Remote Sensing of Environment 128 (2013) 268–275

Contents lists available at SciVerse ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Inference for lidar-assisted estimation of forest growing stock volume

Ronald E. McRoberts [a,*], Erik Næsset [b], Terje Gobakken [b]

[a] Northern Research Station, U.S. Forest Service, 1992 Folwell Avenue, Saint Paul, MN, USA
[b] Department of Ecology and Natural Resource Management, Norwegian University of Life Sciences, Ås, Norway

ARTICLE INFO

Article history:
Received 3 August 2012
Received in revised form 4 October 2012
Accepted 6 October 2012
Available online 10 November 2012

Keywords:
Nonlinear logistic regression model
stratified estimator
model-assisted estimator
model-based estimator

ABSTRACT

Estimates of growing stock volume are reported by the national forest inventories (NFI) of most countries and may serve as the basis for aboveground biomass and carbon estimates as required by an increasing number of international agreements. The probability-based (design-based) statistical estimators traditionally used by NFIs to calculate estimates are generally unbiased and entail only limited computational complexity. However, these estimators often do not produce sufficiently precise estimates for areas with small sample sizes. Model-based estimators may overcome this disadvantage, but they also may be biased and estimation of variances may be computationally intensive. For a minor region within Hedmark County, Norway, the study objective was to compare estimates of mean forest growing stock volume per unit area obtained using probability- and model-based estimators. Three of the estimators rely to varying degrees on maps that were constructed using a nonlinear logistic regression model, forest inventory data, and lidar data. For model-based estimators, methods for evaluating quality of fit of the models and reducing the computational intensity were also investigated. Three conclusions were drawn: the logistic regression model exhibited no serious lack of fit to the data; estimators enhanced using maps produced greater precision than estimates based on only the plot observations; and third, model-based synthetic estimators benefit from sample sizes for larger areas when applied to smaller subsets of the larger areas.

Published by Elsevier Inc.

1. Introduction

Forest growing stock volume is one of the two variables most commonly assessed by national forest inventories (NFI), the other being forest area. Estimators of growing stock volume also may serve as the basis for estimates of aboveground biomass and carbon which are increasingly required for reporting for international agreements. NFIs traditionally calculate estimates of growing stock volume using probability-based (design-based) simple random sampling, stratified, or model-assisted estimators. However, estimates are generally reported only for large areas such as countries, regions within countries, states, and provinces because of sample size constraints. Model-based or model-dependent approaches (Hansen et al., 1983) rely more heavily on models and auxiliary variables to produce estimates and may produce more precise estimates for small areas for which probability-based estimates are not feasible. In addition, model-based estimators produce estimates are consistent with maps in the sense that they represent the aggregation of map unit predictions. Their primary disadvantages are that they are not necessarily unbiased and they are often computationally intensive.

McRoberts (2010) reported comparisons of approaches to inference for estimates of forest area using models, inventory observations of the categorical forest/non-forest variable, and Landsat imagery. The basic results were that the greater the degree to which the models and their predictions were used, the greater the increase in the precision of estimates. The results of similar comparative studies are not known to have been reported for continuous response variables such as forest growing stock volume and biomass, possibly because auxiliary data useful for predicting below-canopy forest attributes have not generally been available. However, the recent literature includes multiple reports of strong relationships between these attributes and lidar metrics. For example, Næsset (2002) reported that 80–93% of the variability in field measured forest stem volume could be explained by models that use lidar metrics, and Næsset and Gobakken (2008) reported that 88% of the variability in aboveground biomass could be explained by models using lidar metrics. Similar results have also been reported for multiple other studies including Frazer et al. (2011), Li et al. (2008), and Zhao et al. (2009). Thus, a study of different approaches to inference for parameters related to below-canopy attributes such as growing stock volume using models based on lidar metrics is now feasible.

Statistical inference requires expression of an estimate in probabilistic terms; for this study the expressions took the form of approximate 95% confidence intervals calculated as $\hat{\mu} \pm 2 \cdot \sqrt{\hat{\mathrm{Var}}\left(\hat{\mu}\right)}$. Thus, the study focused on estimation of means and their variances for purposes of facilitating construction of confidence intervals. The objective of the study was to compare estimates of means and their variances for forest growing stock volume per unit area (VOL) (m^3/ha) obtained using four statistical estimators: (1) a simple random sampling estimator, (2) a

* Corresponding author. Tel.: +1 651 649 5174; fax: +1 651 649 5285.
E-mail address: rmcroberts@fs.fed.us (R.E. McRoberts).

0034-4257/$ – see front matter. Published by Elsevier Inc.
http://dx.doi.org/10.1016/j.rse.2012.10.007

stratified estimator, (3) a model-assisted regression estimator, and (4) a model-based estimator. The latter three estimators relied on a map of growing stock volume constructed using a nonlinear logistic regression model, NFI plot data, and lidar metrics.

2. Data

The study area was mostly in the municipalities of Åmot and Stor-Elvdal in Hedmark County, Norway (Fig. 1). Airborne lidar data were acquired between 15 July 2006 and 12 September 2006 from a height of approximately 1700 m with average aircraft speed of 75 ms^{-1}. The pulse repetition frequency was 50 kHz, the scan frequency was 31 Hz, the maximum scan angle was 16°, which corresponded to an average swath width of approximately 975 m, the mean footprint diameter was approximately 50 cm, and the average point density was 0.7 pulses m^{-2}. Data for only single echoes or the first of multiple echoes were used. For each plot and population unit, height distributions were estimated for first echoes from tree canopies, i.e., heights greater than 2 m. Echoes with heights less than 2 m were considered to have been reflected from non-tree objects such as shrubs, grass, or the ground. For each plot and population unit, heights corresponding to the 10th, 20th, …, and 100th percentiles of the distributions were calculated and denoted h_1, h_2, …, and h_{10}, respectively. Canopy densities were calculated as the proportions of echoes with heights greater than 0%, 10%, …, and 90% of the range between 2 m above ground and the 95th height percentile and were denoted d_0, d_1, …, and d_9, respectively (Gobakken & Næsset, 2008).

Field measurements were obtained from 250-m^2, circular Norwegian NFI field plots located at the intersections of a 3-km × 3-km grid (Tomter et al., 2010). On each plot, all trees with diameters at-breast-height (dbh, 1.3 m) of at least 5 cm were callipered. Tree heights were measured on an average of 10 sample trees per plot selected with probability proportional to stem basal area, and heights for the remaining trees were predicted using height-dbh models (Fitje & Vestjordet, 1977; Vestjordet, 1968). The volume of each sample tree was estimated using species-specific volume models with dbh and either measured height or predicted height as independent variables (Braastad, 1966; Brantseg, 1967; Vestjordet, 1967). The ratio of the mean volume estimate for trees with predicted heights and the mean volume estimate for trees with measured heights was used to adjust the former volume estimates. The estimate of total plot volume, VOL, was calculated as the sum of volume estimates for individual trees. Although the plot-level values of VOL were estimates, the uncertainty of the estimates was considered negligible relative to plot-to-plot variability and was ignored. A variogram analysis indicated no meaningful spatial correlation among plot VOL observations. Differential Global Navigation Satellite Systems (GPS and the Russian GLONASS) were used to determine the position of the center of each plot.

To minimize the effects of forest change between the plot observation dates and the 2006 date of the lidar acquisition, only the 145 plots measured between 2005 and 2007 were used for this study. Thus, the study area was defined as the geographic area represented by the portion of the Latin Square sampling design used by the Norwegian NFI inventoried between 2005 and 2007 (Fig. 1). The study area includes 1259 km^2 and features altitudinal variations ranging from 204 m to 1134 m above sea level (asl) with a mean of 570 m asl. For analytical purposes, the study area was tessellated into square 250-m^2 cells that served as population units. To assess the effects on estimates of study areas of different sizes, a second study area with the same center but of approximately half the size and including only 69 NFI plots was also selected. The dominant tree species are Norway spruce (*Picea abies* (L.) Karst.) and Scots pine (*Pinus sylvestris* L.). Mean and maximum plot volumes were 0.117 m^3 and 1.985 m^3, respectively.

3. Methods

All analyses were based on three underlying assumptions: (1) a finite population consisting of N units in the form of the square,

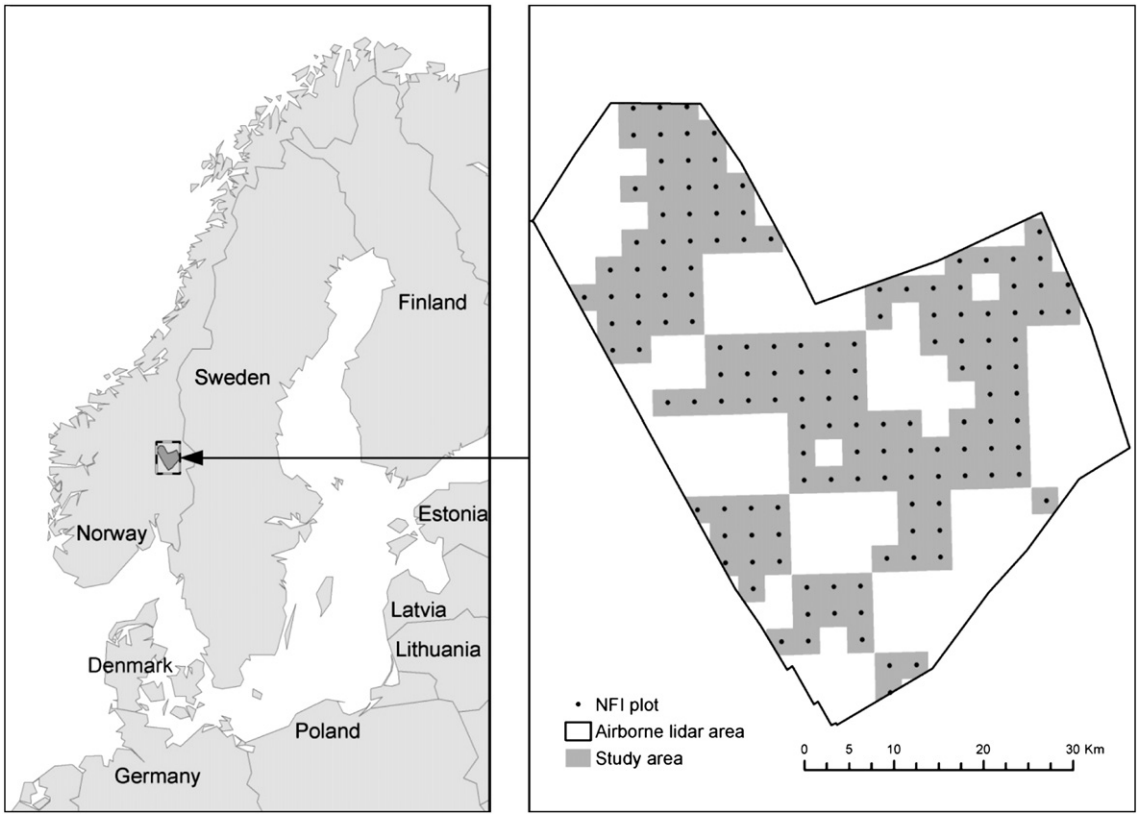

Fig. 1. Study area and sampling design.

250-m^2 cells, (2) an equal probability sample of n population units in the form of the NFI plots, and (3) availability of auxiliary data in the form of lidar metrics for all plots and population units. In the following sections, the terms *cell* and *population unit* are considered synonymous.

3.1. Nonlinear logistic regression model

A nonlinear logistic regression model was used to describe the relationship between VOL for NFI plots and associated lidar metrics. The model has the mathematical form,

$$y_i = f(\mathbf{X}_i;\boldsymbol{\beta}) = \frac{\beta_{J+2}}{1 + \exp\left(\beta_{J+1} + \sum_{j=1}^{J}\beta_j x_{ij}\right)} + \varepsilon_i, \quad (1)$$

where i indexes plots or population units, x_{ij} is the jth lidar metric, the βs are parameters to be estimated, and ε_i is a residual error term. An advantage of the logistic model expressed by Eq. (1) over a linear model is that all predictions are non-negative and are constrained by the lower horizontal asymptote of $\hat{y} = 0$ and the upper horizontal asymptote of $\hat{y} = \hat{\beta}_{J+2}$ which is estimated from the sample data. This logistic regression model should not be confused with the binomial or multinomial logistic regression models that are often used with categorical data (Agresti, 2007).

3.2. Probability-based estimators

Properties of probability-based estimators derive from the probabilities of selection of population units into the sample, thus these estimators are characterized as *probability-based* (Hansen et al., 1983), although they are also characterized as *design-based*. Probability-based inference is based on three assumptions: (1) population units are selected for the sample using a probability-based randomization scheme; (2) the probability of selection for each population unit is positive and known; and (3) the observation of the response variable for each population unit is a fixed value. Estimators are derived to correspond to sampling designs and typically are unbiased or nearly unbiased, meaning that the expectation of estimates obtained with the estimators over all samples that could be obtained with the sampling design is the true value of the population parameter.

3.2.1. Simple random sampling estimator

The simplest approach to probability-based inference is to use the familiar simple random sampling (SRS) estimators for means and their variances,

$$\hat{\mu}_{SRS} = \frac{1}{n}\sum_{i=1}^{n} y_i \quad (2)$$

and

$$\hat{Var}\left(\hat{\mu}_{SRS}\right) = \frac{\sum_{i=1}^{n}\left(y_i - \hat{\mu}_{SRS}\right)^2}{n(n-1)}, \quad (3)$$

where i indexes the n sample observations and y_i is the observation for the ith population unit selected for the sample. The primary advantages of the SRS estimators are that they are intuitive, simple, and unbiased when used with an SRS design. However, even though the estimator is unbiased, the estimate obtained with any particular sample may still deviate substantially from the true value. Nevertheless, estimates obtained with the SRS estimator with large samples are often used as a standard for comparison for estimates obtained with other estimators. The disadvantage of the SRS estimators is that variances are frequently large, particularly for small sample sizes. Although $\hat{Var}(\hat{\mu}_{SRS})$ from Eq. (3) may be biased when used with systematic sampling, it is usually conservative in the sense that it over-estimates the true variance (Särndal et al., 1992, p. 83). For this study, finite population correction factors were ignored because of the small sampling intensity of one 250-m^2 plot per 9 km^2 of land area.

3.2.2. Stratified estimator

For areal estimation, the stratified (STR) estimator can use a map to decrease the variances of estimates and thereby enhance the inference by shortening the confidence interval. However, because the validity of the inference is still based on probabilities of selection of population units into the sample, the stratified estimator is characterized as probability-based. The essence of stratified estimation is to assign population units to groups or strata, calculate within stratum sample plot means and variances, and then calculate the population estimate as a weighted average of the within stratum estimates where the weights are proportional to the stratum sizes. Stratified estimation requires accomplishment of two tasks: (1) calculation of the strata weights as the relative proportions of the population area corresponding to strata and (2) assignment of each sample unit to a single stratum. The first task is accomplished by calculating the strata weights as proportions of population units in strata. The second task is accomplished for this study by assigning the NFI plots to strata on the basis of the stratum assignments of the population units containing the plot centers. The Forest Inventory and Analysis program of the U.S. Forest Service, which conducts the NFI of the United States of America, has devoted considerable effort to investigating stratified estimation using remotely sensed data (Hansen & Wendt, 2000; Liknes et al., 2004, 2009; McRoberts et al., 2002a, 2002b, 2006, 2012; Nelson et al., 2005; Westfall et al., 2011).

Stratified estimates of means and variances are calculated using estimators provided by Cochran (1977):

$$\hat{\mu}_{STR} = \sum_{h=1}^{H} w_h \hat{\mu}_h, \quad (4)$$

and

$$\hat{Var}\left(\hat{\mu}_{STR}\right) = \sum_{h=1}^{H} w_h^2 \frac{\hat{\sigma}_h^2}{n_h}, \quad (5a)$$

where

$$\hat{\mu}_h = \frac{1}{n_h}\sum_{i=1}^{n_h} y_{hi},$$

$$\hat{\sigma}_h^2 = \frac{1}{n_h - 1}\sum_{i=1}^{n_h}\left(y_{hi} - \hat{\mu}_h\right)^2,$$

h = 1,…H denotes strata; y_{hi} is the ith sample observation in the hth stratum; w_h is the weight for the hth stratum; n_h is the number of plots assigned to the hth stratum; $\hat{\mu}_h$ and $\hat{\sigma}_h^2$ are the sample estimates of the within stratum mean and variance, respectively; and STR denotes the stratified estimator.

NFIs often use permanent plots whose locations are based on systematic grids or tessellations and use sampling intensities that are constant over large geographic areas, if not the entire population. In such cases, even though stratified sampling is not possible, increase in precision may still be achieved by using stratified estimation subsequent to the sampling, a technique characterized as post-sampling stratification or simply post-stratification. Cochran (1977, p. 135)

provides a modified stratified estimator of the variance for use with post-stratification and the resulting random within-strata sample sizes,

$$\hat{\mathrm{Var}}\left(\hat{\mu}_{STR}\right) = \sum_{h=1}^{H} w_h \frac{n_h}{n}\frac{\hat{\sigma}_h^2}{n_h} + \frac{1}{n}\sum_{h=1}^{H}(1-w_h)\frac{n_h}{n}\frac{\hat{\sigma}_h^2}{n_h}, \tag{5b}$$

where n is the total sample size over all strata. For large sample sizes, particularly when the strata weights are nearly equal to the proportions of plots assigned to strata, Eqs. (5a) and (5b) produce similar estimates.

The utility of a stratification for increasing precision is often evaluated using relative efficiency (RE) calculated as,

$$RE = \frac{\hat{\mathrm{Var}}\left(\hat{\mu}_{SRS}\right)}{\hat{\mathrm{Var}}\left(\hat{\mu}_{STR}\right)}, \tag{6}$$

where values of RE greater than 1.0 indicate increasing effectiveness of the stratification.

3.2.3. Model-assisted estimator

Model-assisted estimators use models based on auxiliary data to enhance inferences but rely on the probability sample for validity. For this study, the model-assisted regression (MA) estimators of means and variances are provided by Särndal et al. (1992):

$$\hat{\mu}_{MA} = \frac{1}{N}\sum_{i=1}^{N}\hat{y}_i - \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i), \tag{7}$$

where N is the population size, $\hat{y}_i$ is obtained from Eq. (1) using the model parameter estimates and $\varepsilon_i = 0$. The first term in Eq. (7), $\frac{1}{N}\sum_{i=1}^{N}\hat{y}_i$, is simply the mean of the model predictions, $\hat{y}_i$, for all population units, and the second term, $\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)$, is an estimate of bias calculated over the sample units and compensates for systematic model prediction errors. The variance can be approximated as,

$$\hat{\mathrm{Var}}(\hat{\mu}_{MA}) = \frac{1}{n(n-1)}\sum_{i=1}^{n}\left(\varepsilon_i - \bar{\varepsilon}\right)^2, \tag{8}$$

where $\varepsilon_i = \hat{y}_i - y_i$.

The primary advantage of MA estimators is that they capitalize on the relationship between the sample observations and their model predictions to reduce the variance of the estimate of the population mean. In this regard, they are potentially preferable to the STR estimator because they use the model predictions for individual population units, whereas the STR estimator aggregates the population units and their predictions into strata.

3.3. Model-based inference

The assumptions underlying model-based (MOD) inference differ considerably from the assumptions underlying probability-based inference. First, the observation for a population unit is a random variable whose value is considered a realization from a distribution of possible values, rather than a fixed value as is the case for probability-based inference. Second, the basis for a MOD inference is the model, not the sample as is the case for probability-based inference. Randomization for MOD inference enters through the random realizations from the distributions for population units, whereas randomization for probability-based inference enters through the random selection of population units into the sample.

The mean and standard deviation of the distribution of Y for the ith population unit are denoted μ_i and σ_i, respectively. The mean is estimated as $\hat{\mu}_i = \hat{y}_i$ from Eq. (1), and σ_i is estimated as the standard deviation of residual deviations between observations and $\hat{\mu}_i$. The MOD estimator of the population mean, μ_{MOD}, is based on the set of estimates, $\{\hat{\mu}_i, i = 1, 2, \ldots, N\}$, of the means for individual population units and is expressed as,

$$\hat{\mu}_{MOD} = \frac{1}{N}\sum_{i=1}^{N}\hat{\mu}_i = \frac{1}{N}\sum_{i=1}^{N} f\left(\mathbf{X}_i;\hat{\boldsymbol{\beta}}\right), \tag{9}$$

where $f\left(\mathbf{X}_i;\hat{\boldsymbol{\beta}}\right)$ is the nonlinear logistic regression model prediction from Eq. (1) evaluated at the parameter estimates, $\hat{\boldsymbol{\beta}}$. The corresponding variance estimator is,

$$\hat{\mathrm{Var}}\left(\hat{\mu}_{MOD}\right) = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\hat{\mathrm{Cov}}\left(\hat{\mu}_i, \hat{\mu}_j\right). \tag{10}$$

The covariances required for Eq. (10) may be estimated as,

$$\hat{\mathrm{Cov}}\left(\hat{\boldsymbol{\mu}}_i, \hat{\boldsymbol{\mu}}_j\right) = Z'_i \hat{V}_{\hat{\beta}} Z_j,$$

where $z_{ij} = \frac{\partial f\left(x_i;\hat{\beta}\right)}{\partial \beta_j}$ and $\hat{V}_{\hat{\beta}}$ is the estimated parameter covariance matrix (McRoberts, 2006). The latter matrix can be approximated as $\hat{V}_{\hat{\beta}} = \left(Z'WZ\right)^{-1}$ where, in the absence of spatial correlation among observations of the response variable, W is a diagonal matrix with $w_{ii} = \hat{\sigma}_i^{-2}$.

3.4. Analyses

3.4.1. Nonlinear logistic regression model

The nonlinear logistic regression model was fit to the data for the entire study area using iterative weighted least squares techniques to accommodate heteroskedasticity. A subset of the lidar metrics was selected using an iterative, stepwise approach based on model efficiency (Vanclay & Skovsgaard, 1997) calculated as $Q^2 = 1 - \frac{SS_{res}}{SS_{mean}}$, where

$$SS_{res} = \sum_{i=1}^{n}(\hat{y}_i - y_i)^2$$

and

$$SS_{mean} = \sum_{i=1}^{n}(y_i - \bar{y})^2.$$

With this approach, the single lidar metric with the greatest value of Q^2 was first selected. Next, all combinations of the first metric selected with each of the remaining metrics were evaluated, and the combination with the greatest value of Q^2 was selected. This procedure of selecting the lidar metric that best augmented the previously selected metrics continued until a Q^2 value was obtained for each possible number of metrics. Selection of the optimal subset of metrics was based on a subjective assessment of the combination of metrics beyond which only negligible increases in Q^2 were found.

MOD estimators are not necessarily unbiased, or even nearly unbiased, as is often the case for probability-based estimators. Because unbiasedness for these estimators is closely linked to correct model specification, the quality of fit of the model to the data should be assessed. If the model is correctly specified, graphs of observations versus model predictions for a continuous response variable should feature points that lie along the 1:1 line with intercept 0 and slope 1. In addition to graphing observations against predictions, a three-step approach was used to assess quality of fit: (1) all plot

observation/model prediction pairs, $(y_i, \hat{y}_i)$, were ordered with respect to $\hat{y}_i$; (2) the ordered pairs were grouped into categories of equal numbers of pairs, and the group means of the plot observations and the corresponding model predictions were calculated; and (3) a graph of the observation means versus the model prediction means was constructed (Hosmer & Lemeshow, 1989). If the model is correctly specified, a graph of means of observations against means of predictions should again lie along the 1:1 line.

3.4.2. Stratified estimator

For the STR estimator, strata were constructed using the logistic model predictions, $\hat{y}_i$, which were standardized and scaled to the [0,100] interval:

$$s_i = 100 \cdot \frac{\hat{y}_i - \hat{y}_{min}}{\hat{y}_{max} - \hat{y}_{min}}, \quad (11)$$

where $\hat{y}_{min}$ and $\hat{y}_{max}$ are the minimum and maximum predictions, respectively, over all population units. Each unit was then assigned to one of the 101 standardized classes [0,0], (0,1], …, (99,100]. The classes were aggregated into strata subject to two constraints: (1) classes aggregated together must be contiguous and (2) each stratum must include at least 10 plots to ensure the accuracy of the within-strata variance estimates (McRoberts et al., 2012; Westfall et al., 2011). Stratifications featuring two to six strata were constructed by selecting strata boundaries to minimize $\hat{Var}(\hat{\mu}_{STR})$ calculated using Eq. (5b) or, equivalently, to maximize RE calculated using Eq. (6). Although the model was applied in the same manner and with the same parameter estimates, separate stratifications were constructed for the entire and the half study areas.

Legitimate concerns may be raised regarding violations of assumptions resulting from using stratifications based on the sample observations to stratify the same sample units from which the observations are obtained. Breidt and Opsomer (2008) coined the term *endogenous post-stratification* to describe this approach and demonstrated that when strata are constructed by dividing the range of predictions obtained from a linear model calibrated using the response variable observations, the detrimental effects of violating the stratification assumptions are negligible, even for small sample sizes. Dahlke et al. (2013) also used similar data to show that when a non-parametric prediction procedure was used, estimates of standard errors were asymptotically similar to those obtained using the familiar Horvitz–Thompson estimator (Horvitz & Thompson, 1952). These findings permit the use of Eqs. (5a) and (5b), even when the model does not fit the data as well as desired.

3.4.3. Model-assisted estimator

The model was applied for the entire and half study areas using the same model parameter estimates. Models whose parameters are estimated using data for areas larger than the areas to which they are applied are characterized as *synthetic* estimators (Särndal et al., 1992) (Section 4.3). When estimates were calculated using Eqs. (7) and (8), only data for plots with centers in the two study areas were used.

3.4.4. Model-based estimator

A disadvantage of MOD estimators is that calculation of $\hat{Var}(\hat{\mu}_{Mod})$ using Eq. (10) is computationally intensive because of the double summation and the large number of population units for even relatively small study areas. For example, the current study area consists of more than 5×10^6 population units, which means that the number of covariance calculations necessary for Eq. (10) is on the order of 10^{13}. However, McRoberts (2010) noted that Eq. (10) is just a two-dimensional mean over all units in the population and that $\hat{Var}(\hat{\mu}_{MOD})$ can be approximated by sampling from the population. Using only population units located at the intersections of an equally-spaced, two-dimensional, perpendicular grid superimposed on the study area,

$$\hat{Var}(\hat{\mu}_{MOD}) \approx \frac{1}{n_{grid}^2} \sum_{i=1}^{n_{grid}} \sum_{j=1}^{n_{grid}} Z'_i V_{\hat{\beta}} Z_j, \quad (12)$$

where n_{grid} is the number of grid lines in each dimension. For a grid width of n_p, defined as the number of population units separating the grid lines, the computational intensity necessary to calculate $\hat{Var}(\hat{\mu}_{MOD})$ is reduced by a factor of approximately n_p^2. McRoberts (2010) also reported that for study areas much smaller than used for this study, the effects of this approach on $\hat{Var}(\hat{\mu}_{MOD})$ were negligible for grid widths as great as $n_p = 10$. For this study, grid widths of $n_p = 1$ to $n_p = 100$ were used, and the effects on $\hat{Var}(\hat{\mu}_{MOD})$ were compared.

4. Results and discussion

4.1. Nonlinear logistic regression model

Analyses of combinations of lidar metrics used as independent variables for predicting VOL with the nonlinear logistic regression model indicated that little improvement beyond $Q^2 = 0.84$ was achieved for more than the four variables: h_1, h_{10}, d_0, and d_9. Visual inspections of the graph of VOL observations versus model predictions (Fig. 2) and the graph of the means of observations versus means of model predictions for the ordered groups (Fig. 3) indicated no obvious model lack of fit. A simple linear model was fit to the VOL observations as the dependent variable and the VOL predictions as the independent variable, and a second simple linear model was fit to the means of the VOL observations as the dependent variable and the means of the VOL predictions as the independent variable. For both models, an F-test comparing estimates of the intercepts and slopes jointly to (0,1) indicated no statistically significant differences at $\alpha = 0.05$. Although this test does not take into account the uncertainty in the model predictions, doing so would produce an even less significant result.

4.2. Probability-based estimators

4.2.1. Simple random sampling estimator

The SRS estimators yielded $\hat{\mu}_{SRS} = 74.26\,m^3/ha$ with $SE(\hat{\mu}_{SRS}) = 7.45\,m^3/ha$ for the entire study area and $\hat{\mu}_{SRS} = 82.46\,m^3/ha$ with $SE(\hat{\mu}_{SRS}) = 11.96\,m^3/ha$ for the half study area (Table 1). Because the SRS estimator is unbiased, SRS estimates obtained using large samples

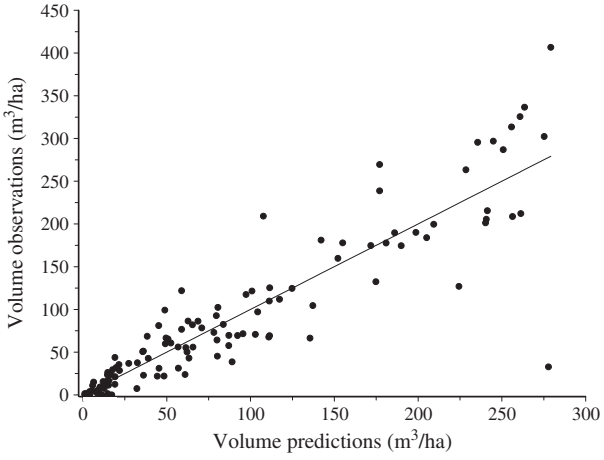

Fig. 2. Growing stock volume (VOL) observations versus nonlinear logistic model predictions for sample data used to construct the model.

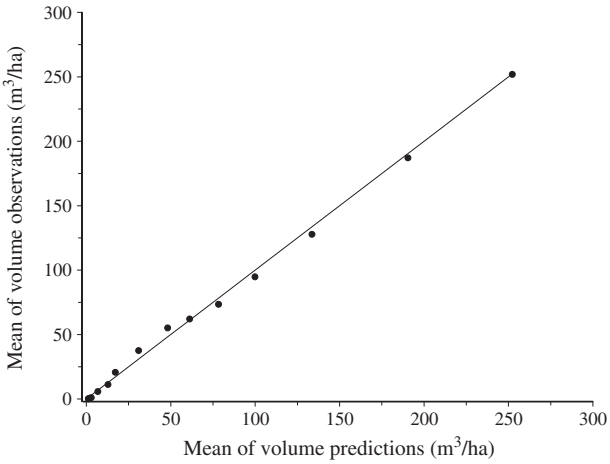

Fig. 3. Means of growing stock volume (VOL) observations versus means of nonlinear logistic model predictions for the sample data used to construct the model.

are often used as a standard for comparison for estimates obtained with other estimators.

4.2.2. Stratified estimator

The STR estimates of the mean for the entire study area ranged from $\hat{\mu}_{STR} = 79.10\,m^3/ha$ to $\hat{\mu}_{STR} = 81.16\,m^3/ha$, depending on the number of strata, and estimates for the half study area ranged from $\hat{\mu}_{STR} = 88.58\,m^3/ha$ to $\hat{\mu}_{STR} = 94.12\,m^3/ha$, again depending on the number of strata (Table 1). For the entire and half study areas, the stratified estimates were approximately 8% and 11% larger, respectively, than the SRS estimates. These differences are attributed to differences between strata weights and the proportions of plots assigned to strata.

Standard errors ranged approximately from 4.3 m^3/ha to 5.1 m^3/ha for the entire study area and approximately from 7.7 m^3/ha to 8.8 m^3/ha for the half study area, depending on the number of strata. As proportions of estimates of the mean, the standard errors were in the range 0.054–0.63 for the entire study area and in the range 0.084–0.096 for the half study area. Selection of strata boundaries to minimize $\hat{Var}(\hat{\mu}_{STR})$ produced RE of approximately 3.0 for four or more strata for the entire study area and approximately 2.4 for the half study area. However, little additional increase in RE was realized for more than four strata (Table 1). In this context, RE can be interpreted as the multiplicative factor by which the sample size would have to be increased to achieve the same reduction in variance using the SRS estimators as was realized with the STR estimators. For forest inventory applications, the cost of tripling the sampling size would be prohibitive; thus, RE = 3.0 is of considerable economic importance, subject to the cost of the lidar data. Of importance, although no lack of fit of the nonlinear regression model to the data was indicated, lack of fit does not induce bias into the STR estimator but rather just limits the degree to which stratifications reduce variance estimates. Similarly, assignment of plots to incorrect strata does not induce bias as long the assignment is consistent, but again just inhibits increases in RE.

The stratifications had the desired effect of aggregating plots into strata whose estimates of means and variances or standard errors were clearly distinguishable (Table 2). The greater SE for five and six strata than for four strata for the half study area can be attributed to two factors. First, the requirement of 10 plots per stratum limited the possible stratifications. Second, larger numbers of strata result in smaller within-strata sample sizes and larger variances of within-strata means.

Table 1
Estimates of mean growing stock volume per unit area (VOL) (m^3/ha).

Estimator	Entire study area (n = 145)		Half study area (n = 69)	
	$\hat{\mu}$ (m^3/ha)	SE($\hat{\mu}$) (m^3/ha)	$\hat{\mu}$ (m^3/ha)	SE($\hat{\mu}$) (m^3/ha)
Simple random sampling (SRS)	74.25	7.45	82.46	11.96
Stratified (STR)				
2 strata	80.49	4.99	94.12	8.34
3 strata	81.16	4.34	91.31	7.40
4 strata	80.04	4.27	91.09	7.31
5 strata	79.17	4.19	91.17	7.33
6 strata	79.10	4.18	88.58	8.16
Model assisted (MA)	81.87	2.95	90.71	5.04
Model-based (MOD)[a]				
$n_p = 1$	81.85	3.50	88.38	3.95
$n_p = 5$	81.85	3.50	88.38	3.94
$n_p = 10$	81.85	3.52	88.38	3.98
$n_p = 25$	81.85	3.53	88.38	3.98
$n_p = 50$	81.85	3.57	88.38	3.96
$n_p = 100$	81.85	3.60	88.38	4.08

[a] n_p is the spacing in population units between grid lines used to calculate $SE(\hat{\mu}_{Mod})$; $\hat{\mu}_{Mod}$ is calculated using all population units.

4.2.3. Model-assisted estimator

The MA estimates of the mean for both the entire and half study areas were very similar to the STR estimates but larger than the SRS estimates (Table 1). Bias estimates for the MA estimator were small, −0.03 m^3/ha or less than 0.1% of the mean for the entire study area and −2.32 m^3/ha or less than 2.6% of the mean for half the study area. Standard errors for estimates of the means were also relatively small, 2.95 m^3/ha for the entire study area and 5.04 m^3/ha for the half study area (Table 1). As percentages of estimates of the means, the standard errors were 3.6% and 5.6% for the entire and half study areas, respectively. Small bias estimates, which reflect the means of differences between VOL observations and model predictions, and small variance estimates can be attributed to the good quality of fit of the nonlinear logistic regression model to the data. Values of RE calculated using Eq. (6) were 6.4 and 5.6 for the entire and half study areas, respectively. These REs indicate that to achieve the same precision with the SRS estimator as was achieved with the MA estimator, the sample size would have to be increased from n = 145 to n = 928 for the entire study area and from n = 69 to n = 386 for the half study area, both of which would be prohibitively expensive.

4.3. Model-based estimator

As noted in Section 4.1, neither visual analyses of graphs of VOL observations versus predictions nor statistical tests of significance for the estimates of coefficients for linear models fit to these data indicated any lack of fit of the nonlinear logistic regression model (Table 1). In addition, the estimate of bias for the MA estimator was negligible relative to the estimate of mean VOL (Section 4.2.3). Finally, $\hat{\mu}_{MOD} = 81.85\,m^3/ha$ was within approximately one SRS standard error of $\hat{\mu}_{SRS} = 74.25\,m^3/ha$. This evidence suggests that any lack of fit of the nonlinear logistic model to the VOL data was insufficient to induce serious bias into the MOD estimator of the mean. However, a degree of caution must be exercised; use of the same data to assess

Table 2
Stratified estimates for the entire study area.

Stratum	Stratum weight	Number of plots	Proportion of plots	Growing stock volume (VOL)	
				Mean (m^3/ha)	SE (m^3/ha)
1	0.34	51	0.35	5.05	1.30
2	0.23	29	0.20	51.79	7.38
3	0.26	45	0.31	96.45	9.98
4	0.17	20	0.14	233.40	18.23
Total	1.00	145	1.00	77.34	4.37

lack of fit as was used to estimate model parameters may lead to optimistic results.

The MOD estimates of the mean for the entire and half study areas were similar to the STR and MA estimates but larger than the SRS estimates. For $n_p = 1$, the standard errors for the estimates of the means were approximately 3.50 m^3/ha for the entire study area and 3.95 m^3/ha for the half study area. As percentages of estimates of the means, the standard errors were less than 4.5% for both the entire and the half study areas. Any detrimental effects of using two-dimensional, equally-spaced, perpendicular grids to calculate $\hat{\mathrm{Var}}\left(\hat{\mu}_{\mathrm{MOD}}\right)$ were negligible for grid widths of up to $n_p = 100$ population units (Table 1).

In contrast to the probability-based estimators whose standard errors are largely dependent on sample size and increased substantially for the half study area ($n = 69$) compared to the entire study area ($n = 145$), the MOD standard error for the half study area increased only slightly relative to that for the entire study area. MOD standard errors depend on the mathematical form of the model, the model parameter estimates, the covariance matrix for the model parameter estimates, and values of the independent variables. The elements of the model parameter covariance matrix, in turn, depend on the same factors plus residual variability and sample size. If the relationship between the response variable and the independent variables is stationary for the entire study area, then the model parameter estimates and the parameter covariance matrix obtained for the entire study area may be applied to any subset of the entire study area without adjustment. In such cases, the estimator is characterized as *synthetic* (Särndal et al., 1992) with the desirable feature that the positive effects of the larger sample size for the entire study area accrue to standard errors for any subset of that study area.

Because of the effects on variance estimators of different assumptions underlying probability-based and model-based inference, formal and rigorous comparisons of variance estimates obtained with the two kinds of estimators are generally not considered possible.

4.4. Summary

The STR, MA, and MOD estimates were all greater than the SRS estimates. The differences are attributed to differences between the distribution of VOL observations in the sample and the distribution of VOL predictions for the population. Differences of the magnitudes observed can be attributed to model lack of fit, for which there is no evidence for this study, or to the effects of random sampling. Because differences among the estimates for the probability-based estimators were either not statistically significant or only marginally statistically significant, and because differences between the MOD and SRS estimates were also small, the differences for this study are attributed to the effects of random sampling.

For the probability-based estimators, the benefits of greater use of the model predictions are reflected in standard errors for the MA estimator being smaller than for the STR estimator, which, in turn, were smaller than for the SRS estimator. For the entire study area, an informal comparison indicated that the standard error for the MA estimator was smaller than the SE for the MOD estimator, but for the half study area, the relationship was reversed. This result illustrates an important difference between the two estimators. Whereas probability-based estimators suffer from the detrimental effects of smaller sample sizes for smaller areas, synthetic MOD estimators may capitalize on the larger sample sizes for the larger study areas for which the model is correctly specified, even when applied to smaller subsets of the larger study areas. Thus, whereas variances for probability-based estimators increase for smaller study areas with smaller sample sizes, variances for synthetic MOD estimators may remain nearly constant. However, the price to be paid for this advantage is that the relationship between response and independent variables must be stationary, and the model must be correctly specified.

5. Conclusions

Three conclusions were drawn from this study. First, the nonlinear logistic regression model adequately described the relationship between the plot-level growing stock volume observations and the lidar metrics. Second, the three estimators that used the map in the form of the nonlinear logistic model predictions produced smaller variance estimates than the simple random sampling estimator. Thus, use of the auxiliary lidar data enhanced inferences with greater use resulting in smaller variance estimates. Third, unlike probability-based estimators, synthetic model-based estimators can capitalize on the beneficial effects of larger samples sizes for larger study areas when the model is applied to smaller subsets of the study areas.

References

Agresti, A. (2007). *An introduction to categorical data analysis.* Hoboken, NJ: Wiley-Interscience.

Braastad, H. (1966). Volume tables for birch. *Meddelser norske SkogforsVes.*, *21*, 265–365 (In Norwegian with English summary).

Brantseg, A. (1967). Volume functions and tables for Scots pine. South Norway. *Meddelser norske SkogforsVes*, *22*, 689–739 (In Norwegian with English summary).

Breidt, F. J., & Opsomer, J. D. (2008). Endogenous post-stratification in surveys: Classifying with a sample-fitted model. *The Annals of Statistics*, *36*(1), 403–427.

Cochran, W. G. (1977). *Sampling techniques* (3rd edition). New York: Wiley.

Dahlke, M., Breidt, F. J., Opsomer, J., & Van Keilegom, I. (2013). Nonparametric endogenous post-stratification. *Statistica Sinica*, *23*(1). http://dx.doi.org/10.5705/ss.2011.272.

Fitje, A., & Vestjordet, E. (1977). Stand height curves and new tariff tables for Norway spruce. *Meddelser norske SkogforsVes.*, *34*, 23–62 (In Norwegian with English Summary.).

Frazer, G. W., Magnussen, S., Wulder, M. A., & Niemann, K. O. (2011). Simulated impact of sample plot size and co-registration error on the accuracy and uncertainty of LiDAR-derived estimates of forest stand biomass. *Remote Sensing of Environment*, *115*, 636–649.

Gobakken, T., & Næsset, E. (2008). Assessing effects of laser point density, ground sampling intensity, and field plot sample size on biophysical stand properties derived from airborne laser scanner data. *Canadian Journal of Forest Research*, *38*, 1095–1109.

Hansen, M. H., Madow, W. G., & Tepping, B. J. (1983). An evaluation of model-dependent and probability-sampling inferences in sample surveys. *Journal of the American Statistical Association*, *78*, 776–793.

Hansen, M. H., & Wendt, D. G. (2000). Using classified Landsat Thematic Mapper data for stratification in a statewide forest inventory. In R. E. McRoberts, G. A. Reams, & P. C. Van Deusen (Eds.), *Proceedings of the first annual forest inventory and analysis symposium. 2–3 November 1999, San Antonio, Texas. U.S. For. Serv. Gen. Tech. Rep. NC-213.* (pp. 20–27) (Available at: http://fia.fs.fed.us/symposium/proceedings/, Last accessed: 20 February 2012)

Horvitz, D. G., & Thompson, D. J. (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association*, *47*, 663–685.

Hosmer, D. W., & Lemeshow, S. (1989). *Applied logistic regression.* New York: John Wiley & Sons, Inc.

Li, Y., Andersen, H. -E., & McGaughey, R. (2008). A comparison of statistical methods for estimating forest biomass from light detection and ranging data. *Western Journal of Applied Forestry*, *23*(4), 223–231.

Liknes, G. C., Nelson, M. D., Gormanson, D. D., & Hansen, M. (2009). The utility of the cropland data layer for forest inventory and analysis. In R. E. McRoberts, G. A. Reams, P. C. Van Deusen, & W. H. McWilliams (Eds.), *Proceedings of the eighth annual forest inventory and analysis symposium; 2006 October 16–19; Monterey, CA. Gen. Tech. Report WO-79* (pp. 259–264). Washington, DC: Department of Agriculture, Forest Service.

Liknes, G. C., Nelson, M. D., & McRoberts, R. E. (2004). Evaluating classified MODIS satellite imagery as a stratification tool. In R. E. McRoberts, & T. Mowrer (Eds.), *Proceedings of the Joint Meeting of the 6th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences and the 15th Annual Conference of the International Environmetrics Society* (June 28 – July 1 2004, Portland, Maine, USA).

McRoberts, R. E. (2006). A model-based approach to estimating forest area. *Remote Sensing of Environment*, *103*, 56–66.

McRoberts, R. E. (2010). Probability- and model-based approaches to inference for proportion forest using satellite imagery as ancillary data. *Remote Sensing of Environment*, *114*, 1017–1025.

McRoberts, R. E., Gobakken, T., & Næsset, E. (2012). Post-stratified estimation of forest area and growing stock volume using lidar-based stratifications. *Remote Sensing of Environment*, *125*, 157–166.

McRoberts, R. E., Holden, G. R., Nelson, M. D., Liknes, G. C., & Gormanson, D. D. (2006). Using satellite imagery as ancillary data for increasing the precision of estimates for the forest inventory and analysis program of the USDA forest service. *Canadian Journal of Forest Research*, *36*, 2968–2980.

McRoberts, R. E., Nelson, M. D., & Wendt, D. G. (2002). Stratified estimation of forest area using satellite imagery, inventory data, and the k-nearest neighbors technique. *Remote Sensing of Environment*, *82*, 457–468.

McRoberts, R. E., Wendt, D. G., Nelson, M. D., & Hansen, M. D. (2002). Using a land cover classification based on satellite imagery to improve the precision of forest inventory area estimates. *Remote Sensing of Environment*, *81*, 36–44.

Næsset, E. (2002). Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sensing of Environment*, *80*, 88–99.

Næsset, E., & Gobakken, T. (2008). Estimation of above- and below-ground biomass across regions of the boreal forest zone using airborne laser. *Remote Sensing of Environment*, *112*, 3079–3090.

Nelson, M. D., McRoberts, R. E., Liknes, G. C., & Holden, G. C. (2005). Comparing forest/non-forest classifications of Landsat TM imagery for stratifying FIA estimates of forest land area. In R. E. McRoberts, G. A. Reams, P. C. Van Deusen, W. H. McWilliams, & C. J. Cieszewski (Eds.), *Proceedings of the fourth annual forest inventory and analysis symposium: General Technical Report NC-252* (pp. 121–128). St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station (Available at: http://fia.fs.fed.us/symposium/proceedings/, Last accessed: 20 February 2012)

Särndal, C. -E., Swensson, B., & Wretman, J. (1992). *Model assisted survey sampling*. New York: Springer-Verlag, Inc. (694 pp.).

Tomter, S. M., Hylen, G., & Nilsen, J. -E. (2010). Development of Norway's national forest inventory. In E. Tomppo, T. Gschwantner, M. Lawrence, & R. E. McRoberts (Eds.), *National forest inventories — Pathways for common reporting* (pp. 411–424). Springer.

Vanclay, J. K., & Skovsgaard, J. P. (1997). Evaluating forest growth models. *Ecological Modelling*, *98*, 1–12.

Vestjordet, E. (1967). Functions and tables for volume of standing trees. Norway spruce. *Meddelser norske SkogforsVes.*, *22*, 539–574 (In Norwegian with English summary).

Vestjordet, E. (1968). Volum av nyttbart virke hos gran og furu basert på relativ høyde og diameter i brysthøyde eller ved 2,5 m fra stubbeavskjær. *Meddelser norske SkogforsVes*, *25*, 411–549 (In Norwegian with English Summary).

Westfall, J. A., Patterson, P. L., & Coulston, J. W. (2011). Post-stratified estimation: Within-strata and total sample size recommendations. *Canadian Journal of Forest Research*, *41*, 1130–1139.

Zhao, K., Popescu, S., & Nelson, R. (2009). Lidar remote sensing of forest biomass: A scale-invariant estimation approach using airborne lasers. *Remote Sensing of Environment*, *113*, 182–196.