New Models for Predicting Diameter at Breast Height from Stump Dimensions

James A. Westfall

Models to predict dbh from stump dimensions are presented for 18 species groups. Data used to fit the models were collected across thirteen states in the northeastern United States. Primarily because of the presence of multiple measurements from each tree, a mixed-effects modeling approach was used to account for the lack of independence among observations. The heterogeneous error variance was described as a function of stump diameter, which allowed for more accurate representation of prediction intervals. Application of the mean response model (fixed-effects parameters only) to independent data indicated an average absolute error between 0.2 and 0.7 in. for most groups. An additional advantage is that random-effect parameters allow the model to be calibrated to local conditions if some additional data are available. An example is provided that indicates the local calibration results in a mean residual value that is closer to zero compared with the mean response model. Efforts in other locales to use stump information to inform dbh predictions can obtain the same advancements by adopting a similar modeling methodology.

Keywords: mixed-effects model, stand reconstruction, heterogeneous variance, prediction interval

There are various reasons for reconstruction of sizes of removed trees, including reviewing harvesting practices, assessing damage due to catastrophic events, creating historical records of past management activities, and establishing loss due to timber theft (Wharton 1984, Corral-Rivas et al. 2007). The need to estimate the attributes of removed trees for which no readily usable information is available has led to numerous studies in which tree stump dimensions are used to predict tree characteristics. Most commonly, the relationships between stump characteristics and tree dbh are described via regression models. Most early works consisted of development of lookup tables or ordinary least-squares (OLS) linear models that were limited to species of high commercial importance (Cunningham et al. 1947, Hampf 1957, Bones 1960). Wider ranges of tree species were accommodated in later works. In addition to stump diameter, McClure (1968) used a log transformation of stump height to develop linear regression models for 53 species in the southern United States. A similar methodology was used by Alemdag and Honer (1977) to describe stump/dbh relationships for 11 species in Canada and by Raile (1978) to present linear models for dbh/stump diameter ratios for more than 20 species occurring in the Lake States (Minnesota, Wisconsin, and Michigan). Relationships between stump diameter and dbh for seventeen species in the northeastern United States were described via OLS linear regression models by Wharton (1984). Other research on prediction of dbh using stump information includes simple OLS linear regression models for lodgepole pine (Schlieter 1986), linear and geometric models for white and black oak in Michigan (Ojave et al. 1991), a linear model with a heteroscedastic variance function for baldcypress (Parresol 1993), and nonlinear models using stump diameter as a predictor for southern Indiana oaks (Weigel and Johnson 1997). Recently, Corral-Rivas et al. (2007) developed both linear and nonlinear models to predict dbh from stump diameter for five pine species in central Mexico. It is notable that relatively little work on prediction of dbh from stump dimensions has been done over the last two decades.

The most flexible prediction models were those that used both stump diameter and stump height (McClure 1968, Raile 1978). The models were calibrated using least-squares regression analysis, despite the violation of the assumption of independence of observations (multiple data points were taken from each stump or tree). The proper treatment of correlated observations is necessary to avoid bias in variance estimates (Swindel 1968, Sullivan and Reynolds 1976). Advances in statistical theory now allow for appropriate treatment of the data in the model fitting process. Particularly, correlations between observed data points can be taken into account such that unbiased model error estimates can be obtained (Gregoire et al. 1995). In this report, models that predict dbh from stump dimensions are presented. A mixed-effects modeling approach was taken to account for the within-tree correlations in the data (Garber and Maguire 2003, Trincado and Burkhart 2006). The mixed-model approach was further exploited to illustrate how the models may be refined for a particular area of interest.

Methods

Data

The data used in this study were collected by the Forest Inventory and Analysis (FIA) program of the US Forest Service as part of a tree taper study. The geographic range encompassed 13 states in the northeastern United States, including West Virginia, Maryland, Delaware, New Jersey, Pennsylvania, Ohio, New York, Massachusetts, Rhode Island, Connecticut, Vermont, New Hampshire, and Maine. Data were collected assuming species groupings described by Scott (1981), which are also used by FIA for tree volume estimation (implying similar tree form). The groupings represent a compromise between individual species and broad aggregations (e.g., hardwood and softwood). Tree species frequency and tree size information...
obtained from FIA inventory plots were used to allocate the sample. Geographic dispersion and elimination of the potential need to account for correlations between trees was accomplished by measuring only one tree per species group at a given sampling location. Figure 1 depicts the spatial distribution of sample locations.

Measurements were taken on a total of 2,464 trees over a range of tree species and sizes (Table 1). Height/diameter information was obtained using a Barr & Stroud dendrometer. Measurement points occurred at heights of approximately 1, 2, and 3 ft, where heights were recorded to nearest 0.1 ft and tree diameters were recorded to nearest 0.1 in.; dbh was measured at 4.5 ft of height. There were 7,371 height/diameter data pairs obtained.

Analysis
To predict dbh from stump measurements, there are three primary sources of information: (1) species, (2) the diameter of the stump (d), and (3) the height at which the diameter d occurs (h). The latter two pieces of information, along with the fact that dbh is measured at a height of 4.5 ft, can be used to describe the relationship between the stump dimensions and dbh for a given species or species group. For this study, two measures of distance between the stump height (h) and 4.5 ft were considered. First, the ratio of $4.5/h$ provides a relative metric of how far $h$ is from the point where dbh is measured. Also, the difference $(4.5 - h)$ was calculated to represent the absolute distance between the two points of interest on the tree bole. These two variables, in addition to stump diameter, were used to specify the prediction model,

$$\hat{dbh}_i = d_i \times (4.5/h)^{\hat{\beta}_0} + \beta_1 (4.5 - h) + \varepsilon_i$$  \hspace{1cm} (1)

where $\hat{dbh}_i$ = estimated diameter breast height (in.) for tree $i$; $d_i$ = stump diameter (in.) for tree $i$; $h_i$ = stump height (ft) for tree $i$; $\hat{\beta}_0$, $\beta_1$ = estimated fixed-effects parameters; and $\varepsilon_i$ = random error for tree $i$.

Equation 1 is conditioned such that $\hat{dbh}_i = d_i$ when $h_i = 4.5$ ft. To account for the multiple observations per tree, random-effects parameters were added to Equation 1 to indirectly estimate the within-tree correlations using the variance/covariance matrix of the model parameters.

During model development, it was also noted that the error variance increased with increasing stump diameter (surprisingly, the magnitude of error was unrelated to stump height). To ascertain the nature of this increase, the variances of the $e_i$ for each value of $d_i$ were calculated and analyzed (observed diameters greater than 34 in. were grouped into 0.5-in. classes). The nonlinear trend was accounted for via the error variance formulation given in Equation 3.

$$dbh_i = d_i \times (4.5/h)^{\hat{\beta}_0} + (\hat{\beta}_1 + \theta_d)(4.5 - h_i) + \varepsilon_i$$  \hspace{1cm} (2)

$$\varepsilon_i \sim N(0, \sigma^2 d_i^{\hat{\beta}_0})$$  \hspace{1cm} (3)

where $\hat{\beta}_0$, $\hat{\beta}_1$, $\theta_d$ = estimated fixed-effects parameter; $\sigma^2$ = estimated model error variance; $\theta_{d)}$ = random-effects parameters for tree $i$; $\theta_d - N(0, \sigma^2_d)$; $h = 1, 2$; and other terms are as previously defined.

Results and Discussion
The regression analyses were conducted separately for each of the 18 species groups. Table 2 reports the estimated values for the fixed-effects parameters ($\beta_{0-1}$), model error variance ($\sigma^2$), variance of random-effects ($\sigma_{d-2}^2$), and covariance between random-effects ($\sigma_{1-2}$) for Equations 2 and 3. All fixed effects parameter estimates were statistically significant ($\alpha = 0.05$) for all groups except the $\beta_1$ parameter for three of the species groups (5, 10, and 14). The estimated parameters for $\beta_0$ in the error variance function 3 provided an accurate depiction of the relationship between residual variance and stump diameter. Figure 2 shows this relationship evaluated over the entire data set. Also shown is the constant error variance that would have been estimated in a traditional application of OLS regression. Figure 3 depicts the observed correlation between residual variance and stump height, as well as a linear regression through the data points. The slope of the regression line was not statistically different.
Table 1. Tree frequency, tree size, and sample size information by species for 18 species groups.

<table>
<thead>
<tr>
<th>Species group</th>
<th>Species name</th>
<th>No. of trees</th>
<th>Minimum</th>
<th>Mean</th>
<th>Maximum</th>
<th>dbh (in.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Eastern white pine</td>
<td>102</td>
<td>3.2</td>
<td>12.6</td>
<td>36.7</td>
<td>104</td>
</tr>
<tr>
<td>2</td>
<td>Red pine</td>
<td>38</td>
<td>3.8</td>
<td>10.9</td>
<td>19.5</td>
<td>114</td>
</tr>
<tr>
<td>3</td>
<td>Black spruce</td>
<td>30</td>
<td>3.5</td>
<td>7.7</td>
<td>15.7</td>
<td>90</td>
</tr>
<tr>
<td>4</td>
<td>Red spruce</td>
<td>73</td>
<td>2.8</td>
<td>10.4</td>
<td>22.9</td>
<td>219</td>
</tr>
<tr>
<td>5</td>
<td>White pine</td>
<td>45</td>
<td>3.1</td>
<td>10.0</td>
<td>19.4</td>
<td>135</td>
</tr>
<tr>
<td>6</td>
<td>Balsam fir</td>
<td>150</td>
<td>3.4</td>
<td>9.5</td>
<td>21.3</td>
<td>450</td>
</tr>
<tr>
<td>7</td>
<td>Eastern hemlock</td>
<td>133</td>
<td>3.4</td>
<td>13.6</td>
<td>33.5</td>
<td>399</td>
</tr>
<tr>
<td>8</td>
<td>Norway spruce</td>
<td>12</td>
<td>4.2</td>
<td>10.3</td>
<td>16.1</td>
<td>36</td>
</tr>
<tr>
<td>9</td>
<td>Scotch pine</td>
<td>8</td>
<td>3.7</td>
<td>7.8</td>
<td>13.3</td>
<td>24</td>
</tr>
<tr>
<td>10</td>
<td>Table Mountain pine</td>
<td>5</td>
<td>7.8</td>
<td>10.5</td>
<td>13.9</td>
<td>15</td>
</tr>
<tr>
<td>11</td>
<td>Virginia pine</td>
<td>26</td>
<td>3.1</td>
<td>9.7</td>
<td>16.8</td>
<td>78</td>
</tr>
<tr>
<td>12</td>
<td>Jack pine</td>
<td>3</td>
<td>6</td>
<td>8.4</td>
<td>10.3</td>
<td>9</td>
</tr>
<tr>
<td>13</td>
<td>Larch (introduced)</td>
<td>7</td>
<td>4.4</td>
<td>11.3</td>
<td>16</td>
<td>21</td>
</tr>
<tr>
<td>14</td>
<td>Lodgepole pine</td>
<td>30</td>
<td>5.2</td>
<td>10.8</td>
<td>23.5</td>
<td>90</td>
</tr>
<tr>
<td>15</td>
<td>Pitch pine</td>
<td>30</td>
<td>3.4</td>
<td>9.8</td>
<td>15.3</td>
<td>89</td>
</tr>
<tr>
<td>16</td>
<td>Shortleaf pine</td>
<td>7</td>
<td>6.4</td>
<td>10.6</td>
<td>15.2</td>
<td>21</td>
</tr>
<tr>
<td>17</td>
<td>Tamarack (native)</td>
<td>18</td>
<td>4</td>
<td>9.1</td>
<td>14.2</td>
<td>54</td>
</tr>
<tr>
<td>18</td>
<td>Atlantic white-cedar</td>
<td>29</td>
<td>5</td>
<td>9.1</td>
<td>17.9</td>
<td>103</td>
</tr>
<tr>
<td>19</td>
<td>Eastern hemlock</td>
<td>133</td>
<td>3.4</td>
<td>13.6</td>
<td>33.5</td>
<td>412</td>
</tr>
<tr>
<td>20</td>
<td>Norway spruce</td>
<td>12</td>
<td>4.2</td>
<td>10.3</td>
<td>16.1</td>
<td>36</td>
</tr>
<tr>
<td>21</td>
<td>Scotch pine</td>
<td>8</td>
<td>3.7</td>
<td>7.8</td>
<td>13.3</td>
<td>24</td>
</tr>
<tr>
<td>22</td>
<td>Table Mountain pine</td>
<td>5</td>
<td>7.8</td>
<td>10.5</td>
<td>13.9</td>
<td>15</td>
</tr>
<tr>
<td>23</td>
<td>Virginia pine</td>
<td>26</td>
<td>3.1</td>
<td>9.7</td>
<td>16.8</td>
<td>78</td>
</tr>
<tr>
<td>24</td>
<td>Jack pine</td>
<td>3</td>
<td>6</td>
<td>8.4</td>
<td>10.3</td>
<td>9</td>
</tr>
<tr>
<td>25</td>
<td>Larch (introduced)</td>
<td>7</td>
<td>4.4</td>
<td>11.3</td>
<td>16</td>
<td>21</td>
</tr>
<tr>
<td>26</td>
<td>Lodgepole pine</td>
<td>30</td>
<td>5.2</td>
<td>10.8</td>
<td>23.5</td>
<td>90</td>
</tr>
<tr>
<td>27</td>
<td>Pitch pine</td>
<td>30</td>
<td>3.4</td>
<td>9.8</td>
<td>15.3</td>
<td>89</td>
</tr>
<tr>
<td>28</td>
<td>Shortleaf pine</td>
<td>7</td>
<td>6.4</td>
<td>10.6</td>
<td>15.2</td>
<td>21</td>
</tr>
<tr>
<td>29</td>
<td>Tamarack (native)</td>
<td>18</td>
<td>4</td>
<td>9.1</td>
<td>14.2</td>
<td>54</td>
</tr>
<tr>
<td>30</td>
<td>Atlantic white-cedar</td>
<td>29</td>
<td>5</td>
<td>9.1</td>
<td>17.9</td>
<td>103</td>
</tr>
</tbody>
</table>
Table 2. Estimates (and standard errors) for fixed-effects parameters ($\beta_0$, $\beta_1$, $\beta_2$), model error variance ($\sigma^2$), variance of random-effects ($\sigma^2_1$, $\sigma^2_2$), and covariance between random-effects ($\sigma_{12}$) from Equations 2 and 3 for 18 species groups.

<table>
<thead>
<tr>
<th>Species group</th>
<th>$\beta_0$ (SE)</th>
<th>$\beta_1$ (SE)</th>
<th>$\beta_2$ (SE)</th>
<th>$\sigma^2$ (SE)</th>
<th>$\sigma^2_1$ (SE)</th>
<th>$\sigma^2_2$ (SE)</th>
<th>$\sigma_{12}$ (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.1096 (0.0047)</td>
<td>0.0588 (0.0154)</td>
<td>1.4462 (0.1575)</td>
<td>0.0024 (0.0010)</td>
<td>0.0008 (0.0002)</td>
<td>0.0145 (0.0031)</td>
<td>-0.0003 (0.0005)</td>
</tr>
<tr>
<td>2</td>
<td>-0.1334 (0.0062)</td>
<td>0.0740 (0.0134)</td>
<td>2.1841 (0.1557)</td>
<td>0.0005 (0.0002)</td>
<td>0.0023 (0.0004)</td>
<td>0.0071 (0.0012)</td>
<td>-0.0021 (0.0006)</td>
</tr>
<tr>
<td>3</td>
<td>-0.1353 (0.0084)</td>
<td>0.1451 (0.0184)</td>
<td>2.4605 (0.1222)</td>
<td>0.0002 (0.0001)</td>
<td>0.0024 (0.0005)</td>
<td>0.0065 (0.0009)</td>
<td>-0.0022 (0.0007)</td>
</tr>
<tr>
<td>4</td>
<td>-0.1162 (0.0057)</td>
<td>0.0686 (0.0137)</td>
<td>2.5778 (0.1356)</td>
<td>0.0002 (0.0001)</td>
<td>0.0013 (0.0003)</td>
<td>0.0074 (0.0014)</td>
<td>-0.0000 (0.0004)</td>
</tr>
<tr>
<td>5</td>
<td>-0.1117 (0.0030)</td>
<td>0.1242 (0.0121)</td>
<td>2.4099 (0.1874)</td>
<td>0.0004 (0.0002)</td>
<td>0.0012 (0.0001)</td>
<td>0.0062 (0.0006)</td>
<td>-0.0000 (0.0004)</td>
</tr>
<tr>
<td>6</td>
<td>-0.1631 (0.0103)</td>
<td>0.1517 (0.0305)</td>
<td>2.2783 (0.1585)</td>
<td>0.0005 (0.0002)</td>
<td>0.0040 (0.0009)</td>
<td>0.0157 (0.0034)</td>
<td>-0.0037 (0.0016)</td>
</tr>
<tr>
<td>7</td>
<td>-0.1323 (0.0072)</td>
<td>0.2442 (0.0231)</td>
<td>1.8478 (0.1376)</td>
<td>0.0011 (0.0004)</td>
<td>0.0014 (0.0003)</td>
<td>0.0074 (0.0014)</td>
<td>-0.0006 (0.0004)</td>
</tr>
<tr>
<td>8</td>
<td>-0.0966 (0.0043)</td>
<td>0.0260 (0.0114)</td>
<td>1.7993 (0.1326)</td>
<td>0.0010 (0.0003)</td>
<td>0.0016 (0.0003)</td>
<td>0.0074 (0.0014)</td>
<td>-0.0018 (0.0008)</td>
</tr>
<tr>
<td>9</td>
<td>-0.1074 (0.0064)</td>
<td>0.0685 (0.0148)</td>
<td>1.9470 (0.1657)</td>
<td>0.0008 (0.0003)</td>
<td>0.0024 (0.0004)</td>
<td>0.0091 (0.0016)</td>
<td>-0.0007 (0.0005)</td>
</tr>
<tr>
<td>10</td>
<td>-0.0720 (0.0039)</td>
<td>0.1214 (0.0167)</td>
<td>2.1169 (0.1677)</td>
<td>0.0008 (0.0003)</td>
<td>0.0024 (0.0004)</td>
<td>0.0060 (0.0020)</td>
<td>-0.0002 (0.0009)</td>
</tr>
<tr>
<td>11</td>
<td>-0.1743 (0.0075)</td>
<td>0.1376 (0.0170)</td>
<td>1.7486 (0.1622)</td>
<td>0.0022 (0.0008)</td>
<td>0.0024 (0.0006)</td>
<td>0.0060 (0.0020)</td>
<td>-0.0002 (0.0009)</td>
</tr>
<tr>
<td>12</td>
<td>-0.1171 (0.0055)</td>
<td>0.0714 (0.0143)</td>
<td>2.1741 (0.1358)</td>
<td>0.0006 (0.0002)</td>
<td>0.0015 (0.0003)</td>
<td>0.0097 (0.0019)</td>
<td>-0.0006 (0.0005)</td>
</tr>
<tr>
<td>13</td>
<td>-0.1193 (0.0078)</td>
<td>0.1099 (0.0261)</td>
<td>1.7150 (0.1319)</td>
<td>0.0013 (0.0004)</td>
<td>0.0014 (0.0003)</td>
<td>0.0076 (0.0015)</td>
<td>-0.0000 (0.0005)</td>
</tr>
<tr>
<td>14</td>
<td>-0.1651 (0.0038)</td>
<td>0.1609 (0.0261)</td>
<td>2.4990 (0.1610)</td>
<td>0.0004 (0.0002)</td>
<td>0.0021 (0.0002)</td>
<td>0.0130 (0.0015)</td>
<td>-0.0002 (0.0009)</td>
</tr>
<tr>
<td>15</td>
<td>-0.1329 (0.0082)</td>
<td>0.0659 (0.0224)</td>
<td>1.6448 (0.1551)</td>
<td>0.0022 (0.0008)</td>
<td>0.0026 (0.0005)</td>
<td>0.0133 (0.0032)</td>
<td>-0.0023 (0.0011)</td>
</tr>
<tr>
<td>16</td>
<td>-0.1578 (0.0055)</td>
<td>0.0615 (0.0156)</td>
<td>1.8147 (0.1755)</td>
<td>0.0009 (0.0004)</td>
<td>0.0007 (0.0003)</td>
<td>0.0087 (0.0029)</td>
<td>-0.0018 (0.0003)</td>
</tr>
<tr>
<td>17</td>
<td>-0.1662 (0.0075)</td>
<td>0.1258 (0.0206)</td>
<td>1.4846 (0.1383)</td>
<td>0.0034 (0.0012)</td>
<td>0.0024 (0.0004)</td>
<td>0.0173 (0.0030)</td>
<td>-0.0008 (0.0008)</td>
</tr>
<tr>
<td>18</td>
<td>-0.1382 (0.0104)</td>
<td>0.1010 (0.0278)</td>
<td>1.8050 (0.1674)</td>
<td>0.0019 (0.0008)</td>
<td>0.0015 (0.0004)</td>
<td>0.0151 (0.0036)</td>
<td>-0.0002 (0.0006)</td>
</tr>
</tbody>
</table>

Figure 2. Observed and modeled error variance versus stump diameter across all species groups. Note: Observed diameters greater than 34 in. were grouped into 0.5-in. classes. OLS, ordinary least-squares.

from zero ($P = 0.71$). The variances of the random effects were also significant across all groups, indicating that there were inherent between-tree differences in stump dimension/dbh relationships within species groups. Examination of residuals versus predicted values indicated no systematic trends that would indicate model mis-specification.

Given that application of Equation 2 will be to predict dbh for new observations, prediction error intervals are of primary interest. Quantifying model error can be accomplished through Equation 3. In this setting, random-effects parameters are assigned their expected value of zero. Prediction intervals were computed from (Draper and Smith 1981)

$$
\hat{dbh}_i \pm t_{n-p,0.025} \sqrt{\sigma^2 + \sigma^2_1 h_1^2 \left[1 + x_i (X'X)^{-1} x_i \right]} 
$$

where $n$ = sample size; $p$ = number of estimated parameters; $t_{n-p,0.025}$ = $t$ statistic ($n - p$ degrees of freedom, probability $\alpha/2 = 0.025$); $X$ = regression design matrix for fixed-effects parameters; and $x_i$ = vector from design matrix associated with tree $i$.

As an example, the results for group 16 (hickory) were used to illustrate how the prediction error changes with stump diameter (Figure 4). The model predictions assumed a stump height of 1 ft, and the interval width represents the 95% confidence level. It is shown that the interval is relatively small when tree size is small, e.g., the interval is roughly ± 0.4 in. when dbh is estimated to be 9 in. In contrast, the interval increases to ± 1.9 in. when dbh is estimated to be 30 in. Note that this differs from the classical prediction intervals, where the interval is smallest at the mean of the predictor variable(s) and widens elsewhere (Figure 4).
Figure 3. Observed error variance versus stump height and linear trend line across all species groups. Note: The slope of the trend line was not different from zero ($P = 0.71$).

Figure 4. Mean response and 95% confidence intervals for prediction of new observations using variance from Equation 5 and using variance from ordinary least-squares (OLS) (hickory; species group 16).
Most commonly, the model will be applied by setting the ran-
dom-effects parameters to zero (mean response model). To evaluate
model performance under this scenario, independent data from
studies conducted on seven national forests were used. Forests
within the region where the model fitting data were collected in-
cluded the Monongahela, Green Mountain, White Mountain, and
Allegheny national forests. To obtain more data for some of the
species groups, some forests outside the region were also includ-
ed (Hiawatha, Hoosier, and Chequamegon-Nicolet national forests).
Table 3 provides the means and standard deviations for raw resid-
ual, absolute residual, and mean residual (residual/dbh) by species
group.

These results show the mean residual is generally within ± 0.5
in., although eastern hemlock (group 4) is somewhat higher. Mean
absolute residuals ranged between approximately 0.2 to 0.7 in. for
groups except hemlock (group 4), basswood (group 13), and oak
(group 14). These results compare favorably with those for the same
region reported by Wharton (1984), in which mean error often exceeded 1.0 in. The mean relative residuals indicate that the aver-
age amount of error should be less than 4% of the true dbh For most
species groups, some forests outside the region were also included
(Hiawatha, Hoosier, and Chequamegon-Nicolet national forests).

Table 3. Means (and standard deviations) for raw residual, mean absolute residual, and mean relative residual from mean response model applied to independent data from seven national forests.

\[
\begin{array}{cccc}
\text{Species group} & n & \text{Mean residual} & \text{Mean absolute residual} & \text{Mean relative residual} \\
1 & 160 & 0.077 (0.460) & 0.319 (0.388) & 0.006 (0.038) \\
2 & 19 & 0.057 (0.279) & 0.256 (0.149) & 0.008 (0.033) \\
3 & 12 & -0.212 (0.518) & 0.310 (0.462) & -0.020 (0.047) \\
4 & 8 & -0.751 (1.012) & 0.857 (0.866) & -0.062 (0.074) \\
5 & 112 & 0.300 (0.353) & 0.355 (0.298) & 0.035 (0.039) \\
6 & 53 & -0.463 (0.878) & 0.518 (0.846) & -0.060 (0.107) \\
7 & 213 & -0.393 (1.148) & 0.648 (1.025) & -0.036 (0.087) \\
8 & 46 & -0.025 (0.946) & 0.646 (0.685) & -0.002 (0.063) \\
9 & 63 & -0.393 (0.951) & 0.588 (0.842) & -0.034 (0.077) \\
10 & 46 & 0.017 (0.850) & 0.577 (0.769) & -0.002 (0.046) \\
11 & 68 & 0.162 (0.555) & 0.426 (0.387) & 0.012 (0.046) \\
12 & 47 & -0.186 (1.106) & 0.657 (0.903) & -0.014 (0.072) \\
13 & 25 & -0.528 (1.707) & 1.160 (1.343) & -0.032 (0.113) \\
14 & 99 & 0.483 (1.224) & 0.917 (0.940) & 0.025 (0.064) \\
15 & 21 & -0.009 (0.672) & 0.513 (0.420) & 0.000 (0.040) \\
16 & 26 & 0.177 (0.778) & 0.567 (0.550) & 0.011 (0.053) \\
17 & 41 & -0.065 (1.034) & 0.646 (0.803) & -0.003 (0.066) \\
18 & 88 & 0.355 (0.572) & 0.497 (0.452) & 0.028 (0.045) \\
\end{array}
\]

An alternative approach to implementation is to obtain predi-
cctions of random-effects for new observations using additional in-
formation collected where the model will be applied. For instance,
Lappi (1991) calibrated a height/diameter model from local height
and diameter measurements. Trincado and Burkhart (2006) de-
scribe how to localize a taper model by using upper-stem measure-
ments for the trees of interest. The most likely approach to local
calibration is measurement of a "stump" diameter/height and dbh
from several nearby trees for each species group present. In this case,
a set of random coefficients applicable at the stand level can be
obtained via (Vonesh and Chinchilli 1997)

\[
\theta = D\{ZDZ' + R\}^{-1}(y - Xb)
\]

where \(\theta\) = vector of predicted random-effect parameters; \(B = \) regres-
sion design matrix for random-effects parameters; \(F = \) matrix of
partial derivatives of Equation 2 with respect to each fixed parameter
\((b_{i,s})\) evaluated at \(d_{i,s}\) for each calibration tree; \(Z = FB; R = \)
predicted variance/covariance matrix of residual errors; \(D = \)
predicted variance/covariance matrix of random-effects (from Table 2);
\(y = \) vector of observed tree dbh values; \(X = \) regression design matrix
for fixed-effects parameters; and \(b = \) vector of fixed-effects param-
eters (from Table 2).

Independent data from 31 sugar maple trees (species group 7)
were used to illustrate the process. For each tree, stump diameter at
stump height = 1 ft and dbh information were available. Sixteen
trees were randomly chosen to represent the harvested trees for
which predictions of dbh are desired. Fifteen trees were used to
estimate the random-effects parameters via Equation 5. The results
were \(\theta_1 = 0.0100\) and \(\theta_2 = -0.0687\). Thus, the localized prediction
model for sugar maple is given as

\[
dbh_i = \hat{d}_i \times (4.5/h_i)^{-0.1158 \pm 0.0190}
+ (0.1047 + (-0.0687)) (4.5 - h_i) + \epsilon_i
\]

Applying Equation 6 to the 16 harvest trees shows that the dis-
tribution of the residuals is shifted to be better centered about zero,
compared with the residuals resulting from applying the nonlocal-
ized mean model (Figure 5). The local calibration does result in
poorer predictions for some trees, but better estimates of dbh for the
entire sample are gained; the mean residual for the mean model
(fixed-effects only) was -0.16 in., whereas the mean residual for the
locally calibrated model was -0.04 in.

Conclusion

The models presented differ from previous efforts to estimate
dbh from stump dimensions in three ways. First, the inclusion
of random-effects parameters allows for unbiased estimates of error
variance, which directly affect inferences regarding estimated pa-
rameters and estimation of confidence/prediction intervals. Second,
the ability to locally calibrate the model provides an alternative to
using the mean response over an often large geographic area. Third,
the heterogeneous error variance was described as a function of
stump diameter, which allows for more realistic prediction intervals
than those based on an (often invalid) assumption of homogeneous
error. This represents marked improvements over earlier efforts, in
which such features were lacking.

The models are applicable to many species occurring in the
northeastern United States. Given that many of these species also
occur outside the area used in this study, the model may be used
elsewhere. However, it is recommended that local calibration be
performed if possible, and if not, considerable caution should be
exercised, as unknown biases may produce inaccurate predictions.
Ultimately, the most attractive option for other geographic areas
would be to collect data and adopt a similar modeling methodology.
The predicted values may be used to further construct the missing trees (e.g., via height-diameter models) or as input into volume or biomass models that the user may wish to use.

**Literature Cited**


