledge of the probability of a given outcome. The bootstrap distributions should approximate the variation we might expect from running the model on numerous independent datasets. From the variation in the model parameters, we are able to produce ranges of height growth prediction error falling under a given probability of occurrence. By evaluating these ranges under several combinations of input variables that represent extreme situations, we are able to visualize the stability of the model under each situation. Each of the four components of the model can be investigated separately, which allows us to determine which components of the model might benefit from reformulation. In this case we find that the model is less stable in extremely high site index, especially under low vegetation competition. Other than the computing time involved with the bootstrap, most of the analysis is fairly quick and easy to perform. FOR. SCI. 56(6):592–602.

Keywords: vegetation, density, non-linear model, sensitivity analysis

THE ACCURACY AND PRECISION of predictions from a fitted growth model are usually of primary importance. Unfortunately, it is very easy to overfit a model to one particular dataset (Huang et al. 2003). This situation would probably result in biased predictions when the model is applied to other datasets. A misspecification of the model or a poor draw of training data from the population could also lead to biased predictions. A validation is typically performed to reassure the modeler that the fitted model is free of such problems (Vanclay and Skovsgaard 1997).

Despite the necessity of such a procedure, several issues arise when the model is tested on additional datasets. One issue is that truly independent datasets can be difficult to find. The common alternative, data-splitting, reduces the precision of model parameter estimates (Harrell 2001, p. 90) and may even reduce the accuracy when influential cases are omitted from the modeling dataset. Even if a single independent dataset is found, its analysis can be misleading. Indeed, a single sample cannot provide a very powerful estimator of any quantity. An additional issue, as shown in Yang et al. (2004), is that the results from a single validation dataset are also highly dependent on the particular test used.

Rather than testing the model against one or a limited number of independent datasets, some modern computer-intensive techniques allow one to evaluate model performance using only one complete dataset. The bootstrap estimator for regression models presented in Efron and Tibshirani (1993, p. 113) is one such technique. As presented, it is a tool to obtain a robust estimate of model prediction error as part of model validation. By sampling with replacement from the original dataset, a bootstrap procedure creates b new pseudo-datasets of the same size. The model is then fit to each pseudo-dataset, and the new predictions are collected and analyzed. These pseudo-datasets, under some basic assumptions about how the data were obtained, should be representative of the variation one would see while repeatedly sampling the entire population.

The bootstrap process described above is based on the idea that any model-building dataset is just one sample from a population of datasets obtainable from the entire population. Even if the specified model form is perfectly correct, the parameters estimated from the fitting dataset still represent only one point in a distribution of available parameter estimates. The probability that these parameters match the true population parameters is always exactly zero. Because we cannot hope to obtain the true population parameters, knowledge of the probability of varying degrees of model prediction error would be highly valuable.

Given data from previous studies or an expert opinion on the probable range of model parameters, such an estimate can be obtained with sensitivity analysis (Saltelli et al. 2004, p. 42). Under simple models, such as linear models, this analysis is rather straightforward. Under more complicated scenarios, such as nonlinear models with high parameter

Nicholas R. Vaughn, College of Forest Resources, University of Washington, Box 352100, Seattle, WA 98195—Phone: (206) 616-5785; Fax: (206) 685-0790; nrv2@u.washington.edu. Eric C. Turnblom, College of Forest Resources, University of Washington—ect@u.washington.edu. Martin W. Ritchie, US Forest Service, Pacific Southwest Research Station—mritchie@fs.fed.us.

Acknowledgments: We thank all individuals involved with the RVMM project, and, in particular, Steven Radosevich and Robert Shula for their kind donation of the RVMM dataset. In addition, we owe much gratitude to the SMC and all its contributors. Jim Flewelling, David Marshall, and David Briggs each provided comments that contributed greatly to the completion of this work.

Manuscript received April 19, 2009, accepted July 13, 2010

correlation, the analysis becomes much more difficult (Elston 1992). Typically, when analytical solutions are not tractable, some form of sampling-based analysis is used (Helton et al. 2006).

As an alternative, with the same bootstrap technique used for simple model validation, we can obtain a solid empirical estimate of the parameter distribution from which our model parameters come. We can input this information directly into the model and map variation of parameter estimates to variation in model prediction (Davison and Hinkley 1997, p. 285 and 355). Using this technique, we can ascribe a probability to our model, producing errors of a given level when it is applied to new data. Not only will we have a better understanding of the stability of the model when applied to new data from the population at large, but we will also be able to apportion errors to individual components of the model. This should enable the discovery of components, causing a higher probability of large errors. These components may benefit from a respecification.

Given the utility of such an analysis, it is surprising that there are relatively few examples of the bootstrap technique in the forestry literature. Schreuder et al. (1987) provided a thorough review of literature up to that time using both the bootstrap and the related jackknife (Gregoire 1984) procedures. Although both are focused on sampling, there are several cited examples of success and failure with each techniques. More recently, the bootstrap has been used to accomplish a wider variety of tasks. Some of these tasks are to estimate model parameters (Fang et al. 2001), to estimate bias and variance of stand density metrics (Ducey and Larson 1999), to model snag abundance (Fan et al. 2004), to optimize stand management Liang et al. (2006), and to characterize the genetic variation of an eastern conifer species (Potter et al. 2008).

However, no examples of bootstrap use in a sensitivity analysis could be found in the forestry literature. One has to depart from the field of forestry to find examples of similar usage of the bootstrap. Mennemeyer and Cyr (1997) used a bootstrap to develop confidence intervals for the expected values in a medical treatment decision tree. Here the bootstrap was really the only tool capable of the task. Pasta et al. (1999) also dealt with medical decisionmaking. They used bootstrap methods to develop distributions of some model parameters for a more traditional sampling-based sensitivity analysis.

In this article, we wished to demonstrate that the bootstrap procedure has a place in forest modeling beyond simple model validation. We thoroughly evaluated the predictive performance of a growth model with a bootstrap-based sensitivity analysis. For this purpose, we present an annual height growth model for young stands of Douglas-fir in the Pacific Northwest. This height growth equation incorporates functions to estimate the multiplicative effects of tree position, stand density, and competing shrub vegetation cover. This model is highly nonlinear. In addition, an iterative procedure is used to produce parameters for an annual model from data with varying growth periods. These two attributes combine to make it difficult to analyze the model using traditional methods. Using the bootstrap-derived parameter distribution, we are able to assess sensitivity of model output to the variation one might expect to occur between data sets drawn from the population at large. This analysis was done to identify potential weak points in the model, those parts that perform wildly under extreme stand conditions.

Methods

This modeling effort was undertaken to build a growth model that would expand on the applicable range of the CONIFERS (Ritchie and Hamann 2008) growth modeling project. The existing CONIFERS model covered many coniferous species growing in the region of northern California and southern Oregon. Two large database projects with tree-level data covering young Douglas-fir in much of the Pacific Northwest were consolidated into one large modeling dataset. The resulting model has been embedded as an option under the same CONIFERS user interface. The model described is applicable only to stands of pure Douglas-fir less than 20 years old, although age itself is not a predictor in the model.

Data

This section summarizes the information provided by Vaughn (2007). Modeling data were provided by the Stand Management Cooperative (SMC) and the Regional Vegetation Management Model (RVMM) project. The SMC is a consortium of landowners in the Pacific Northwest established in 1985 to pool resources to provide high-quality information on the long-term effects of silvicultural treatments (Chappell et al. 1988). The SMC data used for this project are a product of a planting density trial (the "Type III" installations). There are 34 such installations with 3–23 plots, each from 0.04–0.20 ha in size (Silviculture Project TAC 1991). The RVMM project began in the early 1990s for the purpose of developing a growth model to weigh various vegetation management options. The RVMM dataset contains data from 98 research plots, each 0.04 ha in size, in each of the Coastal and Cascade Mountains (196 plots total) (Shula 1998). The SMC installations and the RVMM plots range from southern Oregon to the Canadian border, with one installation in British Columbia (Figure 1). At each measurement, crews recorded the total height and diameter of each tree in the plot. On both SMC and RVMM plots, crews measured a subsample of trees for crown width and height to crown base. Remeasurement interval was 2 years in the RVMM plots and 2 to 4 years in the SMC plots. In total, more than 20,000 Douglas-fir tree growth records from 388 plot measurements made up the model-fitting data.

To sample the vegetation, on both the SMC and RVMM projects, crews located four 0.0247-ha subplots within each tree plot. Each subplot was split into four equal-sized quadrants and crews took ocular estimates of cover for each species in the quadrant. We averaged these 16 quadrant samples to get a mean plot level value of each species. To create a single value estimate for plot vegetation cover, we summed cover estimates for all species classified as a fern or as a shrub. The unit of this number is percent cover, but

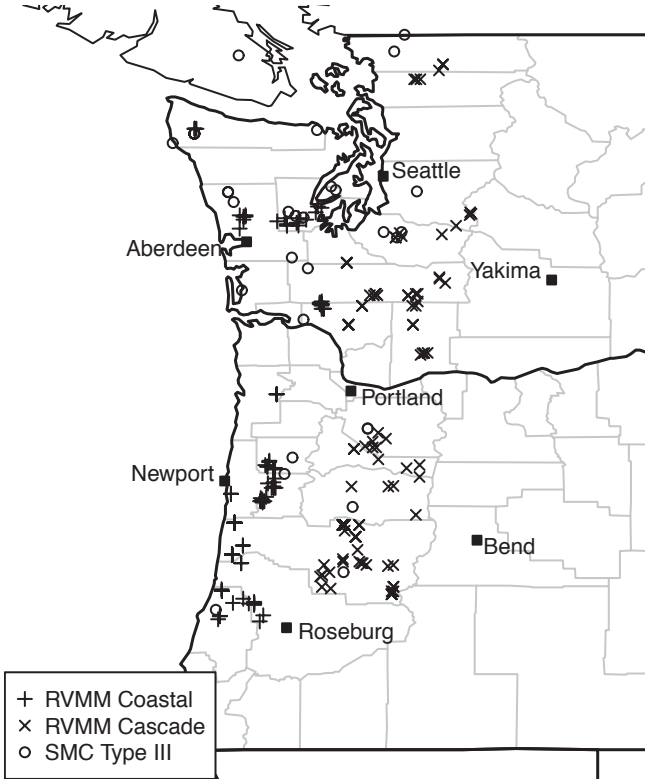

Figure 1. Locations of the study tree plots within the Pacific Northwest. Plots from the two datasets of the RVMM project: +, Coastal; ×, Cascade; ○, SMC installations, which contain multiple tree plots.

it ranges from 0 to 212 because in many cases vegetation layers overlapped within a plot.

Height Growth Function

A single function of two predictor variables, initial height and site index, explained a large share of the total variation in height growth. Site index was computed from the data with the curves of Flewelling et al. (2001), which were designed for use with younger plantations of Douglas-fir. This nonlinear function, shown as Equation 1, is referred to as the base function.

$$\widehat{\Delta H}_{ij} = b(H_{(0)ij}, S_i) \tag{955}$$

$$= \frac{1}{b_1 + b_2 H_{(0)ij}^{-1} S_i^{-1.5} + b_3 S_i^{-0.5} + b_4 H_{(0)ij}^{0.5}},$$

where $\widehat{\Delta H}_{ij}$ is the 1-year change in total height in meters of tree j in plot i, $H_{(0)ij}$ is the initial total height in meters of tree j in plot i, S_i is the site index (meters at base age 30 years total) of plot i, and b_1 through b_4 are parameters estimated by the model-fitting procedure.

It is common for height growth models to be composed of a potential growth function and a modifier function (Hann and Ritchie 1988). In a similar fashion, the growth estimates from this base function are modified by a series of three multiplier functions, shown in Equations 2 to 4. All three of these modifier functions are affected by stand top height ($H_{(\text{top})i}$) or the mean height of the "largest" 99 trees per hectare in the given plot. On plots with sampled trees shorter than breast height, we used total height as the standard to select the largest trees. On plots with all sampled trees above breast height (1.3 m), we defined largest trees as those with the greatest dbh.

$$r(H_{(0)ij}, H_{(\text{top})i}) = \exp(r_1 * \exp(H_{(\text{top})i}^{r_2}) * \log(H_{(0)ij}/H_{(\text{top})i})), \tag{956}$$

$$v(H_{(\text{top})i}, V_i) = \exp(-\exp(v_1 + v_2 * (H_{(\text{top})i}))\sqrt{V_i}), \tag{957}$$

$$d(H_{(\text{top})i}, T_i) = 1 + d_1 * (H_{(\text{top})i} - d_2) * (T_i - 741.32), \tag{958}$$

where $r(H_{(0)ij}, H_{(\text{top})i})$ is the relative height modifier function, $v(H_{(\text{top})i}, V_i)$ is the vegetation cover modifier function, $d(H_{(\text{top})i}, T_i)$ is the density modifier function, $H_{(0)ij}$ is the initial total height in meters of tree j in plot i, $H_{(\text{top})i}$ is the top height in meters of plot i, $H_{(0)ij}/H_{(\text{top})i}$ is the relative height of tree j in plot i, V_i is the mean vegetation (shrub and fern) cover in plot i, T_i is the stems per hectare in plot i, and $r_{1,2}$, $v_{1,2}$, and $d_{1,2}$ are model parameters.

The relative height modifier function in Equation 2, with a positive value for r_1, gives the taller trees in a stand a faster growth rate and the smaller trees a slower growth rate. The amount of increase or reduction depends on the current top height of a given stand. Younger stands do not see as much intertree competition as do older stands, and the modifier function reflects this observation.

The vegetation modifier function in Equation 3 also changes as the stand gets taller. In this case the amount of vegetation in a stand will have decreasing impact on the individual tree growth rates as the stand climbs out of the influence of the vegetation competition. The value of v_2 in this function dictates the rate at which vegetation effects diminish with increasing top height.

Equation 4 adjusts growth rates for the number of living trees per hectare in the stand. Previous work has shown that higher densities actually increase the early height growth of a Douglas-fir stand (Scott et al. 1998, Turnblom 1998, Woodruff et al. 2002). This phenomenon was observed in the datasets used to build this model, and we therefore attempted to incorporate the shifting density effect into this modifier function. The initial increase in height growth at higher densities diminishes with increasing top height until reaching the "crossover" point. At this point, the effects of stand density revert to the more well-known pattern of reduced height growth at higher densities. The location of this crossover point in top height scale is given by the parameter d_2.

Because the data came from plots measured on differing interval lengths, we used the iterative procedure, described by McDill and Amateis (1993), to obtain annualized estimates of the model parameters. In the first iteration, the response variable is the total period height growth divided by the number of years in the growing period (a naive estimate of 1-year growth. In subsequent iterations, the

Table 1. Descriptive statistics of the predictor values from the modeling dataset. All variables are plot-level statistics except for initial height, which is a tree-level statistic.

Predictor	Minimum	25%	50%	Mean	75%	Maximum
Site index (m)	8.4	21.6	25.0	24.2	27.3	31.8
Density (trees/ha)	166.4	793.4	1376.0	1657.0	2265.0	8000.0
Vegetation cover (%)	0.0	27.4	48.6	55.1	74.4	212.4
Top height (m)	0.7	3.2	5.4	6.0	7.8	14.4
Initial height (m)	0.1	2.2	3.7	4.2	5.8	16.2

current parameter estimates are used to compute the proportion of the predicted period growth attributable to the first year of the period. This proportion is then multiplied by the observed period growth to create the new response variable, an improved estimate of 1-year growth). Iteration is allowed to continue until no parameter estimates change in any of the first five significant digits. We used a logistic mortality model to reduce the stems per hectare and basal area per hectare represented by each tree in the tree list. In addition, we used a linear vegetation cover change model to predict yearly change in vegetation at each time step during the iterations. Parameters at each time step were estimated using the nls function in the R statistical program (R Development Core Team 2009).

Simple descriptive statistics of the five model predictor variables contained in the modeling dataset are shown in Table 1. The first four predictors are measured at the plot level. The data cover a moderate range of each predictor, but lower values of S_i, T_i, and $H_{(\text{top})i}$ are not as well represented. Of particular importance to potential users of the model is the range of initial tree height and stand top height from which the model is fit.

Bootstrap-Assisted Sensitivity Analysis

The simple bootstrap procedure consists of fitting the model to multiple pseudo-datasets. Each of these pseudo-datasets is created by resampling, with replacement, from the original dataset. Because of the grouped structure of the dataset, resampling was done at the plot level. Each pseudo-dataset contained the same number of plots from the SMC dataset and the Coastal and Cascade partitions of the RVMM dataset, as the original data. If a plot was chosen, then the data from all trees at every measurement of the plot were included in the pseudo-dataset. Therefore, data from plots selected n times in a given resample would be duplicated n times in the current pseudo-dataset. Because the number of trees in a plot varied, this selection would result in datasets of differing numbers of trees but with an equal number of plots.

We refit the model to each of 2,048 pseudo-datasets and recorded the parameter estimates (β_i^*) for each. The number of bootstrap runs, 2,048, represents a compromise between computing time limits and large sample size requirements. This procedure should produce a sparse approximation, because of the high dimensionality of the true joint distribution of the 10 model parameters. However, when used, this parameter distribution is transformed to a single dimension, eliminating any problems with sparsity.

It is important to note that resampling techniques such as this make extensive reuse of the same data. As such, these methods are dependent on a representative sample of data from the population of interest. It is ideal that a dataset contain information about all growing conditions occurring in the area of application. The data came from plots that span most of western Washington and Oregon. Although undoubtedly not every environment in the region is represented, this dataset purposely encompasses a wide range of growing conditions over a period of 13 years (1991–2004).

Table 2. Thirty-three sets of predictor values selected for use in the analysis of model sensitivity

Set	1	2	3	4	5	6	7	8	9	10	11
Site index (m)	18	32	18	32	18	32	18	32	18	32	18
Density (trees/ha)	250	250	1,250	1,250	250	250	1,250	1,250	250	250	1,250
Top height (m)	2.0	2.0	2.0	2.0	13.0	13.0	13.0	13.0	2.0	2.0	2.0
Initial height (m)	1.2	1.2	1.2	1.2	7.8	7.8	7.8	7.8	2.2	2.2	2.2
Vegetation cover (%)	10	10	10	10	10	10	10	10	10	10	10
	12	13	14	15	16	17	18	19	20	21	22
Site index (m)	32	18	32	18	32	18	32	18	32	18	32
Density (trees/ha)	1,250	250	250	1,250	1,250	250	250	1,250	1,250	250	250
Top height (m)	2.0	13.0	13.0	13.0	13.0	2.0	2.0	2.0	2.0	13.0	13.0
Initial height (m)	2.2	14.3	14.3	14.3	14.3	1.2	1.2	1.2	1.2	7.8	7.8
Vegetation cover (%)	10	10	10	10	10	90	90	90	90	90	90
	23	24	25	26	27	328	29	30	31	32	33
Site index (m)	18	32	18	32	18	32	18	32	18	32	25
Density (trees/ha)	1,250	1,250	250	250	1,250	1,250	250	250	1,250	1,250	1,376
Top height (m)	13.0	13.0	2.0	2.0	2.0	2.0	13.0	13.0	13.0	13.0	5.4
Initial height (m)	7.8	7.8	2.2	2.2	2.2	2.2	14.3	14.3	14.3	14.3	4.3
Vegetation cover (%)	90	90	90	90	90	90	90	90	90	90	49

Table 3. Parameters from the full height growth model in equations 1 through 4 with the estimated and bootstrap SEs

Coefficient	Estimate	Estimated SE	Bootstrap SE
b_1	−5.77e−01	4.20e−02	2.06e−01
b_2	7.95e+01	5.23e+00	2.28e+01
b_3	4.75e+00	2.31e−01	1.52e+00
b_4	1.18e−01	6.53e−03	4.55e−02
v_1	−1.82e+00	3.66e−02	2.36e−01
v_2	−1.81e−01	6.70e−03	3.79e−02
d_1	−1.91e−05	8.84e−07	6.12e−06
d_2	7.51e+00	1.30e−01	1.05e+00
r_1	5.70e−02	5.39e−03	1.69e−02
r_2	2.27e−01	1.67e−02	4.52e−02

The stability of the model is dependent on the values of the five predictor variables. In an attempt to cover as much of the range of each predictor as possible, we chose values to represent low and high values (10th and 90th percentile) of each predictor. Rather than being based on percentiles, the low and high values for initial height were taken to be 60 and 120%, respectively, of the top height value in each set. We selected 32 predictor value sets, representing every one of the 32 possible combinations of high and low values for the five predictors. In addition, we created one set of predictor variables representing the median of each variable, making a total of 33 sets (Table 2). The 2,048 bootstrap parameter sets (β^*) were then applied to the model to produce a distribution of bootstrap predictions (ΔH_{ij}) for each of 33 sets of predictor values. Investigation of these distributions should give insight into stand conditions for which the model is stable and to which situations the model should not be applied.

Results

The estimated parameter values for the height growth model as fit to the original dataset are shown in Table 3. The second column of Table 3 shows the estimated standard errors of each model parameter as reported by the R function nls. The model fit with a root mean square error of 0.2405 m. However, this value is only approximate because the "actual" 1-year growth is not measured but predicted as well from the proportion of full-period growth expected to occur in the first year. The residuals were found to increase with tree height, but we determined that the magnitude of difference did not warrant the further complexity associated with any data transformation.

Visualizations of the four components of the growth model are displayed in Figure 2 as fit to the entire dataset. The first panel is a contour plot of height growth prediction as determined by site index and initial height. For a given site index, one can see that predicted height growth follows a well-known pattern; it increases sharply with initial height

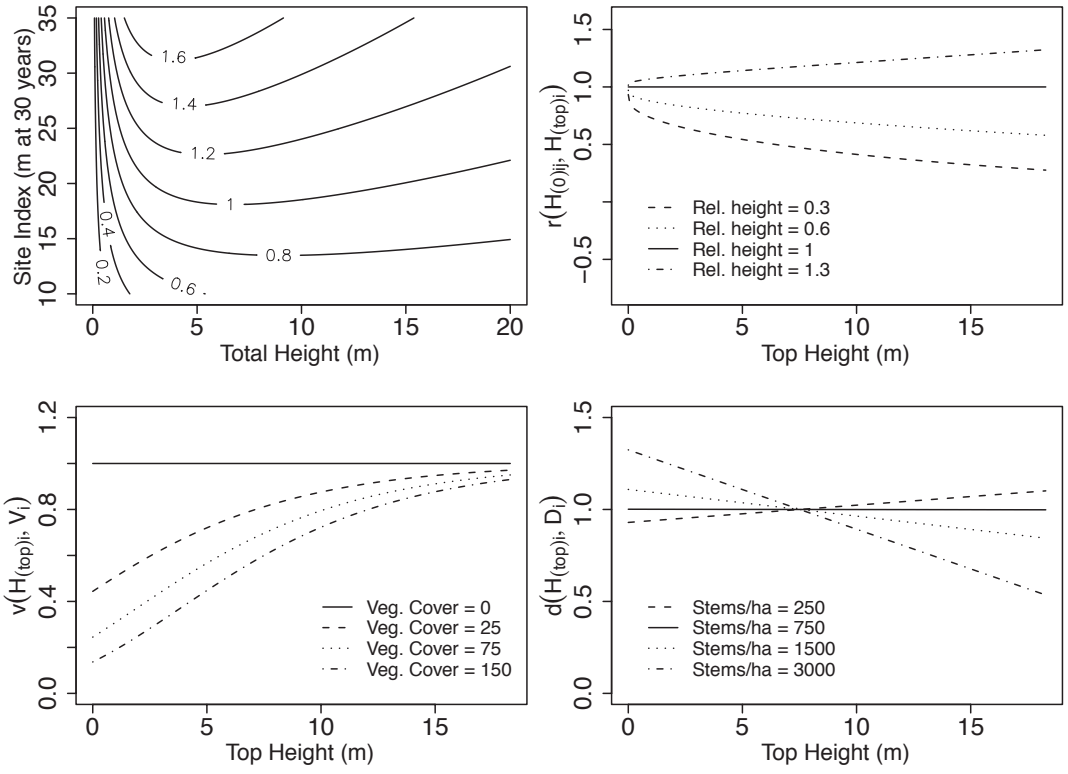

Figure 2. Values of the four components of the height growth model. The top left panel contains a contour plot of the base function height increment for a range of initial height and site index values. The top right panel depicts the relative height modifier function (Equation 2) for several values of relative height (Rel. height). The bottom left panel depicts the vegetation modifier function (Equation 3) for several values of vegetation cover (Veg. Cover). The bottom right panel depicts the density modifier function (Equation 4) for several values of stand density (stems/ha). The first component predicts height growth, whereas the last three components produce values that increase/reduce this height growth through multiplication.

before peaking and decreasing with further increases of initial height. The next three panels show the multiplicative effects of the three modifier functions as top height increases. As expected, there is a strong influence of top height on the effects of vegetation cover (Knowe et al. 1992), stand density (Scott et al. 1998), and relative height (Ritchie and Hann 1986).

The bootstrap procedure took about 23 hours to perform on a fairly modern desktop computer, while splitting the runs across the processor's four cores. As the bootstrap procedure is inherently parallel, access to a cluster computing environment would drastically reduce the time involved. Of the 2,048 bootstrap runs, 23 needed manual adjustment of the parameter starting values to converge. The number of annualizing iterations required for convergence ranged from 2 to 254, with the median equal to 27. Ninety percent of the runs took between 20 and 39 iterations.

The third column of Table 3 contains estimates of the SD of the parameters, as produced by the bootstrap procedure. Figure 3 shows histograms of the bootstrap estimates of all 10 parameters in the model. A vertical dotted line is placed at 0 for reference, and the approximate density functions derived from normal theory are scaled to fit and overlaid. These histograms make clear the extent of difference between the theoretical standard errors and the bootstrap estimates when assumptions are not met. In Figure 4, the nature of the relationships among the parameters is shown with a scatterplot matrix of the bootstrap estimates. The

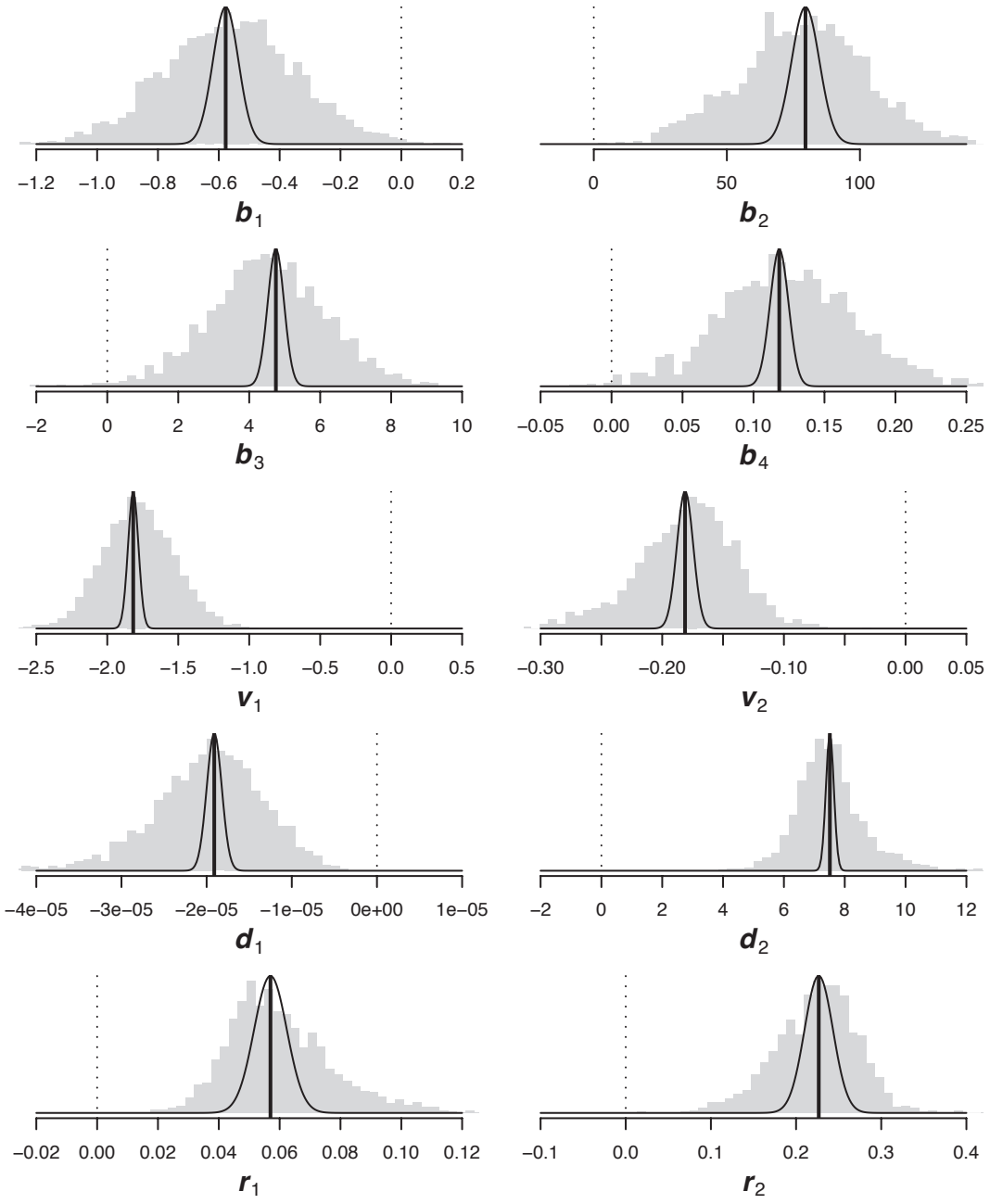

Figure 3. Bootstrap distributions of the full-height growth model (Equations 1–4) parameters are shown in gray. The scaled theoretical distributions appear as solid black lines. Zero on the horizontal axis is indicated with a dotted line.

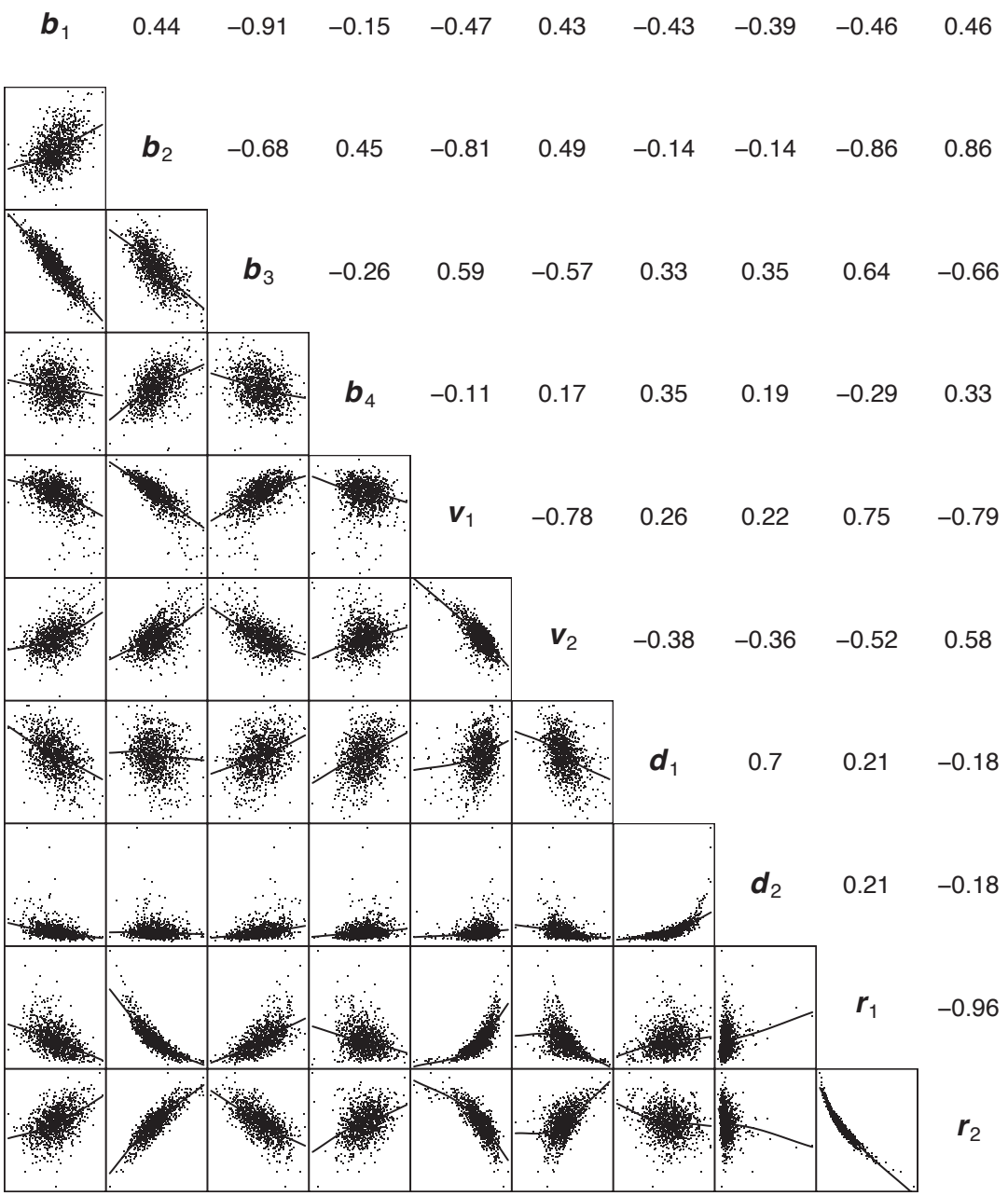

Figure 4. A scatterplot matrix of the 2,048 bootstrap coefficient estimates. The upper portion contains the linear correlation coefficients between the model parameter estimates.

entries in the upper triangle of this figure are the linear correlation coefficients. As is often the case with nonlinear models, several parameters are highly related in curvilinear fashion.

To visualize the effect of sampling variation on the model function, Figure 5 shows approximate 95% confidence bands around the response curve of each component. These bands were created using the 0.025 and 0.975 quantiles of the prediction line at each of a sequence of points along the horizontal scale. This result illustrates the drastic effect on the base model of bootstrap dataset variation, especially for initial heights less than 5 m. In addition, the shape of the vegetation prediction surface varied heavily. The changes in the other two components were not as drastic.

Although the model surface did shift moderately across bootstrap data sets, more important to prediction performance is how these shifts interact to affect the model predictions. Shifts in one model component were probably countered by opposing shifts in the other components. Figure 6 shows a box and whisker plot for the model predictions at each of the predictor value sets listed in Table 2. The vertical axis represents the deviance of the predictions (each ΔH_{ij}) by the model under the bootstrap parameter sets from the equivalent prediction by the original model (ΔH_{ij}). The boxes represent the 25th, 50th, and 75th percentile within

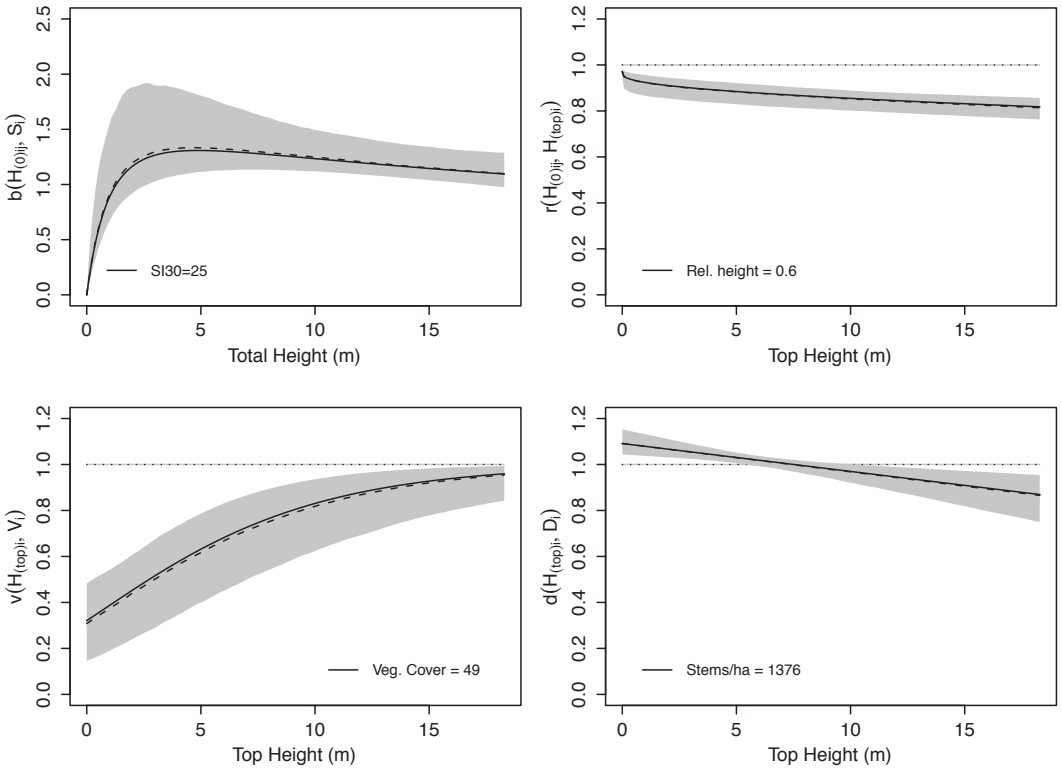

Figure 5. The 95% quantile regions of the four components of the height growth model. The top left panel contains a plot of predicted height growth by initial height at a constant value of site index. The solid line is the original model fit, and the dashed line is the median line over the 2,048 bootstrap coefficient sets. The gray area represents the zone containing the most central 95% of the prediction curves form the bootstrap fits. The top right panel depicts the relative height modifier function for a single value of relative height (Rel. height). The bottom left panel depicts the vegetation modifier function for a single value of vegetation cover (Veg. Cover). The bottom right panel depicts the density modifier function for a single value of stand density (stems/ha).

each parameter set. Whiskers represent the 10th and 90th percentile. Circles represent the 5th and 95th percentile, and × represent the 1st and 99th percentiles. To reiterate, these boxes represent the distribution of predicted mean growth for a single tree in a particular stand, not variation of growth within a group of trees. Applications in which the model is stable should be visible on this figure as those predictor sets with small quantile ranges. The bars at the bottom indicate whether the values of the individual predictor variables are high (open bars) or low (black bars). Of particular importance is the stability of the model under the last predictor set, which best represents the intended zone of application for the model.

We also looked at the fluctuations in the model component functions for individual predictor value sets. An example of this is shown for set number 30 in Figure 7. Here the four boxes represent fluctuations during the bootstrap of the values output by each of the four components of the model: the base function (equation, the relative height growth modifier, the vegetation modifier, and the density modifier [Equations 1, 2, 3, and 4, respectively]). The final three boxes represent modifier functions, for which fluctuations will have an equal effect. A value less than 1 is a reduction in predicted growth and a value greater than 1 is an increase in predicted growth. It is clear from this figure that the vegetation modifier has the highest variation for this particular predictor set.

Discussion

Unlike repeated sampling of the population, the bootstrap resamples a relatively small subset of the population. This heavy reuse of the same data introduces two new potential problems when distribution percentiles are estimated. First, despite the continuous nature of the predictor variables, the sample data are discrete in nature. This discreteness leaks its way into the resulting distribution estimates. Second, each bootstrap dataset contains numerous duplicates, which should result in reduced efficiency for point estimates. Both of these problems become less relevant as the number of original samples increases. In addition, increasing the number of bootstrap runs alleviates the second problem (Efron and Tibshirani 1993, p. 271). We found partial evidence of discreteness in our results, depicted as drastic changes in bar height in the histograms in Figure 3. These effects are minute and will probably have little effect on interval estimates. In addition, 2,048 is a rather large number of bootstrap samples and should result in precise enough point estimates.

Although there were differences between the parameter

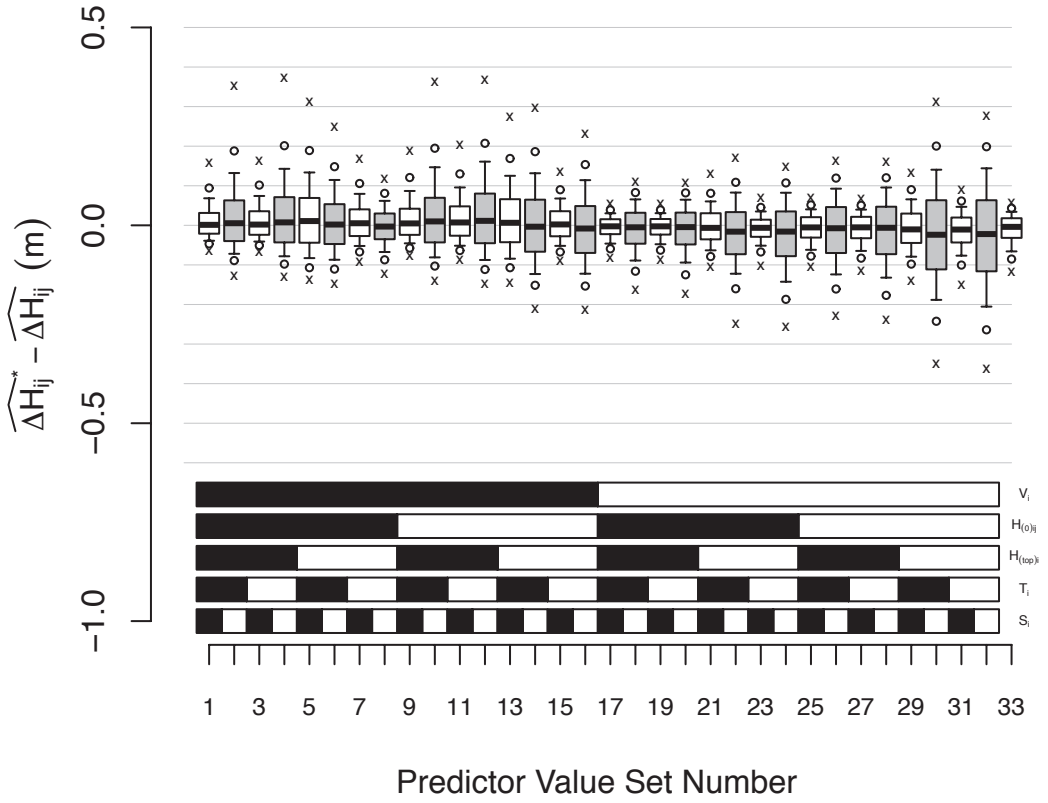

Figure 6. Boxplots representing the distributions of predicted height growth produced by running the model under each of the bootstrapped coefficient sets. Each boxplot is associated with a predictor value set, as listed in Table 2, and is centered around the prediction by the original model. Boxes represent the 25th, 50th, and 75th percentile. Whiskers represent the 10th and 90th percentile. ○, 5th and 95th percentile. ×, 1st or 99th percentile value. ■, low values of the given predictor; □, high values of the given predictor.

standard errors estimated by the bootstrap and those estimated by analytical approximations (Table 3), nothing unexpected occurred. For each parameter the bootstrap distribution was significantly more dispersed. The degree of difference could easily be explained by our neglect of the grouping structure of the data in the original modeling process. This structure was deliberately omitted for two reasons. First, accounting for this structure in a nonlinear model is not a straightforward process, because it is unclear which parameters should carry any random effects. Second, the increase in computation time in switching to a nonlinear mixed-effect model was prohibitive.

The empirical bootstrap distribution of the model parameters, shown in Figure 3, provides a clear picture of the variability of the model parameters. One parameter distribution was of particular interest. The d_2 parameter determines at which top height the "crossover effect" occurs, as described in Methods. The estimate of 7.51 m for parameter d_2 seems high compared with some previous results. Woodruff et al. (2002) found that the crossover point was about age 7. Indeed, it seems unlikely that a plantation would reach a top height of 7.51 m in only 7 years. However, looking at the bootstrap distribution of d_2 in Figure 3, we see that values as low as 5 m occurred with some frequency. Because a plantation could conceivably reach 5 m in 7 years, the two results do not seem incompatible.

Also of note is the finding that, on four occasions, the b_4 parameter was estimated to be less than zero. When this happens, negative values are output by the base function, which in turn produces negative height growth estimates. The number of such values is minimal and is little cause for concern. However, much work went into producing a model that does not predict negative height growth, and it is very interesting that some pseudo-datasets produced a model that was still able to predict such growth. We kept track of only the pseudo-datasets that initially failed to result in convergence, but keeping track of all pseudo-datasets would have enabled us to further investigate this phenomenon.

Because datasets can vary so widely and model parameter estimates can likewise vary, we sought a way to thoroughly evaluate the predictive abilities of the height growth model. When applied to various datasets, we desired to know how inaccurate the model predictions could be. A simple model validation could provide a single number answer to the average expected error but not the limits of probable individual errors. A traditional sensitivity analysis might have been able to meet this goal. However, use of this approach would have introduced several complications. For the traditional approach to be accurate, one needs accurate prior knowledge of model parameter behavior. This information, if it exists, is typically for individual parameters. These individual distributions are combined for the analysis (Helton et al. 2006) but will contain no information about parameter correlation. As shown in Figure 4, there is a large association between model parameter estimates. With no understanding of the joint probability distribution of the

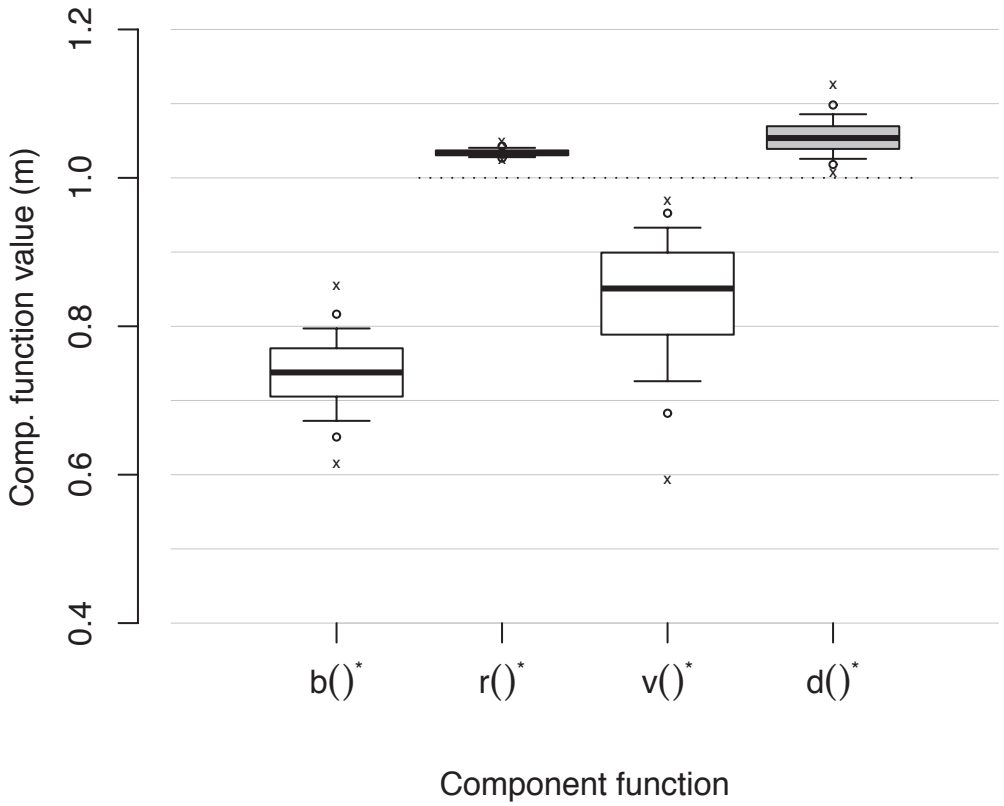

Figure 7. Boxplots representing the distributions of component functions (Equations 2 through 4) produced by running the model under each of the bootstrapped coefficient sets. These boxes all correspond to predictor set 30 as listed in Table 2. Boxes represent the 25th, 50th, and 75th percentile, and whiskers represent the 10th and 90th percentile. ○, 5th and 95th percentile; ×, 1st or 99th percentile value.

parameters, there can be no hope of understanding the effects of combined parameter shifts on model output. Another problem of more traditional sensitivity analysis is deciding how exactly to sample from the parameter space. Much research has been performed to find the efficient methods of doing so (Helton et al. 2005). One notable feature of the bootstrap approach to this problem is that we were able to simultaneously estimate the distribution of the model parameters and sample probabilistically from this multivariate distribution. Thus, we may skip the parameter distribution and sampling problems altogether.

By directly mapping the empirical bootstrap distribution to a distribution of model output values, we can now get estimates of distribution quantiles and thus we can now assign probability values to specific ranges in model prediction error. As a visual example, Figure 6 shows boxplot representations of the height growth prediction distributions. From reading this figure, we should expect that for trees similar to set 6, predictions from our model have a 50% chance of falling within 0.05 m on either side of the true average growth, assuming that the correct model form has been used. The elongated boxes represent the tree and stand conditions for which the particular model building dataset has a large effect on model outcome. These are regions of the predictor space for which the model is not as stable, and therefore the predictions on new data are more likely to be imprecise. To some degree, this variation occurs for all models at the upper and lower extents of the predictor variable ranges because data points in such regions always have more influence.

Another factor affecting the box length in Figure 6 is the increase in height growth variation associated with larger predicted heights. Regardless of this fact, some patterns are still very evident in the figure. It is clear that model performance suffers under high values of site index, especially when the vegetation is in the low range. The cause of this phenomenon could be the lack of data in this region of the predictor variable space. It is easy to imagine that some of the resampled pseudo-datasets may not contain any data representing of this region. However, this effect is probably exacerbated by a misspecification of the model.

To investigate further the causes of the patterned variation evident in Figure 6, we can look at how much of a role each of the model components played in this variation. For a fixed set of input, the output of each model component can be considered a statistic of interest. The bootstrap distributions of the outputs of each model component are displayed as boxplots in Figure 5. The predictor set with the widest distribution of prediction, set 30, was chosen for this analysis. The base function is the main component of the model and can be seen as a prediction for a dominant tree in a stand with no vegetation planted at a density of about 1500 trees/ha. The boxplot of the distribution of base function output shows some variation around a mean of about 0.75 m. This base function prediction is then multiplied by the output of the three modifiers, which all have the same power to affect the model. The box for the vegetation modifier output is much larger than those of the other two modifiers and is clearly a main reason for the variation seen in the height growth predictions for set 30.

When run under alternative datasets, a model almost always performs poorer than on the training data. To understand the variation of predictions from our model on such alternative datasets, we used a bootstrap technique to build empirical distributions of model predictions under several predictor value sets. These distributions not only allowed us to assess likely model errors in operational use but also gave us the opportunity to analyze individual components of the model. We discovered that under regions of predictor space that are represented well in the training data, the model is fairly stable. In extreme situations, the vegetation modifier component led to higher variation in prediction. The bootstrap technique made this entire sensitivity analysis rather simple to perform. Given its versatility, the bootstrap will probably show up more often in the future within the forestry literature.

Literature Cited

CHAPPELL, H.N., R.O. CURTIS, D.M. HYINK, AND D.A. MAGUIRE. 1988. The Pacific Northwest Stand Management Cooperative and its field installation design. P. 1073–1080 in *Forest growth modelling and prediction,* Vol. 2. GTR NC-120. US For. Serv., North Central For. Exp. Stn., Minneapolis, MN.

DAVISON, A.C. AND D.V. HINKLEY. 1997. *Bootstrap methods and their application.* Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge Univ Press, Cambridge, UK.

DUCEY, M.J., AND B.C. LARSON. 1999. Accounting for bias and uncertainty in nonlinear stand density indices. *For. Sci*. 45: 452–457.

EFRON, B., AND R.J. TIBSHIRANI. 1993. *An introduction to the bootstrap,* 1st ed. Monographs on Statistics and Applied Probability. Chapman & Hall, New York, NY.

ELSTON, D.A. 1992. Sensitivity analysis in the presence of correlated parameter estimates. *Ecol. Model*. 64(1):11–22.

FAN, Z., S.S. LEE, S.R. SHIFLEY, F.R. THOMPSON, III, AND D.R. LARSEN. 2004. Simulating the effect of landscape size and age structure on cavity tree density using a resampling technique. *For. Sci*. 50:603–609.

FANG, S., G. GERTNER, AND D. PRICE. 2001. Uncertainly analyses of a process model when vague parameters are estimated with entropy and bayesian methods. *J. For. Res*. 6(1):13–19.

FLEWELLING, J.W., R. COLLIER, R. GONYEA, D.D. MARSHALL, AND E.C. TURNBLOM. 2001. *Height–age curves for planted stands of Douglas-fir with adjustments for density*. SMC Working Paper 1. Stand Management Cooperative, Seattle, WA.

GREGOIRE, T.G. 1984. The jackknife: An introduction with applications in forestry data analysis. *Can. J. For. Res*. 14(4):493–497.

HANN, D.W., AND M.W. RITCHIE. 1988. Height growth rate of Douglas-fir: A comparison of model forms. *For. Sci*. 34(1): 165–175.

HARRELL, F.E. 2001. *Regression modeling strategies: With applications to linear models, logistic regression, and survival analysis.* Springer Series in Statistics. Springer Verlag, New York.

HELTON, J., F. DAVIS, AND J. JOHNSON. 2005. A comparison of uncertainty and sensitivity analysis results obtained with random and Latin hypercube sampling. *Reliability Eng. Syst. Saf*. 89(3):305–330.

HELTON, J.C., J.D. JOHNSON, C.J. SALLABERRY, AND C.B. STORLIE. 2006. Survey of sampling-based methods for uncertainty and sensitivity analysis. *Reliability Eng. Syst. Saf*. 91(10–11): 1175–1209.

HUANG, S., Y. YANG, AND Y. WANG. 2003. A critical look at procedures for validating growth and yield models. P. 271–291 in *Modelling forest systems,* Amaro, A., D. Reed, and P. Soares (eds.). CABI, Wallingford, Oxfordshire, UK.

KNOWE, S.A., T.B. HARRINGTON, AND R.G. SHULA. 1992. Incorporating the effects of interspecific competition and vegetation management treatments in diameter distribution models for Douglas-fir saplings. *Can. J. For. Res*. 22(9):1255–1262.

LIANG, J., J. BUONGIORNO, AND R.A. MONSERUD. 2006. Bootstrap simulation and response surface optimization of management regimes for Douglas-fir/western hemlock stands. *For. Sci*. 52: 579–594.

MCDILL, M.E., AND R.L. AMATEIS. 1993. Fitting discrete-time dynamic models having any time interval. *For. Sci*. 39(3): 499–519.

MENNEMEYER, S.T., AND L.P. CYR. 1997. A bootstrap approach to medical decision analysis. *J. Health Econ*. 16(6):741–747.

PASTA, D.J., J.L. TAYLOR, AND J.M. HENNING. 1999. Probabilistic sensitivity analysis incorporating the bootstrap: An example comparing treatments for the eradication of *Helicobacter pylori*. *Med. Decis. Making* 19(3):353–363.

POTTER, K., W. DVORAK, B. CRANE, V. HIPKINS, R. JETTON, W. WHITTIER, AND R. RHEA. 2008. Allozyme variation and recent evolutionary history of eastern hemlock (*Tsuga canadensis*) in the southeastern United States. *New For*. 35(2):131–145.

R DEVELOPMENT CORE TEAM. (2009). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria.

RITCHIE, M.W., AND J.D. HAMANN. 2008. Individual-tree height-, diameter- and crown-width increment equations for young Douglas-fir plantations. *New For*. 35(2):173–186.

RITCHIE, M.W., AND D.W. HANN. 1986. Development of a tree height growth model for Douglas-fir. *For. Ecol. Manag*. 15(2):135–145.

SALTELLI, A., S. TARANTOLA, AND F. CAMPOLONGO. 2004. *Sensitivity analysis in practice: A guide to assessing scientific models.* John Wiley & Sons, Hoboken, NJ.

SCHREUDER, H.T., H.G. LI, AND C.T. SCOTT. 1987. Jackknife and bootstrap estimation for sampling with partial replacement. *For. Sci*. 33(3):676–689.

SCOTT, W., R. MEADE, R.M. LEON, D.M. HYINK, AND R.E. MILLER. 1998. Planting density and tree-size relations in coast Douglas-fir. *Can. J. For. Res*. 28:74–78.

SHULA, R.G. 1998. *Database development and application to characterize juvenile Douglas-fir and understory vegetation in the Oregon and Washington Coast Range Mountains.* MSc thesis, Oregon State University, Corvallis, OR.

SILVICULTURE PROJECT TAC. 1991. *Stand management cooperative field procedures manual: Douglas-fir version.* Stand Management Cooperative, University of Washington, Seattle, WA.

TURNBLOM, E.C. 1998. *Juvenile plantations exhibit the "crossover" effect*. Technical Report 3. Stand Management Cooperative, College of Forest Resources, University of Washington, Seattle, WA.

VANCLAY, J.K., AND J.P. SKOVSGAARD. 1997. Evaluating forest growth models. *Ecol. Model*. 98(1):1–12.

VAUGHN, N.R. 2007. *An individual-tree model to predict the annual growth of young stands of Douglas-fir (*Pseudotsuga menziesii *(Mirbel) Franco) in the Pacific Northwest.* MSc thesis, University of Washington, Seattle, WA.

WOODRUFF, D.R., B.J. BOND, G.A. RITCHIE, AND W. SCOTT. 2002. Effects of stand density on the growth of young Douglas-fir trees. *Can. J. For. Res*. 32:420–427.

YANG, Y., R.A. MONSERUD, AND S. HUANG. 2004. An evaluation of diagnostic tests and their roles in validating forest biometric models. *Can. J. For. Res*. 34(3):619–629.