Model-assisted estimation of biomass in a LiDAR sample survey in Hedmark County, Norway[1]

Timothy G. Gregoire, Göran Ståhl, Erik Næsset, Terje Gobakken, Ross Nelson, and Sören Holm

Can. J. For. Res. Downloaded from www.nrcresearchpress.com by WAGENINGEN UR on 12/03/13
For personal use only.

Abstract: Inasmuch as LiDAR is becoming an increasingly prominent tool for forest inventory, it is timely to develop a framework to understand the statistical properties of LiDAR-based estimates. A model-assisted approach to estimation and inference when using LiDAR as a tool to inventory aboveground forest biomass is presented. An empirical example is also presented, yet the article's focus is largely methodological. The sampling plan in the example is viewed as a two-stage design, with slightly different primary sampling units between the profiling and scanning laser surveys. A regression estimator is presented that uses biomass data from the Norwegian National Forest Inventory as the response variable and laser-derived variables as covariates. A major thrust of this article is the presentation of the variance of the estimators of total biomass and biomass per hectare as well as variance estimators.

Résumé : Dans la mesure où le LiDAR devient un outil de plus en plus important pour l'inventaire forestier, il est opportun d'élaborer un cadre pour comprendre les propriétés statistiques des estimations basées sur le LiDAR. Nous présentons une approche basée sur la modélisation pour faire des estimations et tirer des conclusions lorsque le lidar est utilisé comme outil d'inventaire de la biomasse forestière aérienne. L'article porte avant tout sur la méthodologie bien qu'un exemple empirique soit présenté. Dans l'exemple, le plan d'échantillonnage est considéré comme un dispositif à deux étapes, avec des unités primaires d'échantillonnage légèrement différentes pour les relevés par profilage et par balayage laser. L'estimateur de régression qui est présenté utilise les données de la biomasse de l'inventaire forestier national de la Norvège comme variable de réponse et les variables dérivées du lidar comme covariables. L'idée maîtresse de cet article est la présentation de la variance des estimateurs de la biomasse totale et de la biomasse par hectare, ainsi que les estimateurs de la variance.

[Traduit par la Rédaction]

Introduction

Both profiling and scanning lasers have been proposed for the estimation of aboveground forest biomass. Sampling designs for laser-aided inventory have varied among applications, yet commonly, they involve two or more levels of sampling and the use of laser pulse data as a source of auxiliary information. In this article, we consider a specific design that was implemented in Hedmark County, Norway, for the purpose, inter alia, of comparing the utility of profiling and scanning laser for biomass estimation. In particular, we present the estimators developed for this survey as well as estimators of their variance in a model-assisted design-based framework. In a companion paper (Ståhl et al. 2011), a model-based approach is considered. For an appreciation of the distinctions between inference based on the sampling design versus on a model, see Gregoire (1998).

The development here draws heavily from Särndal et al. (1992) and from Särndal and Hidiroglou (1989) yet delves into new areas not covered by these or other authors: the systematic selection of primary and secondary sampling units.

Materials and methods

The Hedmark County survey

A large area survey of Hedmark County, Norway, was conducted for the purpose of linking airborne laser measurements of vegetation with field data from the Norwegian National Forest Inventory (NFI) to estimate aboveground forest biomass (AGB). Both profiling and scanning lasers were

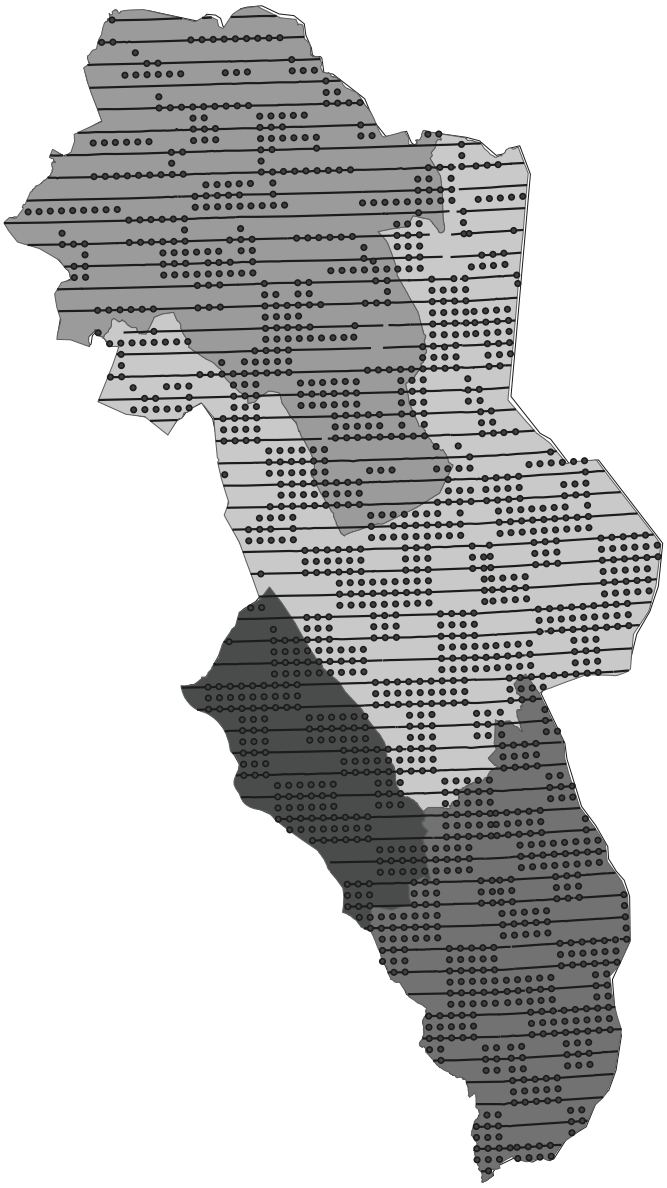
Received 12 January 2010. Accepted 4 October 2010. Published on the NRC Research Press Web site at cjfr.nrc.ca on 20 December 2010.

T.G. Gregoire.[2] School of Forestry and Environmental Studies, Yale University, 360 Prospect Street, New Haven, CT 06511-2104, USA.
G. Ståhl and S. Holm. Department of Forest Resource Management and Geomatics, Swedish University of Agricultural Sciences, 90183 Umeå, Sweden.
E. Næsset and T. Gobakken. Department of Ecology and Natural Resource Management, Norwegian University of Life Sciences, P.O. Box 5003, NO-1432, Ås, Norway.
R. Nelson. 614.4/Biosperic Science Branch, NASA-Goddard Space Flight Center, Greenbelt, MD 20771, USA.

[1]This article is one of a selection of papers from Extending Forest Inventory and Monitoring over Space and Time.
[2]Corresponding author (e-mail: timothy.gregoire@yale.edu).

Can. J. For. Res. Downloaded from www.nrcresearchpress.com by WAGENINGEN UR on 12/03/13
For personal use only.

Fig. 1. Norwegian National Forest Inventory plot locations in Hedmark County. The shading delineates the four administrative units of the county. The 53 horizontal lines are the airborne laser scanning flight lines (spacing of 6 km). The profiling laser was flown at 3 km intervals.



used. A two-way cross-classification (by administrative units (AUs) and by cover type) of the county was conducted with the aim of producing estimates of individual classes as well as the entire county.

The survey may be viewed as a two-stage sampling design with subsequent model-assisted regression estimation of AGB. Hedmark County is located in southeastern Norway bordering Sweden (Fig. 1). Its land area is approximately 27 390 km^2. It comprises four AUs, and a separate estimate of AGB for each AU was sought.

Hedmark County was partitioned by cover into $C = 8$ classes based on existing land use maps and Landsat satellite images. The provision for separate estimates of AGB by cover class also was desired. The eight cover classes (see Table 1) included four productive forest classes: (1) high-, (2) medium-, and (3) low-productivity forests and (4) young forest. The remaining four classes were either nonproductive forest or nonforest classes: (5) nonproductive forest, (6) mountain areas >850 m above sea level, (7) developed areas, e.g., residential areas and infrastructure, and (8) open water.

Predictions of AGB for a network of field plots were available from the NFI of Norway. In Hedmark County, there are 1483 permanent NFI sample plots, measured in 2005–2007, that had been established on a 3 km × 3 km grid almost everywhere. The radius (diameter) of each circular sample plot is 8.92 m (17.84 m), and therefore, its area is 250 m^2. The measurement protocol stipulates that 20% of the permanent plots be remeasured every year. Design-unbiased estimation of tree frequency and aggregate basal area is possible with the NFI, which is a probability sample of forest resources. However, the reliance of AGB on previously fitted models precludes strict design unbiasedness of AGB estimates. Typically, the design-based properties of such estimates are made conditionally on the model-based predictions of individual tree biomass. The work reported here to derive the variance of AGB estimates abides by this conventional practice.

On each sample plot, the diameter at breast height (DBH) of each tree with DBH > 5 cm was measured by calliper. The heights of 10 sample trees per plot, on average, were also measured. Biomass for each tree taller than 1.3 m was predicted with a previously fitted allometric model presented by Marklund (1988), which uses diameter and height as predictor variables. Trees taller than 1.3 m but with DBH < 5 cm were counted and their heights were estimated by means of models presented by Tomter (1998). The spatial coordinates of each plot center were determined with a measurement tolerance of <1 m using differential Global Positioning Systems and Global Navigation Satellite System measurements according to the procedures suggested by Næsset (2001).

Both profiling (portable airborne laser system (PALS); Nelson et al. 2003) and scanning laser (airborne laser scanning (ALS)) height data were collected in 2006 from fixed-wing aircrafts. The flight lines were oriented east–west and followed the NFI grid (see Fig. 1).

For both PALS and ALS, the laser measurements and the metrics derived from them are the source of the auxiliary information, denoted by x, that is linked via regression to AGB observed on the NFI ground plots. Auxiliary information is available for each primary sampling unit (PSU) of the two-stage sample. Therefore, the design accords with Case C of Särndal et al. (1992, section 8.2) wherein "the auxiliary vector value x_k is available for all elements in selected PSUs only...", as opposed to all PSUs in the entire population.

Laser profiling

With airborne laser systems, a short-duration pulse of light is emitted from the instrument and reflected back from the target on the ground, such as trees or humanmade objects or from the terrain. The size of the footprint of the laser beam as it hits the ground might be on the order of, say, 0.1–1 m. The travel time is converted to distance and by extracting the height of the terrain surface, the relative

Table 1. Cover classes by area percentage in Hedmark County and the corresponding number of Norwegian National Forest Inventory (NFI) plots in each class.

	Area (%)	No. of NFI plots
Productive forest		
High	5	92
Medium	13	243
Low	16	306
Young	17	334
All productive forest	51	975
Nonproductive forest and nonforest		
Nonproductive forest	11	167
Mountain areas	28	182
Developed areas	5	82
Open water	5	77
All nonproductive forest and nonforest	49	508

Can. J. For. Res. Downloaded from www.nrcresearchpress.com by WAGENINGEN UR on 12/03/13
For personal use only.

height of the trees and tree canopies can be computed. With a profiling system such as PALS, data are collected for consecutive footprints in a narrow strip beneath the platform. In the current study, PALS height data were collected on m_p = 105 parallel flight lines. The flight lines were spaced 3 km apart and collectively spanned 8309 km. Of the 1483 NFI plots measured in 2005–2007 and made available to this study, 916 plots were overflown within 17.8 m of plot center. In subsequent processing of these PALS data, each flight line was partitioned into a succession of 17.8 m segments.

Laser measurements of canopy height and crown density were extracted from the 17.8 m segment closest to the center of each NFI ground plot. These height and density data were then related to per-unit-area AGB, as measured in the NFI field plots, by fitting linear regression models for all productive forest cover classes combined, as well as a separate regression for each of the four nonproductive forest classes. These models were used subsequently to predict AGB per hectare for each 17.8 m segment.

The general PALS model for biomass per hectare, y', was

$$[1] \qquad y' = \mu(x; \delta) + \epsilon$$

where the $\mu(x; \delta)$ was a linear mean function parameterized by the vector δ. In eq. 1, x represents one or more PALS measurements of canopy height or summary statistics (e.g., quadratic mean canopy height) from a 17.8 m segment. The set of covariates varied somewhat among the regressions fitted for the different cover classes. After having fitted the regression, a prediction of biomass per hectare was obtained as $\widehat{y}' = \mu(x; \widehat{\delta})$. Letting $y = y'/40$ and $\widehat{y} = \widehat{y}'/40$, to convert from biomass per hectare to biomass per 250 m^2 plot, the discrepancy between biomass observed on the NFI plot and predicted biomass per plot is $e(y) = y - \widehat{y}$.

Viewing the PALS survey in Hedmark County as a two-stage sampling design, the flight line is the PSU, whereas each segment is viewed as a secondary sampling unit (SSU). Each SSU was designated as belonging to a single cover class. The AU into which an SSU was assigned depended solely on its geographic location relative to the boundaries of the AU.

Laser scanning

With airborne laser scanning, pulses are distributed continuously along the flight line within a wider swath and not just along a narrow strip as with PALS. The swath width can typically be 500 m or more. For ALS data acquisition, m_a = 53 parallel east–west flight lines, spaced at 6 km intervals, were established. These overflew a total of 705 NFI plots (see Fig. 1). Collectively, the ALS flight lines spanned 4570 km and covered an area of 2297 km^2, or 8.4% of Hedmark County.

The average pulse density for the scanner was 2.8 m^{-2}. In total, about 6.5 billion pulses were acquired.

Similar to the processing of the PALS height data, the ALS height data were related to AGB from the NFI plots. In this effort, logarithmic regression models with additive errors were fitted by cover class to the AGB per plot. For the high-productivity forest class, the following model was fitted for biomass on 250 m^2:

$$[2] \qquad \ln(y) = \mu(x; \beta) + \epsilon$$

where $\mu(x; \beta)$ is nonlinear in the parameters, β, and where x may include topographic variables as well as canopy density and canopy height metrics based on the ALS altimetry. The choice of covariates varied among models; no model was fitted for the open-water cover class (8) because we lacked plot data for that class. The same model for each particular cover class was used in all AUs.

After having fitted the regression, a prediction of biomass per plot is obtained as $\widehat{y} = \exp[\mu(x; \widehat{\beta})]$. As in the PALS case, the discrepancy between observed and predicted biomass per plot is $e(y) = y - \widehat{y}$.

The data from each scanning flight line were parsed so that only those pulses within a 500 m swath around the centerline of the scan were retained. Each swath was partitioned into regular 250 m^2 grid cells. For the ALS application, the

Can. J. For. Res. Downloaded from www.nrcresearchpress.com by WAGENINGEN UR on 12/03/13
For personal use only.

500 m swath constitutes the PSU, whereas the 250 m^2 grid cells within a swath are the SSUs.

Each cell was designated as belonging to a single cover class. As with PALS, the assignment of an ALS cell into an AU was based on only its location within the AU.

Notation

Population parameters of interest

We use τ to denote the total AGB in Hedmark County. The average AGB per hectare in Hedmark County is

$$[3] \qquad \lambda = \tau A^{-1}$$

where A represents the horizontal land area of Hedmark County expressed in hectares.

Using u to index the U AUs in Hedmark County, the total biomass in the uth unit is τ_u, biomass per hectare is λ_u, and its land area is A_u, where $u = 1, ..., U$. Likewise, using c to index the C cover classes, the total biomass in the cth class is τ_c, biomass per hectare is λ_c, and the land area classified into the cth class is A_c, where $c = 1, ..., C$.

For each possible combination of AU and cover class, let τ_{uc} denote total biomass, λ_{uc} denote biomass per hectare, and A_{uc} denote the corresponding land area, where $uc = 1, ..., UC$ and $UC = U \times C$.

Total biomass, as well as biomass per hectare, in Hedmark County aggregates over the various subdivisions in the usual fashion:

$$\tau = \sum_{u=1}^{U} \tau_u = \sum_{c=1}^{C} \tau_c = \sum_{uc=1}^{UC} \tau_{uc}$$

$$\lambda = A^{-1} \sum_{u=1}^{U} A_u \lambda_u = A^{-1} \sum_{c=1}^{C} A_c \lambda_c = A^{-1} \sum_{uc=1}^{UC} A_{uc} \lambda_{uc}$$

Inclusion probabilities

Let the probability with which the pth PSU flight line is selected into the first-stage sample be denoted by π_p. For ALS, the systematic placement of 500 m wide swaths 6 km apart implies that $\pi_p = 500/6000 = 1/12$. This is approximately the same as m_a/M, where M is the number of 500 m wide swaths that could partition Hedmark County. For PALS, we deduced the first-stage inclusion probability in an analogous fashion to provide an approximate first-stage inclusion probability $\pi_p = 1/214$.

In the parlance of systematic sampling, the sampling interval is $I_a = 12$ when intervals are measured in units of 500 m wide swaths for ALS, and for PALS, it is $I_p = 214$.

Let the pth PSU be partitioned into N_p SSUs (250 m^2 cells for ALS and 17.8 m line segments for PALS). From these, we presume that n_p are randomly selected into the second-stage sample. These are the 250 m^2 NFI plots. Conditioned on having selected the pth PSU at the first stage, the inclusion probability of including the NFI plot at the second stage is approximately $\pi_{\mathrm{NFI}|p} = n_p / N_p$. With this approximation, we are not introducing bias into the estimators of τ and λ, which we shall present, by ignoring the systematic placement of the NFI plots. By presuming a simple random second stage of selection, we thereby can appeal to estimators of the variance of the two-stage regression estimators of τ and λ. The consequence of ignoring the systematic nature of the second-stage selection is a likely overestimation of variance (Gregoire and Valentine 2008, p. 55).

We ignore the complication that a PALS line segment will not bisect an NFI plot. We assume that one effect of this error of location will be a less precisely fitted regression than otherwise would be obtained. Another, and potentially more serious, consequence is that the precision of the model-assisted regression estimator of AGB will be degraded quite seriously because of the possibly large residual value resulting from misregistration.

We ignore, too, the complication that an ALS grid cell and an NFI plot will not exactly coincide. We assume that because their areas are identical, the discordance of the actual plot configuration with the presumed configuration will have only minor consequences on estimation and inference.

First-stage sample size by AU

Let $M_u < M$ be the number of 500 m wide swaths that could partition the uth AU.

While m_p and m_a denote the overall number of PSUs by PALS and ALS, respectively, we need to separately indicate the number of PSUs that wholly or partly intersect each AU. For this purpose, let m_{pu} denote the number of PALS flight lines that intersect the uth AU and let m_{au} denote the number of ALS swaths that do likewise. We emphasize that the subset of PSUs in an AU does not alter the first-stage inclusion probability, π_p, of each PSU.

Two-stage estimation of biomass by AU and cover class

In the two-stage setup, an estimator of τ_{uc} from PALS and from ALS share a similar construction. For PALS:

$$[4] \qquad \hat{\tau}^P_{uc} = \sum_{p=1}^{m_{pu}} \frac{\hat{\tau}^P_{puc}}{\pi_p} \approx I_p \sum_{p=1}^{m_{pu}} \hat{\tau}^P_{puc}$$

where $\hat{\tau}^P_{puc}$ is the estimator of τ based on the SSU line segments in the pth PSU flight line. For ALS:

$$[5] \qquad \hat{\tau}^A_{uc} = \sum_{p=1}^{m_{au}} \frac{\hat{\tau}^A_{puc}}{\pi_p} \approx I_a \sum_{p=1}^{m_{au}} \hat{\tau}^A_{puc}$$

Table 2. Norwegian National Forest Inventory plot frequency by administrative unit and cover class.

	Administrative unit				
	1	2	3	4	All
Productive forest					
High	50	17	24	1	92
Medium	81	23	113	26	243
Low	39	34	147	86	306
Young	22	30	140	42	334
All productive forest	292	104	424	155	975
Nonproductive forest and nonforest					
Nonproductive forest	17	12	71	67	167
Mountain areas	0	3	59	120	182
Developed areas	32	31	11	8	82
Open water	21	20	26	10	77
All nonproductive forest and nonforest	70	66	167	105	508

Can. J. For. Res. Downloaded from www.nrcresearchpress.com by WAGENINGEN UR on 12/03/13
For personal use only.

where $\widehat{\tau}^A_{puc}$ is the estimator of τ based on the 250 m^2 SSU grid cells in the pth PSU swath.

Because of the great similarity of eq. 4 to eq. 5, in the sequel, we shall denote both by $\widehat{\tau}_{uc}$, with the understanding that $\widehat{\tau}_{uc} = \widehat{\tau}^P_{uc}$ when referring to the estimator based on PALS and $\widehat{\tau}_{uc} = \widehat{\tau}^A_{uc}$ when referring to the estimator based on ALS. Likewise, we shall use $\widehat{\tau}_{puc}$, with the understanding that $\widehat{\tau}_{puc} = \widehat{\tau}^P_{puc}$ when referring to the estimator based on PALS and $\widehat{\tau}_{puc} = \widehat{\tau}^A_{uc}$ when referring to the estimator based on ALS.

In this more generic notation:

$$\widehat{\tau}_{uc} \approx I \sum_{p=1}^{m_u} \widehat{\tau}_{puc}$$

When evaluating $\widehat{\tau}_{uc}$ with PALS data, $m_u = m_{pu}$ and $I = I_p$. Likewise, when evaluating $\widehat{\tau}_{uc}$ with ALS data, $m_u = m_{au}$ and $I = I_a$.

Both eqs. 4 and 5 rely on the fitted regression that relates biomass per hectare on a PALS line segment or on a 250 m^2 plot, respectively, to laser metrics, namely eq. 1 for PALS and eq. 2 for ALS.

Let $\widehat{y}_{pk}$ be the prediction of biomass on 250 m^2 from the fitted regression corresponding to the kth SSU flight line segment (PALS) or swath (ALS) in the pth PSU. Not all of these will be in the uth AU and cth cover class, yet only the subset of SSUs in the uth AU and cth cover class are used in the estimator $\widehat{\tau}_{puc}$. Therefore, let $\phi_{puc}(y)$ be the following subset of the $\widehat{y}_{pk}$ predictions on the pth PSU:

$$\phi_{puc}(y) = \{\widehat{y}_{pk} | k\text{th SSU of } p\text{th PSU is in } u\text{th AU and } c\text{th cover class}\}$$

Let $\phi_{puc2}(y)$ be the subset of $\phi_{puc}(y)$ that contains only those $\widehat{y}_{pk}$ from SSUs that passed over an NFI plot, i.e., those SSUs that were selected into the second stage of sampling on the pth PSU.

The $\widehat{y}_{pk}$ predictions may be obtained from separate regressions by cover class or AU or from a regression common to two or more cover classes and AUs. In practice, it would matter whether separate or pooled regressions were used for prediction, but for the purpose of deriving the statistical properties of the estimators of AGB, there is no need to distinguish between the two cases.

The regression estimator of total biomass in the ucth class on the pth PSU is

$$[6] \qquad \widehat{\tau}_{puc} = \sum\nolimits_{\phi_{puc}(y)} \widehat{y}_{pk} + \sum\nolimits_{\phi_{puc2}(y)} \frac{y_{pk} - \widehat{y}_{pk}}{\pi_{\mathrm{NFI}|p}} \approx \sum\nolimits_{\phi_{puc}(y)} \widehat{y}_{pk} + \frac{N_p}{n_p} \sum\nolimits_{\phi_{puc2}(y)} e_{pk}(y)$$

where $e_{pk}(y) = y_{pk} - \widehat{y}_{pk}$ is the residual from the fitted regression. Notice that first term of eq. 6 is a sum over a larger set of SSUs than the sum in the second term. As Särndal and Hidiroglou (1989) articulated well, the second term provides a Horvitz–Thompson-like correction for the design-based bias of the first term.

When the pth PSU lacks any SSUs in the ucth class, i.e., when the set $\phi_{puc}(y)$ is null, then $\widehat{\lambda}_{puc} = 0$, by definition, as per Särndal and Hidiroglou (1989, p. 267). In the nonnull case, let $n_{puc} = |\phi_{puc}(y)|$ denote its size. Similarly, when $\phi_{puc2}(y) = \oslash$, the correction term in eq. 6 is zero. Otherwise, let $n_{puc2} = |\phi_{puc2}(y)|$ denote its size. With eq. 6, the total biomass may be

Can. J. For. Res. Downloaded from www.nrcresearchpress.com by WAGENINGEN UR on 12/03/13
For personal use only.

estimated for any combination of AU and cover class that is of interest by means of eq. 4 or 5, as appropriate.

Estimates of biomass per hectare scale down directly as

$$[7] \qquad \widehat{\lambda}_{uc} = \widehat{\tau}_{uc}/A_{uc}$$

The bias and precision of these estimators will depend crucially on the size of the second-stage sample in the *uc*th class. For example, there were 48 NFI plots situated in the medium-productivity forest class in the third AU, yet only six plots are located in the high-productivity forest class in the first AU (see Table 2). The estimated biomass for the former will be considerably more precise and less design-biased (due to the regression estimator) than the estimated biomass for the latter. We turn next to an expression of the variance based on chapter 8 in Särndal et al. (1992).

The approximate variance of $\widehat{\tau}_{uc}$, namely $V(\widehat{\tau}_{uc})$, is

$$[8] \qquad V(\widehat{\tau}_{uc}) \approx V_{\mathrm{I}}(\widehat{\tau}_{uc}) + V_{\mathrm{II}}(\widehat{\tau}_{uc})$$

as in eq. 8.9.24 in Särndal et al. (1992). The first term on the right is the component of variance of $\widehat{\tau}_{uc}$ due to variation in its value among all possible samples of PSUs for the design employed in the sampling. The second term is the corresponding component due to variation among samples of SSUs within their respective PSUs.

Särndal et al. (1992) provided expressions for $V_{\mathrm{I}}(\widehat{\tau}_{uc})$ in their eqs. 4.3.10 and 8.3.2, which in the notation of this paper is

$$[9] \qquad V_{\mathrm{I}}(\widehat{\tau}_{uc}) = \sum_{p=1}^{M_u}\sum_{p'=1}^{M_u}(\pi_{pp'} - \pi_p\pi_{p'})\left(\frac{\tau_{puc}}{\pi_p}\right)\left(\frac{\tau_{p'uc}}{\pi_{p'}}\right)$$

Using r to index the I possible systematic samples of PSUs, and using τ_{ruc} to denote the total biomass in the *uc*th class of the rth first-stage sample, we get from eq. 3.4.4 of Särndal et al. (1992) that eq. 9 can be reexpressed as

$$[10] \qquad V_{\mathrm{I}}(\widehat{\tau}_{uc}) = I\sum_{r=1}^{I}(\tau_r - \tau/I)^2$$

where $I = I_a$ or $I = I_p$, as appropriate. As is well known (Gregoire and Valentine 2008), the expression in eq. 10 does not admit of a design-unbiased estimator even in the simple situation of single-stage element sampling. We are unaware of anyone who has considered the case that we are presently considering, namely systematic sampling of PSUs in a two-stage design.

It seems more fruitful, then, to presume a first-stage design that does permit unbiased estimation of $V_{\mathrm{I}}(\widehat{\tau}_{uc})$ in eq. 8, with the consequent realization that it is usually a conservative approximation to the variance of the estimator of τ_{uc}. The universal default design is simple random sampling without replacement, which leads to

$$[11] \qquad V_{\mathrm{I}}(\widehat{\tau}_{uc}) = \sum_{p=1}^{M_u}\sum_{p'=1}^{M_u}(\pi_{pp'} - \pi_p\pi_{p'})\left(\frac{\tau_{puc}}{\pi_p}\right)\left(\frac{\tau_{p'uc}}{\pi_{p'}}\right) \approx M^2\left(\frac{1}{m} - \frac{1}{M}\right)\sigma^2_{\tau uc}$$

where

$$[12] \qquad \sigma^2_{\tau uc} = (M_u - 1)^{-1}\sum_{p=1}^{M_u}(\tau_{puc} - \tau_{uc}/M_u)^2$$

and where $m = m_a$ or $m = m_p$, as appropriate.

Also in eq. 8:

$$[13] \qquad V_{\mathrm{II}}(\widehat{\tau}_{uc}) = \sum_{p=1}^{M_u}\frac{1}{\pi_p}\left[\sum\sum_{k\in\phi_{pc2(y)}}(\pi_{kk'|p} - \pi_{k|p}\pi_{k'|p})\left(\frac{\epsilon_{pk}}{\pi_{k|p}}\right)\left(\frac{\epsilon_{pk'}}{\pi_{k'|p}}\right)\right] \approx \frac{M}{m}\sum_{p=1}^{M_u}N_p^2\left(\frac{1}{n_p} - \frac{1}{N_p}\right)\sigma^2_{\epsilon puc}$$

which, as explained above, presumes simple random sampling without replacement at the second stage of selection. In eq. 13, ϵ is from eq. 1 or 2 for PALS and ALS, respectively, and

$$[14] \qquad \sigma^2_{\epsilon puc} = (n_{puc2} - 1)^{-1}\sum_{k\in\phi_{puc2}(y)}\epsilon^2_{pk}$$

is the within-PSU variability of biomass among the regression errors for the SSUs that are in the *uc*th class.

To summarize so far:

Can. J. For. Res. Downloaded from www.nrcresearchpress.com by WAGENINGEN UR on 12/03/13
For personal use only.

[15] $$V(\widehat{\tau}_{uc}) \approx M^2\left(\frac{1}{m}-\frac{1}{M}\right)\sigma^2_{\tau uc}+\frac{M}{m}\sum_{p=1}^{M_u}N_p^2\left(\frac{1}{n_p}-\frac{1}{N_p}\right)\sigma^2_{\epsilon puc}$$

Therefore:

[16] $$V(\widehat{\lambda}_{uc}) \approx \frac{1}{A^2_{uc}}V(\widehat{\tau}_{uc})$$

An estimator of $V(\widehat{\tau}_{uc})$ in eq. 15 is (Särndal et al. 1992, eq. 8.9.27)

[17] $$\widehat{V}(\widehat{\tau}_{uc}) = \widehat{V}_{\mathrm{I}}(\widehat{\tau}_{uc}) + \widehat{V}_{\mathrm{II}}(\widehat{\tau}_{uc})$$

where

[18] $$\widehat{V}_I(\widehat{\tau}_{uc}) = \sum_{p=1}^{m_u}\sum_{p'=1}^{m_u}\left(\frac{\pi_{pp'}-\pi_p\pi_{p'}}{\pi_{pp'}}\right)\left(\frac{\widehat{\tau}_{puc\pi}}{\pi_p}\right)\left(\frac{\widehat{\tau}_{p'uc\pi}}{\pi_{p'}}\right)-\sum_{p=1}^{m_u}\left(\frac{1-\pi_p}{\pi_p^2}\right)\widehat{V}_{ypuc} = M^2\left(\frac{1}{m}-\frac{1}{M}\right)\left(s^2_{\widehat{\tau}uc}-m_u^{-1}\sum_{p=1}^{m_u}\widehat{V}_{ypuc}\right)$$

where $m = m_a$ and $m_u = m_{au}$ with ALS and $m = m_p$ and $m_u = m_{pu}$ with PALS. In eq. 18:

[19] $$s^2_{\widehat{\tau}uc} = (m_u-1)^{-1}\sum_{p=1}^{m_u}(\widehat{\tau}_{puc\pi}-\overline{\widehat{\tau}}_{uc\pi})^2$$

[20] $$\widehat{\tau}_{puc\pi} = \sum_{k\in\phi_{puc2}(y)}\frac{y_{pk}}{\pi_{k|p}} \approx \frac{N_p}{n_p}\sum_{k\in\phi_{puc2}(y)}y_{pk}$$

[21] $$\overline{\widehat{\tau}}_{uc\pi} = m_u^{-1}\sum_{p=1}^{m_u}\widehat{\tau}_{puc\pi}$$

Also in eq. 18:

[22] $$\widehat{V}_{ypuc} = \sum\sum_{k,\,k'\in\phi_{puc2}(y)}\left(\frac{\pi_{kk'|p}-\pi_{k|p}\pi_{k'|p}}{\pi_{kk'|p}}\right)\left(\frac{y_{pk}}{\pi_{k|p}}\right)\left(\frac{y_{pk'}}{\pi_{k'|p}}\right) = N_p^2\left(\frac{1}{n_p}-\frac{1}{N_p}\right)s^2_{ypuc}$$

where

[23] $$s^2_{ypuc} = (n_{puc2}-1)^{-1}\sum_{k\in\phi_{puc2}(y)}(y_{pk}-\overline{y}_p)^2$$

[24] $$\overline{y}_p = n^{-1}_{puc2}\sum_{k\in\phi_{puc2}(y)}y_{pk}$$

Also in the variance estimator presented in eq. 17:

[25] $$\widehat{V}_{\mathrm{II}}(\widehat{\tau}_{uc}) = \sum_{p=1}^{m_u}\frac{1}{\pi_p^2}\left[\sum\sum_{kk'\in\phi_{puc2}(y)}\frac{\pi_{kk'|p}-\pi_{k|p}\pi_{k'|p}}{\pi_{kk'|p}}\left(\frac{e_{pk}(y)}{\pi_{k|p}}\right)\left(\frac{e_{pk'}(y)}{\pi_{k'|p}}\right)\right] \approx \frac{M^2}{m^2}\sum_{p=1}^{m_u}N_p^2\left(\frac{1}{n_p}-\frac{1}{N_p}\right)s^2_{epuc}$$

where

[26] $$s^2_{epuc} = (n_{puc2}-1)^{-1}\sum_{k\in\phi_{puc2}(y)}[e_{pk}(y)-\overline{e}_p(y)]^2$$

[27] $$\overline{e}_p(y) = n^{-1}_{puc2}\sum_{k\in\phi_{puc2}(y)}e_{pk}(y)$$

To summarize the variance estimator derived above:

Can. J. For. Res. Downloaded from www.nrcresearchpress.com by WAGENINGEN UR on 12/03/13
For personal use only.

$$[28] \quad \widehat{V}(\widehat{\tau}_{uc}) \approx M^2\left(\frac{1}{m}-\frac{1}{M}\right)\left[s^2_{\widehat{\tau}uc} - m_u^{-1}\sum_{p=1}^{m_u} N_p^2\left(\frac{1}{n_p}-\frac{1}{N_p}\right)s^2_{ypuc}\right] + \frac{M^2}{m^2}\sum_{p=1}^{m_u} N_p^2\left(\frac{1}{n_p}-\frac{1}{N_p}\right)s^2_{epuc}$$

and hence:

$$[29] \quad \widehat{V}(\widehat{\lambda}_{uc}) \approx \frac{1}{A^2_{uc}}\widehat{V}(\widehat{\tau}_{uc})$$

Estimation of biomass by cover class

Estimates of total biomass by *uc* classes may be added to derive marginal estimates for a particular cover class. For example, to estimate the total biomass, τ_c, in the *c*th cover class:

$$[30] \quad \widehat{\tau}^*_c = \sum_{u=1}^{U}\widehat{\tau}_{uc} = \frac{M}{m}\sum_{u=1}^{U}\sum_{p=1}^{m_u}\widehat{\tau}_{puc}$$

However, the variance of $\widehat{\tau}^*_c$ is not the sum of the variances of the $\widehat{\tau}_{uc}$, in contrast with results from stratified sampling (cf. table 5.1 of Gregoire and Valentine 2008). The nonadditivity of the variances stems from the covariances of estimators in different AUs.

As an alternative to $\widehat{\tau}^*_c$, one could estimate the total biomass in the *c*th cover class as

$$[31] \quad \widehat{\tau}_c = \frac{M}{m}\sum_{p=1}^{m}\widehat{\tau}_{pc}$$

in which

$$[32] \quad \widehat{\tau}_{pc} = \sum_{k\in\phi_{pc}(y)}\widehat{y}_{pk} + \sum_{k\in\phi_{pc2}(y)}\frac{y_{pk}-\widehat{y}_{pk}}{\pi_{\mathrm{NFI}|p}} = \sum_{k\in\phi_{pc}(y)}\widehat{y}_{pk} + \frac{N_p}{n_p}\sum_{k\in\phi_{pc2}(y)} e_{pk}(y)$$

where $\phi_{pc}(y)$ is the subset of n_{pc} predictions on the *p*th PSU:

$$\phi_{pc}(y) = \{\widehat{y}_{pk}|k\text{th SSU of } p\text{th PSU is in the } c\text{th cover class}\}$$

and $\phi_{pc2}(y)$ is the subset of $\phi_{pc}(y)$ that contains only those $n_{pc2} \le n_{pc}$ SSUs that were selected into the second stage of sampling on the *p*th PSU.

The approximate variance of $\widehat{\tau}_c$, namely $V(\widehat{\tau}_c)$, is

$$[33] \quad V(\widehat{\tau}_c) \approx V_{\mathrm{I}}(\widehat{\tau}_c) + V_{\mathrm{II}}(\widehat{\tau}_c)$$

where

$$[34] \quad V_{\mathrm{I}}(\widehat{\tau}_c) = \sum_{p=1}^{M}\sum_{p'=1}^{M}(\pi_{pp\prime} - \pi_p\pi_{p\prime})\left(\frac{\tau_{pc}}{\pi_p}\right)\left(\frac{\tau_{p\prime c}}{\pi_{p\prime}}\right) \approx M^2\left(\frac{1}{m}-\frac{1}{M}\right)\sigma^2_{\tau c}$$

where

$$[35] \quad \sigma^2_{\tau c} = (M-1)^{-1}\sum_{p=1}^{M}(\tau_{pc} - \tau_c/M)^2$$

Also in eq. 33:

$$[36] \quad V_{\mathrm{II}}(\widehat{\tau}_c) = \sum_{p=1}^{M}\frac{1}{\pi_p}\left[\sum\sum_{k\in\phi_{puc2}(y)}(\pi_{kk\prime|p} - \pi_{k|p}\pi_{k\prime|p})\left(\frac{\epsilon_{pk}}{\pi_{k|p}}\right)\left(\frac{\epsilon_{p\prime k}}{\pi_{k\prime|p}}\right)\right] \approx \frac{M}{m}\sum_{p=1}^{M}N_p^2\left(\frac{1}{n_p}-\frac{1}{N_p}\right)\sigma^2_{\epsilon pc}$$

where

Can. J. For. Res. Downloaded from www.nrcresearchpress.com by WAGENINGEN UR on 12/03/13
For personal use only.

[37] $$\sigma^2_{\epsilon pc} = (n_{pc2} - 1)^{-1} \sum_{k \in \phi_{pc2}(y)} \epsilon^2_{pk}$$

is the within-PSU variability of biomass among the regression errors for the SSUs that are in the cth class.

An estimator of $V(\widehat{\tau}_c)$ in eq. 33 is (Särndal et al. 1992, eq. 8.9.27)

[38] $$\widehat{V}(\widehat{\tau}_c) = \widehat{V}_{\mathrm{I}}(\widehat{\tau}_c) + \widehat{V}_{\mathrm{II}}(\widehat{\tau}_c)$$

where

[39] $$\widehat{V}_{\mathrm{I}}(\widehat{\tau}_c) = M^2 \left(\frac{1}{m} - \frac{1}{M}\right) \left(s^2_{\hat{\tau}c} - \sum_{p=1}^{m} \widehat{V}_{ypc}\right)$$

and

[40] $$\widehat{V}_{\mathrm{II}}(\widehat{\tau}_c) \approx \frac{M^2}{m^2} \sum_{p=1}^{m} N_p^2 \left(\frac{1}{n_p} - \frac{1}{N_p}\right) s^2_{epc}$$

where

[41] $$s^2_{\hat{\tau}c} = (m-1)^{-1} \sum_{p=1}^{m} (\widehat{\tau}_{pc\pi} - \overline{\widehat{\tau}}_{c\pi})^2$$

[42] $$\widehat{\tau}_{pc\pi} = \sum_{k \in \phi_{pc2}(y)} \frac{y_{pk}}{\pi_{k|p}} \approx \frac{N_p}{n_p} \sum_{k \in \phi_{pc2}(y)} y_{pk}$$

[43] $$\overline{\widehat{\tau}}_{c\pi} = m^{-1} \sum_{p=1}^{m} \widehat{\tau}_{pc\pi}$$

[44] $$\widehat{V}_{ypc} = N_p^2 \left(\frac{1}{n_p} - \frac{1}{N_p}\right) s^2_{ypc}$$

[45] $$s^2_{ypc} = (n_{pc2} - 1)^{-1} \sum_{k \in \phi_{pc2}(y)} (y_{pk} - \overline{y}_p)^2$$

[46] $$\overline{y}_p = n_{pc2}^{-1} \sum_{k \in \phi_{pc2}(y)} y_{pk}$$

[47] $$s^2_{epc} = (n_{pc2} - 1)^{-1} \sum_{k \in \phi_{pc2}(y)} (e_{pk}(y) - \overline{e}_p(y))^2$$

[48] $$\overline{e}_p(y) = n_{pc2}^{-1} \sum_{k \in \phi_{pc2}(y)} e_{pk}(y)$$

Collecting the above results together yields

[49] $$\widehat{V}(\widehat{\tau}_c) = M^2 \left(\frac{1}{m} - \frac{1}{M}\right) \left[s^2_{\hat{\tau}c} - m^{-1} \sum_{p=1}^{m} N_p^2 \left(\frac{1}{n_p} - \frac{1}{N_p}\right) s^2_{ypc}\right] + \frac{M^2}{m^2} \sum_{p=1}^{m} N_p^2 \left(\frac{1}{n_p} - \frac{1}{N_p}\right) s^2_{epc}$$

For estimation of biomass per hectare, $\widehat{V}(\widehat{\lambda}_c) = \widehat{V}(\widehat{\tau}_c)/A_c^2$, as usual.

The above formulae would also apply if one were interested in the aggregation of two or more cover classes, e.g., the four classes of productive forest shown in Fig. 1 and Table 2. To apply the above formulae, the indexing subscript "c" would

Can. J. For. Res. Downloaded from www.nrcresearchpress.com by WAGENINGEN UR on 12/03/13
For personal use only.

implicitly represent all SSUs within any of the productive forest classes, and quantities such as $s^2_{\hat{\tau}c}$ in eq. 41 would have to be computed once for the combined classes.

Estimation of biomass by AU

Estimates of total biomass may be added across cover classes to derive marginal estimates for a particular AU. For example, one way to estimate the total biomass, τ_u, in the uth AU is

$$[50] \qquad \widehat{\tau}^*_u = \sum_{c=1}^{C} \widehat{\tau}_{uc}$$

Here, too, the variance of $\widehat{\tau}^*_u$ is not the sum of the variances of the $\widehat{\tau}_{uc}$. The nonadditivity of the variances again stems from the covariances of estimators in different cover classes.

An alternative, direct estimator of τ_u is

$$[51] \qquad \widehat{\tau}_u = \frac{M}{m}\sum_{p=1}^{m_u} \widehat{\tau}_{pu}$$

in which

$$[52] \qquad \widehat{\tau}_{pu} = \sum_{k\in\phi_{pu}(y)} \widehat{y}_{pk} + \sum_{k\in\phi_{pu2}(y)} \frac{y_{pk} - \widehat{y}_{pk}}{\pi_{\mathrm{NFI}|p}} \approx \sum_{k\in\phi_{pu}(y)} \widehat{y}_{pk} + \frac{N_p}{n_p}\sum_{k\in\phi_{pu2}(y)} e_{pk}(y)$$

where $\phi_{pu}(y)$ is the subset of n_{pu} predictions on the pth PSU:

$$\phi_{pu}(y) = \{\widehat{y}_{pk} | k\text{th SSU of } p\text{th PSU is in the } u\text{th AU}\}$$

and $\phi_{pu2}(y)$ is the subset of $\phi_{pu}(y)$ that contains only those $n_{pu2} \le n_{pu}$ SSUs that were selected into the second stage of sampling on the pth PSU.

The approximate variance of $\widehat{\tau}_u$, namely $V(\widehat{\tau}_u)$, is

$$[53] \qquad V(\widehat{\tau}_u) \approx M^2\left(\frac{1}{m} - \frac{1}{M}\right)\sigma^2_{\tau u} + \frac{M}{m}\sum_{p=1}^{M_u} N_p^2\left(\frac{1}{n_p} - \frac{1}{N_p}\right)\sigma^2_{\epsilon pu}$$

where $\sigma^2_{\tau u}$ and $\sigma^2_{\epsilon pu}$ are defined in Appendix A.

An estimator of $V(\widehat{\tau}_u)$ is

$$[54] \qquad \widehat{V}(\widehat{\tau}_u) \approx M^2\left(\frac{1}{m} - \frac{1}{M}\right)\left[s^2_{\hat{\tau}u} - m^{-1}\sum_{p=1}^{m_u} N_p^2\left(\frac{1}{n_p} - \frac{1}{N_p}\right)s^2_{ypu}\right] + \frac{M^2}{m^2}\sum_{p=1}^{m_u} N_p^2\left(\frac{1}{n_p} - \frac{1}{N_p}\right)s^2_{epu}$$

where $s^2_{\hat{\tau}u}$, s^2_{ypu}, and s^2_{epu} are defined in Appendix A. For the per-hectare estimators, $\widehat{V}(\widehat{\lambda}_u) = \widehat{V}(\widehat{\tau}_u)/A^2_u$.

Results from Hedmark County

For all productive forest classes combined, the PALS (59.5 t·ha^{-1}) and ALS estimates (59.9 t·ha^{-1}) were quite close to the 64.0 t·ha^{-1} estimated solely from the 975 NFI ground plots, as seen in Table 3. Indeed, for individual classes of productive forests, the largest magnitude deviation of PALS from ground NFI estimates was approximately –11 t·ha^{-1} in the medium-productivity forest class, and for ALS, the largest deviation was approximately –7 t·ha^{-1} in both the medium- and low-productivity forest classes. In comparison with the ground NFI estimates, both PALS and ALS underestimated biomass for all but the young productive forest class.

For all classes of nonproductive forest and nonforest combined, biomass estimates from PALS and ALS are quite close to the ground NFI estimate of 9.0 t·ha^{-1}. Judged in terms of absolute deviation from the ground NFI estimate, both PALS and ALS did least well in estimating the biomass of the nonproductive forest class.

The differences between the ground NFI and PALS and ALS estimates are more striking when assessed in terms of the estimates of standard error. The NFI estimates are invariably more precise, in part because they are based on more plots than are used in the second stage of the PALS and ALS sampling that was conducted in Hedmark. Even after accounting for this by multiplying the ground NFI standard error values in Table 3 by the factor (number of NFI plots/number of LiDAR plots)$^{0.5}$, the PALS and ALS estimated

Table 3. Biomass estimates by cover class for Hedmark County.

	PALS (t·ha^{-1})			ALS (t·ha^{-1})			Ground NFI (t·ha^{-1})		
	Mean	SE	n_c	Mean	SE	n_c	Mean	SE	n_c
Productive forest									
High	114.8	14.2	52	119.1	12.8	48	121.3	8.2	92
Medium	83.7	10.7	151	87.6	5.8	105	94.5	3.8	243
Low	42.6	6.6	173	39.8	5.7	141	46.6	2.3	306
Young	41.0	2.7	183	40.6	na	151	40.3	2.8	334
All productive forest	59.5	3.2	559	59.9	3.2	445	64.0	1.7	975
Nonproductive forest and nonforest									
Nonproductive forest	27.3	3.5	103	27.2	na	83	22.5	2.1	167
Mountain areas	6.3	1.3	131	5.1	1.0	95	6.5	0.9	182
Developed areas	6.6	1.1	65	5.4	na	46	1.4	1.0	82
Open water	2.6	0.0	58	2.2	0.0	36	0.6	0.6	77
All nonproductive forest and nonforest	10.6	0.8	357	9.7	na	260	9.0	0.7	508
All cover classes	36.1	1.6	916	35.9	1.8	705	37.6	0.9	1483

Note: SE is the estimated standard error computed by the methods and formulae put forth in this article. Also, n_c represents the number of Norwegian National Forest Inventory (NFI) ground plots that were used to construct the estimates for the cover class. "na" denotes the occurrence of a negative estimate of SE. For the water cover class, SE = 0.0 because aboveground biomass was absent on all NFI ground plots used in the LiDAR calculations.

Can. J. For. Res. Downloaded from www.nrcresearchpress.com by WAGENINGEN UR on 12/03/13
For personal use only.

standard errors exceeded those of the NFI ground estimates, sometimes quite substantially. Given the fundamental difference in the sampling designs, this adjustment based on number of plots can best be regarded as but a crude attempt to put the methods on a comparable basis.

Normally, the precision of alternative estimators is compared by computing the ratio of the variance of one with the variance of the other. In our case, this is complicated by the fact that the LiDAR sampling has a two-stage structure, whereas the NFI does not. In both cases, a systematic sampling design was used and a conservative estimate of standard error was employed. Yet we do not know whether the overestimation of standard error was of comparable magnitude for the NFI and the two LiDAR designs.

Moreover, the accuracy of eq. 8, say, relies on large-sample asymptotic results that surely fail to hold for some combinations of cover class and AU in the LiDAR sampling of Hedmark County.

In view of these caveats, the standard error results in Table 3 are encouraging.

Discussion

The PALS flight lines and ALS swaths were flown at a systematic spacing. To the extent that there is a north–south gradient in AGB in Hedmark County, our treatment of the first stage of sampling as a simple random sampling will result in an overestimation of the variance of these estimators of AGB. Surely some cover classes will have a more pronounced gradient than others, and the actual precision of AGB estimation for these will be greater than our estimates of variance will indicate. We could attempt to estimate the variance among PSUs with successive differences or balanced differences, but we have yet to do so.

We have also ignored the systematic location of NFI plots that serve as our second-stage sample. Lacking an east–west gradient to the AGB, the effect of their systematic location may not be so pronounced.

As with poststratification generally, the sample sizes in each AU and cover class are random. We implicitly have conditioned on the realized sample sizes, which is a commonly accepted practice.

In addition, there was an adaptive element to the inclusion of second-stage NFI plots with PALS: only those plots whose centers were within 17.8 m of a PALS flight line were included in the second-stage sample.

For cases where the NFI plot overlaps two or more ALS cells, the prediction of biomass for an NFI plot following ALS was made more accurate by basing the predictions on the ALS metrics computed from within the plot borders rather than the ALS cell that contains the major portion of the NFI plot.

The design-based estimators of variance do not account for the uncertainty, whether manifest as bias or variance, in the allometric prediction of the biomass on the NFI ground plot. Although this is a commonly accepted practice in forest inventory globally when estimating timber volume with previously fitted models of bole volume, it is a practice that both vitiates design unbiasedness of estimators of τ and likely results in an underestimation of its design-based variance.

In two-stage sampling designs, the PSUs often are chosen with probability proportional to some measure of size of the PSU. The rationale for such unequal probability of first-stage selection is an expected increase in the precision of estimation. We could, for example, employ importance sampling to include PALS flight lines and ALS swath centerlines with a probability density proportional to length. We did not select PSUs with unequal probability in this study, however. Moreover, we weight appropriately by area when scaling down to an estimate of biomass per hectare in the class, as seen in eq. 7. Nonetheless, there is an abiding

Can. J. For. Res. Downloaded from www.nrcresearchpress.com by WAGENINGEN UR on 12/03/13
For personal use only.

feeling that longer flight lines ought to be given greater weight than shorter ones, either by varying the probability with which they are included into the sample or by some form of weighting by length. We have consigned this avenue of research for the future, perhaps with the help of simulation.

Despite these multiple deficiencies, we believe that the work reported here has succeeded on a number of fronts. We have articulated a design-based analysis for the estimation of AGB that puts it on a sound footing within the field of probability sampling. The design utilizes the height data collected along the entire flight line or swath of a profiling or scanning laser, respectively. Previous inventories with a profiling laser have demonstrated how to utilize laser data from the entire flight line to estimate biomass and forest volume at a regional level (Delaware, USA) (Nelson et al. 2004). To date, such studies have been restricted to regions smaller than 5000 km^2 in area.

As far as scanning lasers are concerned, a major focus has been on wall-to-wall inventory of forest stands for management purposes and these methods have even become a preferred method for stand inventory, at least in Scandinavia (Næsset 2007, 2009). Although such stand-based applications seem to perform well compared with conventional inventories (e.g., Næsset 2002, 2004*a*, 2004*b*), the statistical aspects have yet to be addressed in a rigorous manner. However, there is recent progress in this direction (Andersen and Breidenbach 2007; Corona and Fattorini 2008). Applications of scanning lasers within a sampling framework strike us as an obvious choice to consider for regional applications wherein wall-to-wall coverage would be prohibitively costly. Parker and Evans (2004) and Stephens et al. (2008) used scanning lasers in double-sampling applications to assess volume and biomass, but only those laser data in a swath that coincided with predefined plots were utilized in the estimation. Thus, to the best of our knowledge, the present study is an initial demonstration of how the entire data set of laser observations from a scanning laser can be used to estimate the biomass in a fairly large region in a consistent manner.

Because height measurements derived from airborne lasers are highly correlated with forest biomass, airborne lasers can provide timely and fairly accurate estimates over large areas. We believe that such applications will be considered a viable alternative for monitoring and rapid update of forest biomass and carbon stocks that shall be required to comply with future climate treaties and other international agreements. With the present work, we have offered a means of quantifying the reliability of such estimates.

Acknowledgements

This research was funded by the Research Council of Norway (project No. 166482/i10). The authors are grateful to the Norwegian National Forest Inventory program, which made the empirical part of this study possible by giving access to the NFI field plots in Hedmark County. We also wish to acknowledge the efforts made by Blom Geomatics, Norway, for acquiring the laser scanner data and flying the profiling LiDAR instrument operated by one of the coauthors (Dr. Ross Nelson). Finally, we wish to thank our good colleague, Dr. Ole Martin Bollandsås (Norwegian University of Life Sciences), for coordinating much of the efforts to accurately geolocate the field data through a network of continuously operated GPS/GLONASS reference stations.

References

Andersen, H.-E., and Breidenbach, J. 2007. Statistical properties of mean stand biomass estimators in a LiDAR-based double sampling forest survey design. *In* Proceedings of the ISPRS Workshop Laser Scanning 2007 and Silvilaser 2007, Espoo, Finland. Vol. XXXVI. Part 3/W52. *Edited by* P. Rönholm, H. Hyppä, and J. Hyppä. International Society for Photogrammetry and Remote Sensing, Headquarters: Beijing, PR China.. pp. 8–13.

Corona, P., and Fattorini, L. 2008. Area-based lidar-assisted estimation of forest standing volume. Can. J. For. Res. **38**(11): 2911–2916. doi:10.1139/X08-122.

Gregoire, T.G. 1998. Design-based and model-based inference in survey sampling: appreciating the difference. Can. J. For. Res. **28**(10): 1429–1447. doi:10.1139/cjfr-28-10-1429.

Gregoire, T.G., and Valentine, H.T. 2008. Sampling strategies for natural resources and the environment. Chapman & Hall/CRC, Boca Raton, Fla.

Marklund, L.G. 1988. Biomass functions for pine, spruce and birch in Sweden. Department of Forest Survey, Swedish University of Agricultural Sciences, Umeå, Sweden. [In Swedish.]

Næsset, E. 2001. Effects of differential single- and dual-frequency GPS and GLONASS observations on point accuracy under forest canopies. Photogramm. Eng. Remote Sens. **67**: 1021–1026.

Næsset, E. 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. Remote Sens. Environ. **80**(1): 88–99. doi:10.1016/S0034-4257(01)00290-5.

Næsset, E. 2004*a*. Practical large-scale forest stand inventory using small-footprint airborne laser-scanning. Scand. J. For. Res. **19**: 164–179. doi:10.1080/02827580310019257.

Næsset, E. 2004*b*. Accuracy of forest inventory using airborne laser-scanning: evaluating the first Nordic full-scale operational project. Scand. J. For. Res. **19**: 554–557. doi:10.1080/02827580410019544.

Næsset, E. 2007. Airborne laser scanning as a method in operational forest inventory: status of accuracy assessments accomplished in Scandinavia. Scand. J. For. Res. **22**(5): 433–442. doi:10.1080/02827580701672147.

Næsset, E. 2009. Effects of different sensors, flying altitudes, and pulse repetition frequencies on forest canopy metrics and biophysical stand properties derived from small-footprint airborne laser data. Remote Sens. Environ. **113**(1): 148–159. doi:10.1016/j.rse.2008.09.001.

Nelson, R.F., Parker, G., and Hom, M. 2003. A portable airborne laser system for forest inventory. Photogramm. Eng. Remote Sens. **69**: 267–273.

Nelson, R.F., Short, A., and Valenti, M. 2004. Measuring biomass and carbon in Delaware using an airborne profiling LiDAR. Scand. J. For. Res. **19**(6): 500–511 . [Erratum: Vol. 20, pp. 283–284.] doi:10.1080/02827580410019508.

Parker, R.C., and Evans, D.L. 2004. An application of LiDAR in a double-sampling forest inventory. West. J. Appl. For. **19**: 95–101.

Särndal, C.-E., and Hidiroglou, M.A. 1989. Small domain estimation: a conditional analysis. J. Am. Stat. Assoc. **84**(405): 266–275. doi:10.2307/2289873.

Särndal, C.-E., Swensson, B., and Wretman, J. 1992. Model assisted survey sampling. Springer-Verlag, New York.

Ståhl, G., Holm, S., Gregoire, T.G., Gobakken, T., Næsset, E., and Nelson, R. 2011. Model-based inference for biomass estimation

in a LiDAR sample survey in Hedmark County, Norway. Can. J. For. Res. **41**(1). This issue. doi:10.1139/X10-161.

Stephens, P.R., McGaughey, R.J., Daugherty, T., Geard, B., and Loubser, D. 2008. Quality assurance and quality control procedures of airborne scanning LiDAR for a nation-wide carbon inventory of planted forests. *In* Proceedings of Silvilaser 2008: 8th International Conference on LiDAR Applications in Forest Assessment and Inventory, Edinburgh, U.K. *Edited by* R.A. Hill, J. Rosette, and J. Suárez. Available from http://geography.swan.ac.uk/silvilaser/. pp. 563–571.

Tomter, S.M. 1998. Beregning av volum de første år etter bestandsetablering [Volume estimation during the first years after stand establishment]. Norwegian Institute of Land Inventory, Ås, Norway.

Appendix A. Technical details

With regard to eq. 53:

$$[A1] \qquad \sigma^2_{\tau u} = (M_u - 1)^{-1} \sum_{p=1}^{M_u} (\tau_{pu} - \tau_u / M_u)^2$$

and

$$[A2] \qquad \sigma^2_{\epsilon pu} = (n_{pu2} - 1)^{-1} \sum_{k \in \phi_{pu2}(y)} \epsilon^2_{pk}$$

With regard to eq. 54:

$$[A3] \qquad s^2_{\hat{\tau} u} = (m_{au} - 1)^{-1} \sum_{p=1}^{m_{au}} (\widehat{\tau}_{pu\pi} - \overline{\widehat{\tau}}_{u\pi})^2$$

$$[A4] \qquad \widehat{\tau}_{pu\pi} = \sum_{k \in \phi_{pu2}(y)} \frac{y_{pk}}{\pi_{k|p}} \approx \frac{N_p}{n_p} \sum_{k \in \phi_{pu2}(y)} y_{pk}$$

$$[A5] \qquad \overline{\widehat{\tau}}_{uc\pi} = m^{-1}_{au} \sum_{p=1}^{m_{au}} \widehat{\tau}_{puc\pi}$$

$$[A6] \qquad s^2_{ypu} = (n_{pu2} - 1)^{-1} \sum_{k \in \phi_{pu2}(y)} (y_{pk} - \overline{y}_p)^2$$

$$[A7] \qquad \overline{y}_p = n^{-1}_{pu2} \sum_{k \in \phi_{pu2}(y)} y_{pk}$$

$$[A8] \qquad s^2_{epu} = (n_{pu2} - 1)^{-1} \sum_{k \in \phi_{pu2}(y)} (e_{pk}(y) - \overline{e}_p(y))^2$$

$$[A9] \qquad \overline{e}_p(y) = n^{-1}_{pu2} \sum_{k \in \phi_{pu2}(y)} e_{pk}(y)$$

With regard to eq. 54:

$$[A10] \qquad s^2_{\hat{\tau} u} = (m_{pu} - 1)^{-1} \sum_{p=1}^{m_{pu}} (\widehat{\tau}_{pc\pi} - \overline{\widehat{\tau}}_{c\pi})^2$$

$$[A11] \qquad s^2_{ypu} = (n_{pu2} - 1)^{-1} \sum_{k \in \phi_{pu2}(y)} (y_{pk} - \overline{y}_p)^2$$

$$[A12] \qquad s^2_{epu} = (n_{pu2} - 1)^{-1} \sum_{k \in \phi_{pu2}(y)} (e_{pk}(y) - \overline{e}_p(y))^2$$

Can. J. For. Res. Downloaded from www.nrcresearchpress.com by WAGENINGEN UR on 12/03/13
For personal use only.