*For. Sci.* 60(1):34–42
http://dx.doi.org/10.5849/forsci.12-141

biometrics

# Effects of Uncertainty in Model Predictions of Individual Tree Volume on Large Area Volume Estimates

**Ronald E. McRoberts and James A. Westfall**

Forest inventory estimates of tree volume for large areas are typically calculated by adding model predictions of volumes for individual trees. However, the uncertainty in the model predictions is generally ignored with the result that the precision of the large area volume estimates is overestimated. The primary study objective was to estimate the effects of model residual variability and model parameter uncertainty on large area volume estimates and their uncertainties for a study area in northeastern Minnesota, USA. Monte Carlo simulation approaches were used because of the complexities associated with multiple sources of uncertainty and the nonlinear nature of the models. Two conclusions were important. First, for this study, the effects of uncertainty in model predictions on the large area volume estimates and their uncertainties were small when the models were calibrated using an average of 100 or more observations per species and when the average proportion of variance explained by the models was at least 0.95. Second, large area estimates and their uncertainties based on coniferous/deciduous and nonspecific models deviated very little from large area estimates based on species-specific models.

**Keywords:** allometric model, model parameter uncertainty, residual uncertainty, Monte Carlo simulation

Statistical models are routinely used by forest inventory and monitoring programs to predict the volume, biomass, or carbon content for individual trees using measurements of tree attributes such as species, diameter, and height as independent variables. The individual tree model predictions are often aggregated at the plot level and used as training and/or accuracy assessment data for remote sensing-based applications or aggregated at the plot level and then added or averaged over plots to produce large area estimates. When the individual tree model predictions are aggregated at the plot level, inventory and monitoring programs routinely ignore the effects of uncertainty in the model predictions. The problem that results is that all trees with the same values of independent variables receive the same model prediction for volume, whereas in reality different trees with the same values of the independent variables may have quite different volumes. Failure to account for this latter tree-to-tree variability by treating model predictions as observations causes erroneously optimistic precision estimates. From a rigorous statistical perspective, this widely used practice of ignoring model prediction uncertainty cannot be justified. However, from a practical perspective, the practice can perhaps be justified if the effects of model prediction uncertainty are negligible relative to the effects of other sources of uncertainty such as plot-to-plot variability.

Uncertainty in model predictions can be attributed to four primary sources: (1) model misspecification, (2) uncertainty in values of the independent variables, (3) residual variability expressed as differences between observations and predictions obtained from correctly specified models, and (4) uncertainty in the model parameter estimates. Model misspecification is due to the lack of appropriate model calibration data (Breidenbach et al. 2013) or the lack of modeling expertise, must be assessed on a case-by-case basis, and is not a subject of the current study. The effects on the uncertainty of model predictions due to uncertainty in values of the independent variables, whether the result of measurement error and or simply observer-to-observer variability, have been studied extensively (Gertner and Dzialowy 1984, Kangas 1996, Mowrer and Frayer 1986, Gertner 1987, 1990, Mowrer 1991, Gertner and Köhl 1992, McRoberts et al. 1994, McRoberts 1996, McRoberts and Lessard 2001, Westfall and Patterson 2007, Berger et al. 2013). The general consensus is that this source of uncertainty is a major contributor to

Manuscript received November 13, 2012; accepted May 9, 2013; published online August 29, 2013.

Affiliations: *Ronald E. McRoberts (rmcroberts@fs.fed.us), USDA Forest Service, Northern Research Station, Forest Inventory and Analysis, St. Paul, MN. James A. Westfall (jameswestfall@fs.fed.us), USDA Forest Service.*

Acknowledgments: *We acknowledge the assistance of Brian Walters and Dr. Grant Domke, both of the USDA Forest Service, Northern Research Station, St. Paul, MN.*

model prediction uncertainty. Based on the considerable body of literature already available, this source of uncertainty is also not a subject of the current study.

The effects on model predictions of uncertainty from the other two sources, residual variability and model parameter uncertainty, result from attributes of the data and have been studied much less extensively. For a correctly specified model, residual variability is closely related to the quality of fit of the models to the data as measured by indices such as the root mean square error, modeling efficiency (Vanclay and Skovsgaard 1997), and the Akaike information criterion (Akaike 1974). The effects of this source of uncertainty on the uncertainty of model predictions vary by application (Cunia 1965, Gertner 1987, Gertner and Köhl 1992, McRoberts and Lessard 2001).

The effects of uncertainty in model parameter estimates also vary by application (Mowrer and Frayer 1986, Cunia 1987, Gertner 1990, Mowrer 1991, McRoberts and Lessard 2001). For linear models, this source of uncertainty can be easily quantified as the covariance matrix for the model parameter estimates using a first-order Taylor series. For nonlinear models, a first-order Taylor series produces only an approximation for the covariance matrix with the quality of the approximation dependent on the mathematical form of the model and quality of fit of the model to the data (Ratkowsky 1983). When the Taylor series approximations are poor, the Monte Carlo approach described by McRoberts and Lessard (2001, p. 72) can be used.

Several approaches have been used to propagate uncertainty from multiple and varied sources to model predictions and large area estimates. The first approach is based on sampling theory and entails theoretical derivations of expressions for the magnitudes of uncertainty from the different sources and expressions for their simultaneous effect on the uncertainty of estimates. The general approach is similar to the approaches reported by Cunia (1987) and Ståhl et al. (2011, 2013), but has not often been reported in the literature, perhaps because of the theoretical complexities. The second approach, characterized as the method of statistical differentials (Kempthorne and Folks 1971) or error propagation (Mowrer and Frayer 1986, Mowrer 1991, Gertner 1987, 1990), uses first-order Taylor series approximations to produce an additive expression for total uncertainty of the estimate of interest (Gertner 1990). The third approach uses Monte Carlo simulations (Gertner 1987, Gertner and Dzialowy 1984, Kangas 1996, Mowrer 1991, McRoberts et al. 1994, McRoberts 1996, McRoberts and Lessard 2001, Breidenbach et al. 2013). This approach is motivated to a large degree by the near intractability of the mathematics involved when uncertainty from multiple sources must be propagated through complex models. The gain in tractability associated with Monte Carlo approaches may be offset by much greater computational intensity, although increasingly greater computing capabilities mitigate the latter factor.

The overall conceptual objective was to determine whether the practice of ignoring model prediction uncertainty could be justified from a practical perspective. The particular technical objective of the study was to assess the effects of residual variability and uncertainty in model parameter estimates on the uncertainty of large area estimates of mean volume per unit area. The estimates are obtained by aggregating model predictions of volume for individual trees within plots and then calculating the mean over plots. A Monte Carlo approach was used with species-specific models constructed specifically for this study so that residual variability and uncertainty in model parameter estimates could be correctly quantified. The relevance of the study extends beyond well-established inventory and monitoring programs such as the national forest inventories conducted by North American and European countries to the monitoring, reporting, and verification programs that are being established in tropical countries under the auspices of the United Nations REDD (Reducing Emissions from Deforestation and forest Degradation) program.

## Data

### Study Area

The study area was defined by Minnesota Survey Unit 1 for the Forest Inventory and Analysis (FIA) program of the Northern Research Station, USDA Forest Service (Figure 1). The study area includes approximately 33,353 $km^2$ (12,877 $mi^2$) with land use consisting of forestland dominated by aspen-birch and spruce-fir associations, agriculture, wetlands, and water.

### Estimation Data Set

The FIA program conducts the national forest inventory (NFI) of the United States and has established field plot centers in permanent locations using a sampling design that is regarded as producing an equal probability sample (McRoberts et al. 2005). Each FIA plot consists of four 7.32-m (24-ft) radius circular subplots that are configured as a central subplot and three peripheral subplots with centers located at distances of 36.58 m (120 ft) and azimuths of 0, 120, and 240° from the center of the central subplot. Field crews visually estimate the proportion of each subplot that satisfies the FIA definition of forestland: minimum area of 0.4 ha (1.0 acre); minimum crown cover of 10%; stand width, measured as external crown-to-crown distance, of at least 36.6 m (120 ft); and forestland use. Field crews also observe species and measure diameter at breast height (dbh, 1.37 m, 4.5 ft) and height (ht) for all trees with dbh of at least 12.7 cm (5 in.). Volumes for individual trees are estimated using statistical models, aggregated at the plot level, expressed as volume per unit area, and for inventory estimation purposes are considered to be observations without error. For this study, data were used from 2,178 FIA plots on forestland with 50,176 trees representing 38 species (Table 1). For future reference, these data are characterized as the estimation data set.

### Model Calibration Data Sets

The data used to calibrate the volume models in this study (Table 1) were originally collected for a taper model study (Westfall and Scott 2010) encompassing the 24 states of the United States for which the regional FIA program of the Northern Research Station has inventory responsibility (Figure 1). The geographic sources of the data used to calibrate the models for this study were restricted to the states of Michigan, Minnesota, and Wisconsin (Figure 1), which span the ecological province that includes the study area. Exceptions were made for a few minor species for which sample sizes were small (Table 1).

Sample trees were selected based on species frequency and tree size information obtained from FIA plots whose conditions satisfied the definition of forestland. Data were acquired for a range of tree sizes and for all major species groups (Scott 1981) subject to the constraints that data were not acquired for more than one tree per species group at each sample location or for trees exhibiting abnormal conditions such as broken tops or excessive lean. The sampling strategy ensured a wide geographic distribution of observations and circumvented the necessity of accommodating spatial correlations

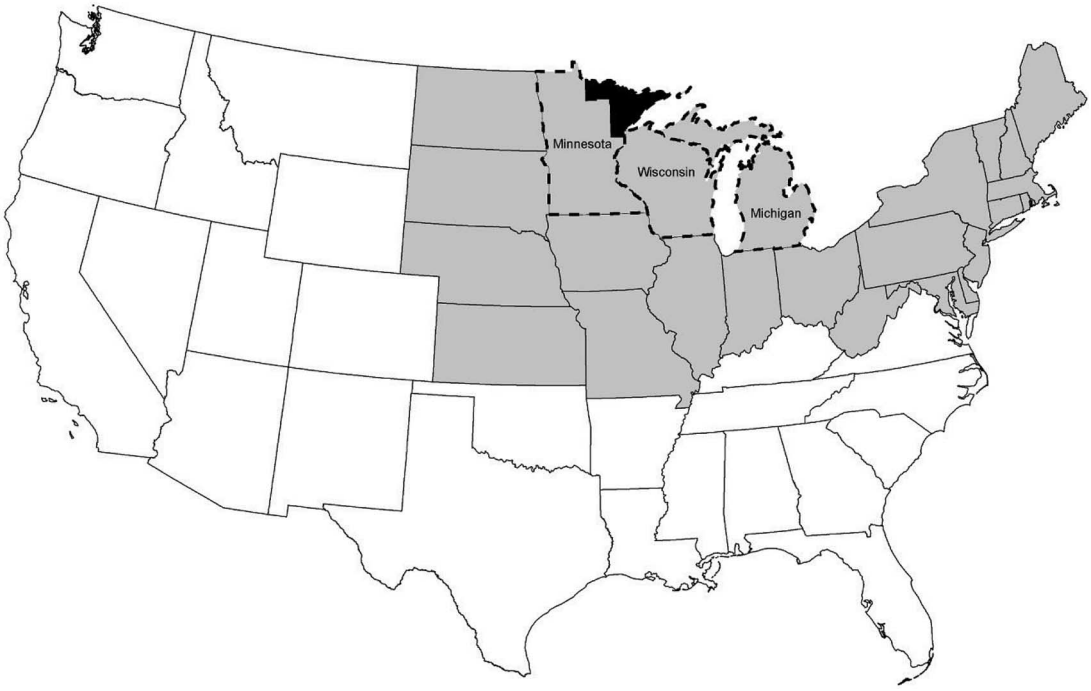


**Figure 1. Minnesota Survey Unit 1 (black), area of Northern Research Station inventory responsibility (gray), and geographic source of model calibration data (Minnesota, Wisconsin, and Michigan).**

among observations for trees at the same sample locations when species-specific models were constructed.

For individual trees, diameter measurements were obtained using a Barr and Stroud dendrometer at heights of 0.3, 0.6, 0.9, 1.4, and 1.8 m and at approximately 2.5-cm taper intervals up to total tree height. Diameter measurements were recorded to the nearest 0.25 cm, and tree heights were recorded to the nearest 0.03 m. Volumes of sections between height measurements were calculated using Smalian's formula (Avery and Burkhart 2002) as the product of mean cross-sectional area and section length. Total stem volumes for individual trees were calculated by adding volumes for all sections, characterized as volume over bark to the top of the tree, and considered to be observations without error. For future reference, these data are characterized as the species-specific calibration data sets.

## Methods

### Volume Model

The individual tree volume models used by NFIs have typically been developed independently of other NFIs and have been revised and modified over long periods of time. As a result, the mathematical forms of these NFI volume models tend to vary considerably. However, in recent years, allometric forms for these models have gained in popularity. Allometry refers to relationships between size and shape, and allometric or power law models are often used to characterize biological phenomena for which the actual mechanisms underlying the relationship are too complex to describe in detail, but predictions are still needed. Allometric models are often of the general form,

$$Y = \beta_0 X_1^{\beta_1} \ldots X_p^{\beta_P} + \varepsilon,$$

or equivalently,

$$Y = \exp[\alpha_0 + \alpha_1 \cdot \ln(X_1) + \cdots + \alpha_p \cdot \ln(X_p)] + \varepsilon,$$

where $Y$ is the response or dependent variable, exp(.) is the exponential function, $X$ is the predictor or independent variables, the $\beta$s and the $\alpha$s are parameters whose estimates are obtained by fitting the model to data, and $\varepsilon$ is a random residual error. This mathematical form has emerged as increasingly popular and has been used in China (Xiang et al. 2011), throughout Europe (Zianis et al. 2005), in the United States (Jenkins et al. 2003), in Norway (Breidenbach et al. 2013), in New Zealand (Beets et al. 2012), and for many recent tropical applications (Brown et al. 1989, Ketterings et al. 2001, Chave et al. 2004, 2005, Litton and Kauffman 2008, Basuki et al. 2009, Kamelarczyk 2009, Návar 2009).

For this study, the relationship between individual tree volume ($V$) as the dependent variable and dbh and ht as the independent variables was modeled using the allometric form,

$$V_i = \beta_0 \cdot \mathrm{dbh}_i^{\beta_1} \cdot \mathrm{ht}_i^{\beta_2} + \varepsilon_i, \quad (1)$$

where $i$ indexes individual trees in the calibration data set. Because of heterogeneity of residual variance, reweighted nonlinear least squares techniques were used to fit the model to the data and estimate the parameters. Residual deviations were calculated as $\varepsilon_i = \hat{V}_i - V_i$ where

$$\hat{V}_i = \hat{\beta}_0 \cdot \mathrm{dbh}_i^{\hat{\beta}_1} \cdot \mathrm{ht}_i^{\hat{\beta}_2} \quad (2)$$

is the model prediction. The quality of fit of the model to the data was assessed using model efficiency (Vanclay and Skovsgaard 1997) calculated as

$$Q^2 = 1 - \frac{SS_{res}}{SS_{mean}}, \quad (3)$$

where $SS_{res} = \sum_{i=1}^{n_{cal}} \varepsilon_i^2$, $SS_{mean} = \sum_{i=1}^{n_{cal}} (V_i - \bar{V})^2$, $\bar{V} = \frac{1}{n_{cal}} \sum_{i=1}^{n_{cal}} V_i$, and $n_{cal}$ is the size of the calibration data set. Although in some cases $Q^2$

**Table 1. Species observed in the study area.**

| Common name | Scientific name | Estimation data set (proportion of trees) | Calibration data sets | |
|---|---|---|---|---|
| | | | No. of trees | $Q^2 = 1 - (SS_{res}/SS_{mean})$ |
| Balsam fir | *Abies balsamea* | 0.1127 | 65 | 0.9806 |
| Tamarack | *Larix laricina* | 0.0540 | 93 | 0.9551 |
| White spruce | *Picea glauca* | 0.0318 | 53 | 0.9953 |
| Black spruce | *Picea mariana* | 0.1619 | 90 | 0.9869 |
| Jack pine | *Pinus banksiana* | 0.0320 | 107 | 0.9714 |
| Red pine | *Pinus resinosa* | 0.0410 | 79 | 0.9792 |
| Eastern white pine | *Pinus strobus* | 0.0107 | 76 | 0.9894 |
| Scotch pine | *Pinus sylvestris* | 0.0001 | 1[1] | |
| Northern white cedar | *Thuja occidentalis* | 0.1119 | 89[2] | 0.9909 |
| Box elder | *Acer negundo* | 0.0010 | 0[3] | |
| Red maple | *Acer rubrum* | 0.0368 | 111 | 0.9785 |
| Silver maple | *Acer saccharinum* | 0.0002 | 118[2] | 0.9722 |
| Sugar maple | *Acer saccharum* | 0.0283 | 132 | 0.9524 |
| Mountain maple | *Acer spicatum* | 0.0001 | 0[3] | |
| Serviceberry | *Amelanchier* spp. | <0.0001 | 0[3] | |
| Yellow birch | *Betula alleghaniensis* | 0.0032 | 73 | 0.9789 |
| Paper birch | *Betula papyrifera* | 0.0895 | 86 | 0.9764 |
| Black ash | *Fraxinus nigra* | 0.0656 | 89 | 0.9639 |
| Green ash | *Fraxinus pennsylvanica* | 0.0045 | 56[2] | 0.9809 |
| Apple | *Malus spp* | 0.0001 | 0[3] | |
| Eastern hophornbeam | *Ostrya virginiana* | 0.0002 | 11[3] | |
| Balsam poplar | *Populus balsamifera* | 0.0206 | 102 | 0.9842 |
| Bigtooth aspen | *Populus grandidentata* | 0.0049 | 93 | 0.9860 |
| Quaking aspen | *Populus tremuloides* | 0.1744 | 156 | 0.9649 |
| Pin cherry | *Prunus pensylvanica* | 0.0002 | 1[3] | |
| Black cherry | *Prunus serotina* | <0.0001 | 162[2] | 0.9812 |
| Common chokecherry | *Prunus virginiana* | <0.0002 | 1[3] | |
| Northern pin oak | *Quercus ellipsoidalis* | 0.0004 | 7[3] | |
| Bur oak | *Quercus macrocarpa* | 0.0020 | 14[3] | |
| Northern red oak | *Quercus rubra* | 0.0019 | 97 | 0.9847 |
| Willow spp. | *Salix spp.* | 0.0002 | 7[3] | |
| Peachleaf willow | *Salix amygdaloides* | 0.0002 | 14[3] | |
| Black willow | *Salix nigra* | 0.0001 | 4[3] | |
| Bebb willow | *Salix bebbiana* | 0.0001 | 0[3] | |
| American mountain ash | *Sorbus americana* | 0.0003 | 0[3] | |
| American basswood | *Tilia americana* | 0.0065 | 64 | 0.9892 |
| American elm | *Ulmus americana* | 0.0024 | 67 | 0.9786 |
| Slippery elm | *Ulmus rubra* | <0.0001 | 4[4] | |
| Coniferous | | 0.5562 | 648 | 0.9866 |
| Deciduous | | 0.4438 | 1,454 | 0.9689 |
| Nonspecific | | 1.0000 | 2,102 | 0.9717 |

[1] Uses jack pine model.
[2] Uses data from throughout region of Northern Research Station (Figure 1).
[3] Uses deciduous model.
[4] Uses American elm model.

is equivalent to the more familiar $R^2$ used for linear regression models, this is not necessarily the case for most nonlinear models (Anderson-Sprecher 1994).

Quality of fit of the models to the calibration data and residual variability were further assessed using a six-step procedure (Hosmer and Lemeshow 1989): (1) the triplets ($\varepsilon_i$, $V_i$, $\hat{V}_i$) were ordered with respect to $\hat{V}_i$; (2) the ordered triplets were grouped to produce at least 10 groups but with group size not exceeding 25; (3) for the $g$th group, $\bar{V}_g = \frac{1}{n_g}\sum_{i=1}^{n_g} V_i$, $\bar{\hat{V}}_g = \frac{1}{n_g}\sum_{i=1}^{n_g} \hat{V}_i$, and $\sigma_g^2 = \frac{1}{n_g - 1}\sum_{i=1}^{n_g} \varepsilon_i^2$ were calculated where $n_g$ is the number of triplets in the $g$th group; (4) a graph of $\bar{V}_g$ versus $\bar{\hat{V}}_g$ was constructed to assess quality of fit of the model to the data; (5) the relationship between $\sigma_g$ and $\bar{\hat{V}}_g$ was estimated using a linear model through the origin as

$$\hat{\sigma}_g = \hat{\gamma} \cdot \bar{\hat{V}}_g \tag{4}$$

where $\gamma$ is a model parameter; and (6) the group model predictions from Equation 4 were graphed against the group estimates from step 3 to assess the quality of fit of the model to the data. If the models expressed by Equations 1 and 4 fit their respective data, the data points for the graphs should lie along the 1:1 line with intercept 0 and slope 1.

Three approaches were considered for estimating the effects of uncertainty in the model parameter estimates. The first approach was based on approximating the covariance matrix for the parameter estimates using a first-order Taylor series linearization as

$$\widehat{\mathrm{Var}}(\hat{\boldsymbol{\beta}}) = Z' \cdot W \cdot Z, \tag{5}$$

where $W$ is an $n_{cal} \times n_{cal}$ diagonal matrix with entries $w_{ii} = \hat{\sigma}_i^{-2}$ and $\hat{\sigma}_i^2$ is calculated from Equation 4 using $\hat{V}_i$ as the independent variable. However, preliminary analyses revealed that for the smaller sample sizes and smaller values of $Q^2$ used for simulations for this study, the covariance matrix approximations were poor. This result is attributed to the nonlinearity of the model in its parameters, which often produces nonelliptical parameter confidence regions that are not well approximated by the elliptical regions produced by Equation 5 (Bates and Watts 1988, Figure 6.9). The second approach is based on Monte Carlo bootstrapping and entails constructing empirical distributions of the model parameter estimates. However, this approach would require fitting the nonlinear model of Equation 1 to tens of thousands of separate data sets, an onerous and labor-intensive task that is difficult to automate.

The third approach, which was selected, was based on linearizing the model using natural logarithm (ln) transformations of both sides of Equation 1, ignoring the residual, which results in

$$\ln(V) = \alpha_1{}^{\text{lin}} + \alpha_2{}^{\text{lin}} \cdot \ln(\text{dbh}) + \alpha_3{}^{\text{lin}} \cdot \ln(\text{ht}) + \varepsilon, \qquad (6)$$

where the $\alpha^{\text{lin}}$ is a vector of parameters to be estimated and $\varepsilon$ is a random residual, although not the same residual as for Equation 1. Graphical analyses revealed no indication of heteroscedasticity for any species or species group, so that unweighted linear regressions could be used on the ln-ln scale when the model of Equation 6 was fitted. For all species and species combinations, the smallest $Q^2$ on the ln-ln scale was $Q^2 = 0.9994$. When back-transforming the model predictions to the original scale, a factor of $\frac{\hat{\sigma}_i^2}{2}$ was used to correct for transformation bias as

$$\hat{V}_i = \exp\left[\hat{\beta}_0 + \hat{\beta}_1 \cdot \ln(\text{dbh}_i) + \hat{\beta}_2 \cdot \ln(\text{ht}_i) + \frac{\hat{\sigma}_i^2}{2}\right], \qquad (7)$$

where exp(.) is the exponential function, ln(.) is the natural logarithm function, and the $\hat{\beta}$s and $\hat{\sigma}_i^2$ are the parameter estimates and the residual variance, respectively, as estimated on the ln-ln scale (Baskerville 1972). Homoscedasticity on the ln-ln scale permits a common value, $\hat{\sigma}^2 = \hat{\sigma}_i^2$, to be used to correct for bias for all predictions for the same species or species group, whereas heteroscedasticity would require a different value for each prediction. Differences between model predictions obtained using Equation 1 and back-transformed model predictions using Equation 7 were negligible for all species and species groups. Preliminary analyses indicated that this procedure produces variability in model predictions on the original scale corresponding to uncertainty in the nonlinear parameters of the model of Equation 1.

In addition to species-specific models, models were constructed and assessed for coniferous and deciduous groupings of species and for all species grouped together. When the latter models were constructed, correlations among observations in the calibration data sets for trees in close physical proximity were ignored based on the relatively low frequency of such observations.

### Simulating Uncertainty

For a given calibration data set size and residual variance, a five-step procedure was used to estimate the uncertainty of plot-level and large area estimates of mean volume per unit area.

**Step 1. —**For each species or species group:

a. The calibration data set was randomly resampled with replacement to produce data sets of sizes $n_{cal} = 25$, $n_{cal} = 50$, $n_{cal} = 100$, $n_{cal} = 250$, and $n_{cal} = 1{,}000$.

b. For each element of each data set, an observation was simulated as the sum of a prediction calculated using Equation 2 and a residual randomly selected from a Gaussian distribution with mean zero and variance obtained from Equation 4.

c. Each data set was transformed to the ln-ln scale.

d. A linear model as per Equation 6 was fit to the transformed data and the vector of parameter estimates, $\hat{\boldsymbol{\alpha}}^{\text{lin}}$, and residual variance, $\hat{\sigma}^2$, were calculated.

**Step 2. —**For the $i$th tree on the $j$th plot in the estimation data set, a volume observation was simulated using the parameter estimates from step 1 as

$$V_{ij} = \exp[\hat{\alpha}_1^{\text{lin}} + \hat{\alpha}_2^{\text{lin}} \cdot \ln(\text{dbh}_{ij}) + \hat{\alpha}_3^{\text{lin}} \cdot \ln(\text{ht}_{ij}) + \frac{\hat{\sigma}_2}{2}] + \lambda\varepsilon,$$

where $\varepsilon$ is a residual randomly selected from a Gaussian distribution with mean 0 and variance calculated from Equation 4 and $\lambda$ is a multiplicative factor selected to produce $Q^2 = 0.85$, $Q^2 = 0.90$, $Q^2 = 0.95$, $Q^2 = 0.98$, and $Q^2 = 0.99$.

**Step 3. —**The total volume for the $j$th plot in the estimation data set was calculated as $V_j = \sum_{i=1}^{n_j} v_{ij}$ where $n_j$ is the number of trees on the plot.

**Step 4. —**The mean and variance of the mean over all plots for the $k$th simulation were calculated as

$$\bar{V}^k = \frac{1}{n_{\text{plot}}} \sum_{j=1}^{n_{\text{plot}}} V_j, \qquad (8)$$

and

$$\widehat{\text{Var}}(\bar{V}^k) = \frac{1}{n_{\text{plot}}(n_{\text{plot}}-1)} \sum_{j=1}^{n_{\text{plot}}} (V_j - \bar{V}^k)^2, \qquad (9)$$

where $n_{\text{plot}}$ is the number of plots.

**Step 5. —**For each value of $n_{cal}$ and $\lambda$, steps 1–4 were replicated and the mean and variance over replications were calculated as per Rubin (1987) as

$$\hat{\mu}_{\text{sim}} = \frac{1}{n_{\text{sim}}} \sum_{k=1}^{n_{\text{sim}}} \bar{V}^k, \qquad (10)$$

and

$$\widehat{\text{Var}}(\hat{\mu}_{\text{sim}}) = \left(1 + \frac{1}{n_{\text{sim}}}\right) \cdot W_1 + W_2, \qquad (11)$$

where $W_1 = \frac{1}{n_{\text{sim}} - 1} \sum_{k=1}^{n_{\text{sim}}} (\bar{V}^k - \hat{\mu}_{\text{sim}})^2$ is the among-simulations variance and $W_2 = \frac{1}{n_{\text{sim}}} \sum_{k=1}^{n_{\text{sim}}} \widehat{\text{Var}}(\bar{V}^k)$ is the mean within-simulation variance. The replications continued until $\hat{\mu}_{\text{sim}}$ and $\widehat{\text{Var}}(\hat{\mu}_{\text{sim}})$ stabilized.

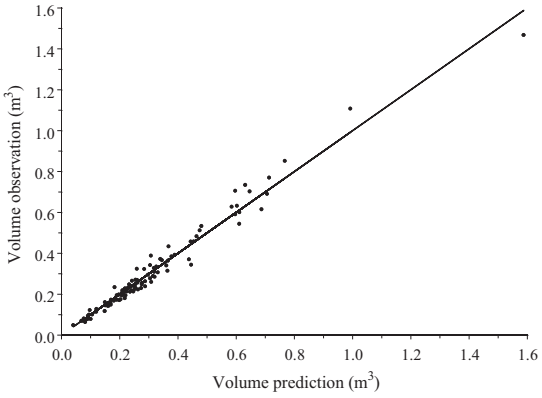


**Figure 2. Individual tree model predictions versus observations for jack pine.**

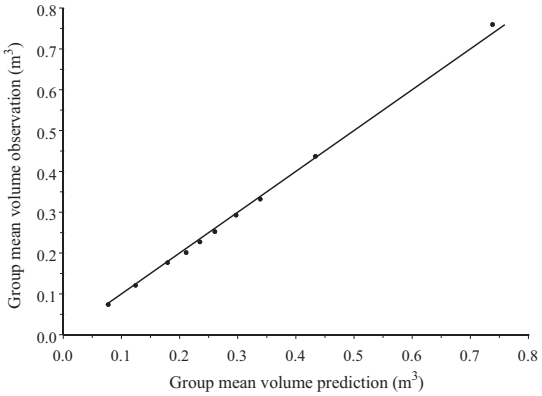


**Figure 3. Group means of individual tree model predictions versus observations for jack pine.**

Correlations among residuals for trees on the same plot in step 2 would be expected to differ by plot, by stand history and density, by tree-to-tree distance, and by tree size, age, and species. However, because estimates of these correlations are not available, a sensitivity analysis was conducted using four levels of among-tree correlations: $\rho = 0.00$, $\rho = 0.25$, $\rho = 0.50$, and $\rho = 0.75$. Although these levels are acknowledged to be only crude approximations, they are sufficient to determine whether correlations among residuals for trees on the same plot have non-negligible effects on large area variance estimates.

This procedure was also used with species-specific values of $n_{cal}$ and $Q^2$ obtained for the calibration data sets.

### Analyses

The estimation data set included some tree species for which there were no observations or only a few observations in the calibration data sets (Table 1). For these species, one of three approaches was used in the following order of priority. First, if sufficient observations for the species were available elsewhere in the 24-state region, they were used to augment the original calibration data sets; models were then constructed using the augmented data sets and applied without regard to the source of the calibration data. Second, a model was used for a similar species for which sufficient calibration data were available. Third, if the species was coniferous, the coniferous model was used, and if the species was deciduous, the deciduous model was used. The proportions of affected trees in the study areas were 0.1166 for the first approach, 0.0002 for the second approach, and 0.0049 for the third approach (Table 1).

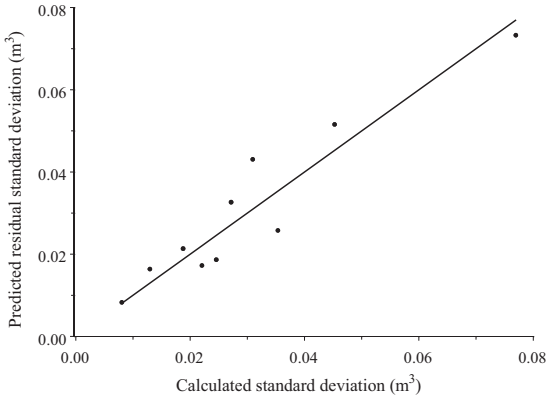


**Figure 4. Group residual standard deviation versus group means of model predictions for jack pine.**

The five-step simulation procedure was implemented using species-specific volume models, models for coniferous and deciduous groupings of species, and a nonspecific model for all species grouped together. Different levels of calibration data set size affected only the covariances for the parameter estimates, but different levels of residual variance affected both covariances for the parameter estimates and individual tree model predictions as per Equation 7. Estimates obtained from the simulations were compared with estimates obtained using model predictions obtained directly from Equation 1 with no incorporation of uncertainty in parameter estimates or residual variability. In addition, estimates obtained using the species-specific models were compared with estimates obtained using the coniferous/deciduous and nonspecific models.

## Results and Discussion

### General

The fits of the models to the data produced $Q^2$ from Equation 3, ranging from $Q^2 = 0.9524$ for sugar maple to $Q^2 = 0.9953$ for white spruce for the species-specific models (Table 1) and $Q^2 = 0.9866$, $Q^2 = 0.9689$, and $Q^2 = 0.9717$ for coniferous, deciduous, and nonspecific models, respectively. Graphs of observations versus predictions for individual trees (Figure 2) and graphs of group observation means versus group prediction means (Figure 3) indicated no systematic lack of fit. Graphs of predicted group standard deviations versus group means revealed no systematic lack of fit (Figure 4), thus indicating that the linear regression models through the origin were sufficient to predict residual variance from volume predictions.

Although estimates of simulation means from Equation 8 and standard errors of means calculated as square roots of variances from Equation 9 typically stabilized by 5,000 simulations (Figure 5), all analyses were based on 10,000 simulations.

The effects of correlations among residuals for trees on the same plot on the uncertainty of large area volume estimates were negligible for all levels and models. A similar finding was reported by Berger et al. (2013) and Breidenbach et al. (2013). Therefore, for this study, further results are reported only for the case of no correlations among residuals.

When no uncertainty was incorporated, differences among estimates of means for the species-specific, coniferous/deciduous, and nonspecific models were relatively small, not more proportionally than 0.025. The differences can be at least partially attributed to differences in species proportions for the calibration and estimation

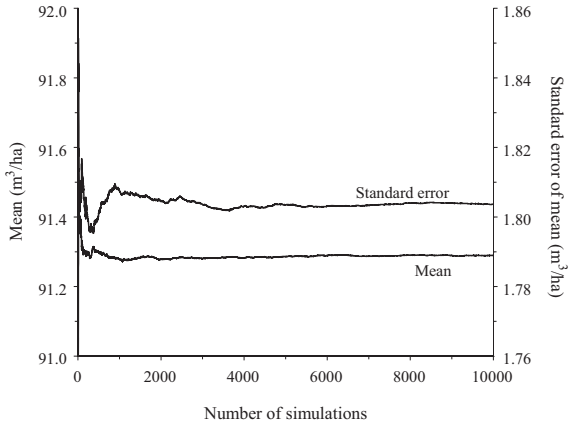


**Figure 5. Means and standard errors of means versus number of simulations.**

data sets. For example, the proportion of deciduous trees in the calibration data sets was approximately 0.69, whereas the proportion in the estimation data set was approximately 0.45. Thus, parameter estimates for the nonspecific model would be skewed toward the deciduous species. However, because species proportions in estimation data sets vary by region, the effects of skewness in the nonspecific model will also vary by region. Differences among the models with respect to the standard errors (SEs) were also relatively small, less proportionally than 0.035. The important result was that SEs for the coniferous/deciduous and nonspecific models were less than those for the species-specific models. This result can be attributed to the combined effects of two factors. First, $Q^2$ values for the coniferous/deciduous and for the nonspecific models were only slightly smaller than those for the species-specific models even though data for multiple species had been combined (Table 1). Second, the much larger calibration data set sizes for the coniferous/deciduous and for the nonspecific models resulted in much smaller covariances for the model parameter estimates.

For the species-specific models with different levels of calibration data set sizes and residual variances, differences between simulation estimates of mean volume per unit area and the mean obtained when no uncertainty was incorporated were small with a maximum proportional difference of less than 0.01. The slight differences can be attributed to the nonlinearity of the models. For models that are nonlinear in the parameters, unlike linear models, changing a parameter by equal amounts in opposite directions does not necessarily produce changes in model predictions that are equal but of opposite direction. Thus, greater differences were realized for smaller calibration sample sizes and greater residual variances because these conditions produced greater covariances and therefore greater deviations of the parameter estimates from their means. For the species-specific models, SEs increased for smaller calibration data set sizes and smaller $Q^2$, as expected, and reflected the effects of greater uncertainty in model predictions due to uncertainty in the parameter estimates and residual variability. For the coniferous/deciduous and nonspecific models, simulation means differed from means obtained when no uncertainty was incorporated proportionally by less than 0.0015. SEs also increased with smaller calibration data set sizes and values of $Q^2$ but not by as much as for the species-specific models.

### Implications for the FIA Program

For this study, the mean size of the calibration data sets for individual species was 66.31 and the mean weighted species-specific $Q^2$ was $Q^2 = 0.974$ where the weights were proportional to the calibration data set sizes. For the species-specific models, estimates of the large area mean and SE when no uncertainty was incorporated were 92.599 and 1.754 $m^3$/ha, whereas the simulation estimates that incorporated uncertainty were 92.614 and 1.779 $m^3$/ha. For the coniferous/deciduous models, the estimates that incorporated no uncertainty were 90.539 and 1.701 $m^3$/ha, whereas the simulation estimates that incorporated uncertainty were 90.547 and 1.712 $m^3$/ha. For the nonspecific models, the estimates that incorporated no uncertainty were 90.395 and 1.695 $m^3$/ha, whereas the simulation estimates that incorporated uncertainty were 90.413 and 1.706 $m^3$/ha. For all three model combinations, differences between estimates of means obtained with and without incorporated uncertainty were proportionally less than 0.001, and differences between SEs were proportionally less than 0.025; for practical purposes, these deviations are negligible. These results suggest that with respect to estimates of mean volume per unit area, little may be lost by using the nonspecific models, and with respect to the SEs of the estimates of the means, nothing is gained by using the species-specific models.

For operational purposes, the FIA program uses the volume models reported by Hahn (1984), not the models constructed for this study. For the Hahn (1984) models, the mean size of the species-specific calibration data sets was 2,156 and the mean weighted $Q^2$ was $Q^2 = 0.825$. These values produced an estimate of mean volume per unit area of 92.599 $m^3$/ha with SE of 1.754 $m^3$/ha when no uncertainty was incorporated and 92.614 $m^3$/ha with SE of 1.779 $m^3$/ha when uncertainty was incorporated. For practical purposes, these estimates are not meaningfully different. Thus, with respect to the uncertainty of model predictions, the much larger calibration data set sizes for the Hahn (1984) models more than compensated for the smaller $Q^2$ values. However, the Hahn (1984) models use a different mathematical form for the volume models and include site index and stand basal area as additional independent variables. A rigorous comparison of the effects of the different mathematical forms of the models is not possible for this study because site index and stand basal area were not available for the calibration data set. Nevertheless, analyses for a variety of two-, three-, and four-parameter models of different mathematical forms that use only dbh and ht indicated that when adjustment was made for different residual variances, the mathematical forms of the models had little effect on confidence intervals around model predictions. Thus, uncertainty in model predictions resulting from residual variability and uncertainty in parameter estimates associated with the Hahn (1984) models can be expected to have only negligible effects on the precision of large area estimates of mean volume per unit area.

In general, residual variability and uncertainty in model parameter estimates characteristic of the data used to construct the models for this study and characteristic of the Hahn models had minimal effects on the uncertainty of large area estimates of mean volume per unit area and their SEs. Similar results are also reported by Berger et al. (2013) for the Austrian NFI, by Breidenbach et al. (2013) for the Norwegian NFI, and by Ståhl et al. (2013) for the Swedish and Finnish NFIs.

**Table 2. Estimates for mean volume per unit area.**

| Simulation parameters | | Models | | | | | |
|---|---|---|---|---|---|---|---|
| | | Species-specific | | Coniferous/deciduous | | Nonspecific | |
| Calibration data set size | $Q^2$ | Mean | SE | Mean | SE | Mean | SE |
| | | (m³/ha) | | | | | |
| No uncertainty added | | 92.597 | 1.752 | 90.445 | 1.698 | 90.344 | 1.692 |
| 1,000 | 0.99 | 92.597 | 1.753 | 90.442 | 1.699 | 90.343 | 1.692 |
| 1,000 | 0.98 | 92.599 | 1.755 | 90.440 | 1.699 | 90.342 | 1.693 |
| 1,000 | 0.95 | 92.612 | 1.758 | 90.442 | 1.701 | 90.349 | 1.694 |
| 1,000 | 0.90 | 92.654 | 1.766 | 90.469 | 1.704 | 90.385 | 1.697 |
| 1,000 | 0.85 | 92.731 | 1.773 | 90.536 | 1.708 | 90.461 | 1.701 |
| 250 | 0.99 | 92.601 | 1.757 | 90.443 | 1.701 | 90.343 | 1.694 |
| 250 | 0.98 | 92.601 | 1.761 | 90.440 | 1.703 | 90.342 | 1.696 |
| 250 | 0.95 | 92.620 | 1.775 | 90.439 | 1.709 | 90.348 | 1.702 |
| 250 | 0.90 | 92.674 | 1.800 | 90.467 | 1.720 | 90.383 | 1.712 |
| 250 | 0.85 | 92.756 | 1.827 | 90.534 | 1.732 | 90.460 | 1.723 |
| 100 | 0.99 | 92.602 | 1.764 | 90.440 | 1.706 | 90.342 | 1.698 |
| 100 | 0.98 | 92.610 | 1.776 | 90.443 | 1.711 | 90.344 | 1.703 |
| 100 | 0.95 | 92.636 | 1.812 | 90.442 | 1.726 | 90.349 | 1.717 |
| 100 | 0.90 | 92.704 | 1.870 | 90.464 | 1.751 | 90.379 | 1.740 |
| 100 | 0.85 | 92.829 | 1.927 | 90.546 | 1.780 | 90.468 | 1.766 |
| 50 | 0.99 | 92.611 | 1.776 | 90.443 | 1.713 | 90.344 | 1.704 |
| 50 | 0.98 | 92.623 | 1.801 | 90.441 | 1.723 | 90.343 | 1.714 |
| 50 | 0.95 | 92.675 | 1.873 | 90.446 | 1.755 | 90.353 | 1.742 |
| 50 | 0.90 | 92.789 | 1.988 | 90.468 | 1.804 | 90.383 | 1.787 |
| 50 | 0.85 | 92.932 | 2.095 | 90.542 | 1.852 | 90.461 | 1.833 |
| 25 | 0.99 | 92.633 | 1.804 | 90.444 | 1.728 | 90.346 | 1.717 |
| 25 | 0.98 | 92.660 | 1.854 | 90.436 | 1.748 | 90.340 | 1.735 |
| 25 | 0.95 | 92.766 | 2.007 | 90.454 | 1.809 | 90.359 | 1.791 |
| 25 | 0.90 | 92.960 | 2.233 | 90.495 | 1.903 | 90.403 | 1.873 |
| 25 | 0.85 | 93.205 | 2.449 | 90.568 | 1.996 | 90.475 | 1.961 |

### Broader Implications

The results reported in Table 2 can be used as crude but general guidelines for assessing the effects of uncertainty in parameter estimates and residual uncertainty on the volume, biomass, and carbon models reported for tropical applications. For the following examples, models similar in form to Equation 6 were used and the reported quality-of-fit measure is understood to be equivalent to $Q^2$ as used for this study. Brown et al. (1989, Table 2) reported aboveground biomass models for dry tropical forests for which calibration data set sizes were as small as $n = 32$ with $Q^2 = 0.67$, and Litton and Kauffman (2008, Table 1) reported models for Hawaii with $n = 36$ and $Q^2 = 0.96$ and with $n = 20$ and $Q^2 = 0.93$. Certainly considerable caution must be exercised when extrapolating from models constructed for boreal and temperate forests to models for tropical forests. Nevertheless, these examples illustrate that individual tree models are reported and presumably used for which uncertainty in volume model predictions resulting from uncertainty in model parameter estimates and residual uncertainty should not be assumed to be negligible. For cases such as these for which species have small calibration data sets, aggregating over multiple species to form larger calibration data sets may produce greater precision with little or no loss of accuracy (Kamelarczyk 2009, p. 15).

## Conclusion

Four conclusions may be drawn from the study. First, the Monte Carlo approach worked well and was not excessively computationally intensive, although the ln-ln transformation was necessary to incorporate uncertainty in the model parameter estimates. Second, no seriously detrimental consequences relative to precision in large area estimates of volume accrue as a result of ignoring model prediction uncertainty for calibration data set sizes and $Q^2$ values for this study or for the Hahn (1984) models. These results provide a practical justification for the widespread inventory and monitoring practice of ignoring the effects of parameter and residual uncertainty on the uncertainty of volume model predictions, although the justification is subject to sufficiently large calibration data sets and $Q^2$ values. Third, the coniferous/deciduous and nonspecific models, which require considerably less effort to construct and maintain, produced only small deviations in large area means relative to those for the species-specific models and actually produced smaller SEs. This conclusion suggests that for large area estimation purposes, little may be gained by using species-specific models. Fourth, the results of the study suggest that efforts by NFIs and other large-scale inventory and monitoring programs to increase the precision of large area volume estimates should focus on reducing the effects of sampling variability by increasing sample sizes, using more efficient sampling designs, and using remotely sensed data to enhance inferences (McRoberts et al. 2012, 2013).

## Literature Cited


AKAIKE, H. 1974. A new look at the statistical model identification. *IEEE Trans. Automat. Control* 19(6):716–723.

ANDERSON-SPRECHER, R. 1994. Model comparisons and $R^2$. *Am. Statist.* 48(2):113–117.

AVERY, T.E., AND H.E. BURKHART. 2002. *Forest measurements*, 5th ed. McGraw-Hill, New York. 456 p.

BASUKI, T.M., P.E. VAN LAAKE, A.K. SKIDMORE, AND Y.A. HUSSIN. 2009.

Allometric models for estimating above-ground biomass in tropical lowland *Dipterocarp* forests. *For. Ecol. Manage.* 257:1684–1694.

BATES, D.M., AND D.G. WATTS. 1988. *Nonlinear regression analysis and its applications.* John Wiley & Sons, New York. 363 p.

BEETS, P.N., M.O. KIMBERLEY, G.R. OLIVER, S.H. PEARCE, J.D. GRAHAM, AND A. BRANDON. 2012. Allometric equations for estimating carbon stocks in natural forest in New Zealand. *Forests* 3:818–839.

BASKERVILLE, G.L. 1972. Use of logarithmic regression in the estimation of plant biomass. *Can. J. For. Res.* 2(1):49–53.

BERGER, A., T. GSCHWANTNER, R.E. MCROBERTS, AND K. SCHADAUER. 2013. Effects of measurement errors on single tree stem volume estimates for the Austrian National Forest Inventory. *For. Sci.* In press.

BREIDENBACH, J., C. ANTÓN-FERNÁNDEZ, H. PETERSSON, P. ASTRUP, AND R.E. MCROBERTS. 2013. Quantifying the contribution of biomass model errors to the uncertainty of biomass stock and change estimates in Norway. *For. Sci.* In press.

BROWN, S., A.J.R. GILLESPIE, AND A.E. LUGO. 1989. Biomass estimation methods for tropical forests with application to forest inventory data. *For. Sci.* 35(4):881–902.

CHAVE, J., C. ANDALO, S. BROWN, M.A. CAIRNS, J.Q. CHAMBERS, D. EAMUS, H. FÖLSTER, ET AL. 2005. Tree allometry and improved estimation of carbon stocks and balance in tropical forests. *Oecolgia* 145:87–99.

CHAVE, J., R. CONDIT, S. AGUILAR, A. HERNANDEZ, S. LAO, AND R. PEREZ. 2004. Error propagation and scaling in tropical forest biomass estimates. *Philos. Trans. R. Soc. Lond. B* 359:409–420.

CUNIA, T. 1965. Some theory on reliability of volume estimates in a forest inventory sample. *For. Sci.* 11(1):115–128.

CUNIA, T. 1987. On the error of continuous forest inventory estimates. *Can. J. For. Res.* 17:436–441.

GERTNER, G.Z. 1987. Approximating precision in simulation projections: An efficient alternative to Monte Carlo. *For. Sci.* 33(1):230–239.

GERTNER, G.Z. 1990. The sensitivity of measurement error in stand volume estimation. *Can. J. For. Res.* 20:800–804.

GERTNER, G.Z., AND P.J. DZIALOWY. 1984. Effects of measurement error on an individual tree-based growth projection system. *Can. J. For. Res.* 14:311–316.

GERTNER, G., AND M. KÖHL. 1992. An assessment of some nonsampling errors in a national survey using an error budget. *For. Sci.* 38:525–538.

HAHN, J.T. 1984. *Tree biomass and volume equations for the Lakes States.* USDA For. Serv., Res. Pap. NC-250, North Central Forest Experiment Station, St. Paul, MN. 10 p.

HOSMER, D.W., AND S. LEMESHOW. 1989. *Applied logistic regression.* John Wiley & Sons, New York. 307 p.

JENKINS, J.C., D.C. CHOJNACKY, L.S. HEATH, AND R.A. BIRDSEY. 2003. National-scale biomass estimators for United States tree species. *For. Sci.* 49(1):12–35.

KAMELARCZYK, K.B.F. 2009. *Carbon stock assessment and modeling in Zambia: A UN-REDD programme study*. UN-REDD Programme MRV Working Pap. 4, Food and Agriculture Organization of the United Nations, United Nations Development Programme and United Nations Environment Programme. 86 p.

KANGAS, A. 1996. On the bias and variance in tree volume predictions due to model and measurement errors. *Scand. J. For. Res.* 11:281–290.

KEMPTHORNE, O., AND L. FOLKS. 1971. *Probability, statistics, and data analysis.* The Iowa State University Press, Ames, IA. 555 p.

KETTERINGS, Q., R. COE, M. VAN NOORDWIJK, Y. AMBAGAU, AND C.A. PALM. 2001. Reducing uncertainty in the use of allometric biomass equations for predicting above-ground tree biomass in mixed secondary forests. *For. Ecol. Manage.* 146:199–209.

LITTON, C.M., AND J.B. KAUFFMAN. 2008. Allometric models for predicting aboveground biomass in two widespread woody plants in Hawaii. *Biotropica* 40(3):313–320.

MCROBERTS, R.E. 1996. Estimating variation in field crew estimates of site index. *Can. J. For. Res.* 26:560–565.

MCROBERTS, R.E., W.A. BECHTOLD, P.L. PATTERSON, C.T. SCOTT, AND G.A. REAMS. 2005. The enhanced Forest Inventory and Analysis program of the USDA Forest Service: Historical perspective and announcement of statistical documentation. *J. For.* 103(6):304–308.

MCROBERTS, R.E., T. GOBAKKEN, AND E. NÆSSET. 2012. Post-stratified estimation of forest area and growing stock volume using LiDAR-based stratifications. *Remote Sens. Environ.* 125:157–166.

MCROBERTS, R.E., J.T. HAHN, G.J. HEFTY, AND J.R. VAN CLEVE. 1994. Variation in forest inventory field measurements. *Can. J. For. Res.* 24:1766–1770.

MCROBERTS, R.E., AND V.C. LESSARD. 2001. Estimating the uncertainty in diameter growth model predictions and its effects on the uncertainty of annual inventory estimates. P. 70–75 in *Proc. of the Second annual Forest Inventory and Analysis symposium*; October 17–18, 2000; Salt Lake City, UT. USDA For. Serv., Gen. Tech. Rep. RS-47, Southern Research Station, Asheville, NC.

MCROBERTS, R.E., E. NÆSSET, AND T. GOBAKKEN. 2013. Inference for lidar-assisted estimation of forest growing stock volume. *Remote Sens. Environ.* 128:268–275.

MOWRER, H.R. 1991. Estimating components of propagated variance in growth simulation model projections. *Can. J. For. Res.* 21:379–386.

MOWRER, H.T., AND W.E. FRAYER. 1986. Variance propagation in growth and yield projections. *Can. J. For. Res.* 16:1196–1200.

NÁVAR, J. 2009. Allometric equations for tree species and carbon stocks for forests of northwestern Mexico. *For. Ecol. Manage.* 257:427–434.

RATKOWSKY, D.A. 1983. *Nonlinear regression modelling: A unified practical approach.* Dekker, New York. 276 p.

RUBIN, D.B. 1987. *Multiple imputation in non-response surveys.* John Wiley & Sons, New York. 287 p.

SCOTT, C.T. 1981. *Northeastern forest survey revised cubic foot volume equations.* USDA For. Serv., Res. Note NE-304, Northeastern Forest Experiment Station, Newtown Square, PA. 10 p.

STÅHL, G., J. HEIKKINEN, H. PETERSSON, J. REPOLA, AND S. HOLM. 2013. Adapting uncertainty assessments from sample based forest inventories to include the effects of model errors. *For. Sci.* In press.

STÅHL, G., J. HOLM, T.G. GREGOIRE, T. GOBAKKEN, E. NÆSSET, AND R. NELSON. 2011. Model-based inference for biomass estimation in a LiDAR sample survey in Hedmark County, Norway. *Can. J. For. Res.* 41:96–107.

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

WESTFALL, J.A., AND P.L. PATTERSON. 2007. Measurement variability error for estimates of volume change. *Can. J. For. Res.* 37:2201–2210.

WESTFALL, J.A., AND C.T. SCOTT. 2010. Taper models for commercial tree species in the northeastern United States. *For. Sci.* 56(6):515–528.

XIANG, W., S. LIU, X. DENG, A. SHEN, S. LEI, D. TIAN, M. ZHAO, AND C. PENG. 2011. General allometric equations and biomass allocation of *Pinus massoniana* trees on a regional scale in southern China. *Ecol. Res.* 26:697–711.

ZIANIS, D., P. MUUKKONEN, R. MÄKIPÄÄ, AND M. MENCUCCINI. 2005. *Biomass and stem volume equations for tree species in Europe*. Silva Fennica Monogr. 4. The Finnish Society of For. Sci., The Finnish Forest Research Institute. 63 p.