A comparison of small-area estimation techniques to estimate selected stand attributes using LiDAR-derived auxiliary variables

Michael E. Goerndt, Vicente J. Monleon, and Hailemariam Temesgen

Abstract: One of the challenges often faced in forestry is the estimation of forest attributes for smaller areas of interest within a larger population. Small-area estimation (SAE) is a set of techniques well suited to estimation of forest attributes for small areas in which the existing sample size is small and auxiliary information is available. Selected SAE methods were compared for estimating a variety of forest attributes for small areas using ground data and light detection and ranging (LiDAR) derived auxiliary information. The small areas of interest consisted of delineated stands within a larger forested population. Four different estimation methods were compared for predicting forest density (number of trees/ha), quadratic mean diameter (cm), basal area (m^2/ha), top height (m), and cubic stem volume (m^3/ha). The precision and bias of the estimation methods (synthetic prediction (SP), multiple linear regression based composite prediction (CP), empirical best linear unbiased prediction (EBLUP) via Fay–Herriot models, and most similar neighbor (MSN) imputation) are documented. For the indirect estimators, MSN was superior to SP in terms of both precision and bias for all attributes. For the composite estimators, EBLUP was generally superior to direct estimation (DE) and CP, with the exception of forest density.

Résumé : Un des défis souvent rencontrés en foresterie est l'estimation des attributs forestiers pour de plus petites zones d'intérêt au sein d'une population plus large qu'est la forêt. Les méthodes d'estimation pour petites zones sont des techniques bien adaptées à l'estimation d'attributs forestiers sur de petites superficies pour lesquelles la taille de l'échantillon existant est faible et l'information auxiliaire est disponible. Quatre méthodes d'estimation pour petites zones ont été sélectionnées et comparées pour estimer une variété d'attributs forestiers sur de petites superficies à l'aide de données terrain et de données auxiliaires dérivées du lidar. Les deux premières méthodes étaient indirectes (la prédiction synthétique (PS) et l'imputation par les plus proches voisins (PPV)); les deux autres étaient composites (la meilleure prédiction empirique linéaire sans biais (PSB) basée sur les modèles de Fay–Herriot et la prédiction composite basée sur la régression linéaire multiple (PC)). Les petites superficies d'intérêt étaient représentées par les peuplements délimités dans une forêt. Les attributs forestiers qui ont été comparés étaient la densité de la forêt (nombre de tiges/ha), le diamètre moyen quadratique (cm), la surface terrière (m^2/ha), la hauteur maximale (m) et le volume des tiges (m^3/ha). La précision et le biais des quatre méthodes d'estimation sont documentés. Dans le cas des estimateurs indirects, l'imputation par les PPV était supérieure à la PS en termes de précision et de biais pour tous les attributs. Dans le cas des estimateurs composites, la PSB était généralement supérieure à l'estimation direct et la PC, excepté pour la densité de la forêt.

[Traduit par la Rédaction]

Introduction

Generally, estimates of forest attributes for areas of interest have been derived by using ground data extracted from plots taken within the area. For forest management and planning purposes, forest areas are partitioned into smaller stands based on attributes such as species composition, forest age, and management history. This partitioning can often make it difficult to obtain precise attribute estimates within the small areas of interest due to small sample sizes. The most basic solution to this problem is to resample each small area using appropriate sample sizes, which can be time consuming and expensive. Small-area estimation (SAE) is an alternative set of techniques well suited to this scenario, as it seeks to obtain precise estimation of selected variables within small areas of interest by incorporating information from the entire population. Increased availability of remotely sensed auxiliary information derived from, e.g., light detection and ranging (LiDAR) technology for entire populations has created new possibilities for the effective use of SAE in estimating selected forest attributes. This study focuses on the use of SAE techniques to obtain precise estimates of selected forest attributes for small areas of interest within a larger forested population using LiDAR-derived auxiliary information.

SAE techniques can effectively increase the precision of forest attribute estimates in many situations where there are

Received 25 May 2010. Accepted 20 November 2010. Published at www.nrcresearchpress.com/cjfr on 24 May 2011.

M.E. Goerndt. Department of Forestry, University of Missouri, 203 Anheuser-Busch Natural Resources Building, Columbia, MO 65211, USA.
V.J. Monleon. Forest Inventory and Analysis, Pacific Northwest Research Station, USDA Forest Service, Corvallis, OR 97331, USA.
H. Temesgen. Department of Forest Engineering, Resources and Management, Oregon State University, Corvallis, OR 97331, USA.

Corresponding author: M.E. Goerndt (e-mail: goerndtm@missouri.edu).

insufficient ground data to achieve the desired level of precision for direct estimates (Petrucci et al. 2005; You and Chapman 2006). There are different types of SAE techniques, the validity of which greatly depends on the structure and composition of both the small area and the available covariates and auxiliary information. Generally, SAE has been divided into three categories: (*i*) direct estimation, (*ii*) indirect estimation, and (*iii*) composite estimation (Costa et al. 2003, 2004). A direct estimator is an estimator that only uses data taken directly from the small area of interest. This implies that there needs to be an adequate sample of data from each area to effectively use direct estimation. Indirect estimators do not necessarily require that there is an adequate sample taken for each area of interest but can "borrow" strength from the auxiliary information within the area and (or) beyond the area to derive a vital component to the estimation process, typically a regression coefficient (Heady et al. 2003). Composite estimation typically combines a direct estimator and an indirect estimator for the particular small area of interest.

A form of indirect estimation that uses information independent of the population of interest is synthetic prediction (SP). In the context of this study, SP pertains to using preexisting models that relate ground-measured forest attributes to LiDAR metrics at either the plot or stand level. There are two primary approaches for predicting forest attributes using LiDAR metrics: (*i*) relating LiDAR metrics to tree-level attributes through single-tree remote sensing (STRS) (Popescu et al. 2003; Chen et al. 2006), and (*ii*) relating LiDAR metrics to area-level forest attributes (Means et al. 2000; Breidenbach et al. 2008). When the goal is to estimate area-level attributes for either a small area or larger population, it has been shown that area-level prediction models are generally more efficient and can also have greater precision than single-tree models, which tend to omit intermediate and suppressed trees within the forest landscape (Popescu 2007; Goerndt et al. 2010). Indirect SAE can include many modeling techniques that are commonly used for estimating forest attributes using auxiliary information, including imputation techniques such as nearest-neighbor (NN) imputation (Rubin 1976; Van Deusen 1997; Moeur 2000).

A method that can incorporate strength from both direct and indirect estimation is a composite predictor (CP), which usually consists of a weighted sum or mean of a direct estimator and indirect estimator for the particular small area of interest. A useful form of indirect SAE is a linear mixed-model known as a Fay–Herriot model (Fay and Herriot 1979; Prasad and Rao 1990; You and Chapman 2006). This is an area-based model in that both the auxiliary and response information used in the modeling process are at the small-area level (e.g., stand level) rather than at a unit level (e.g., plot level). This form of estimator is particularly useful when it is difficult to match unit-level response variables with unit-level auxiliary data. This is a common characteristic when using LiDAR information to estimate forest attributes where the forest inventory plots either have a variable radius or lack accurate spatial coordinates. As a linear mixed model, the Fay–Herriot model incorporates a random effect, which is usually dependent on the small areas within the population of interest. The ultimate objective of using this type of model is to derive empirical best linear unbiased predictions (EBLUP) of the attributes of interest. As a special form of composite prediction, EBLUP depends on incorporation of a direct estimator via a weighting factor and is therefore sample dependent in that it cannot be applied to small areas that lack any ground sample.

The primary goal of this study is to compare selected SAE methods to obtain precise and accurate estimates for selected forest attributes in small areas (stands) of interest within a larger domain. The specific objective is to examine the performance of four synthetic prediction (SP) and composite prediction (CP) methods: synthetic prediction (SP), most similar neighbor (MSN) imputation, empirical best linear unbiased prediction (EBLUP) via Fay–Herriot models, and multiple linear regression-based composite prediction (CP). The performance of the commonly used MSN imputation method is included in this study as a means of comparing CP and SP with a fairly standard method of estimation. The attributes of interest for this study are total stem volume (CuVol, $m^3 \cdot ha^{-1}$), mean height of largest 100 trees per hectare ($Ht \cdot ha^{-1}$), quadratic mean diameter (QDBH, cm), basal area (BA, $m^2 \cdot ha^{-1}$), and density (number of live stems$\cdot ha^{-1}$).

Methods

Study area

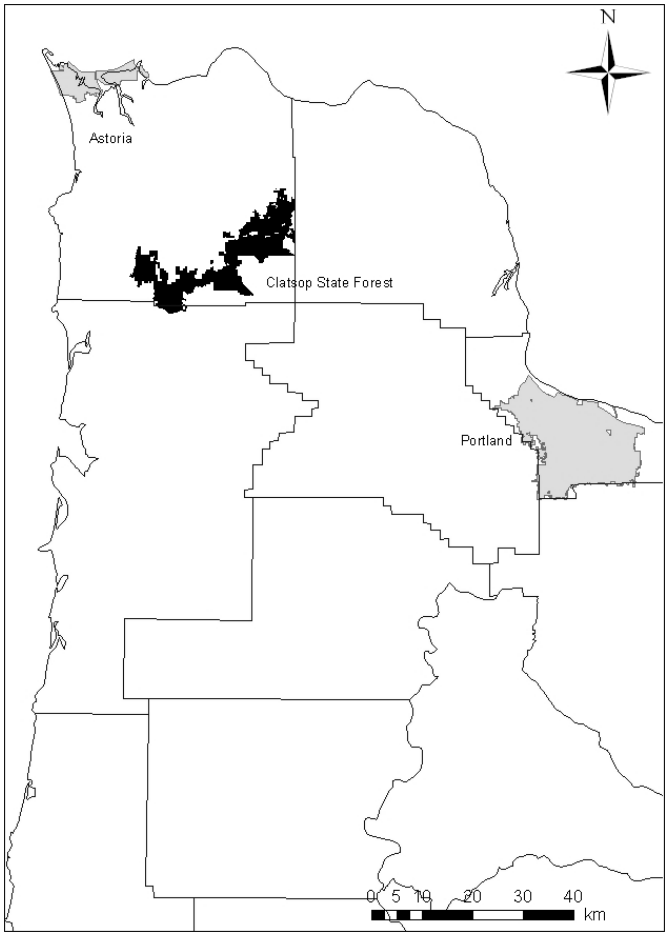
The study was conducted in Clatsop State Forest, located in Clatsop County in northwestern Oregon. The forest covers approximately 32 000 ha, with an elevation range of 60–500 m above sea level. The main tree species found are conifers, including Douglas-fir (*Pseudotsuga menziesii* (Mirbel) Franco), grand fir (*Abies grandis* (Dougl. ex D. Don) Lindl.), western hemlock (*Tsuga heterophylla* (Raf.) Sarg.), and western red cedar (*Thuja plicata* Donn ex D. Don) as apex species. The primary deciduous species are bigleaf maple (*Acer macrophyllum* Pursh) and red alder (*Alnus rubra* Bong.). A map showing the location and relative size of the study area can be seen in Fig. 1.

Field data

The stand level inventory (SLI) system, as developed by the Oregon Department of Forestry (ODF), consists of splitting all of the state-owned forest land area into homogeneous individual stands or groups that are represented by spatial polygons (Oregon Department of Forestry 2008). The classification of the landscape into stands was based primarily on dominant tree species, average tree size (diameter at breast height (DBH)), and stand density. Data used in this study were collected in the Clatsop State Forest from 2002 to 2007. The stands utilized in this study were natural stands that were older than 20 years. Of the 315 existing SLI stands within the Clatsop Forest LiDAR coverage area, 190 stands contained a ground sample. A summary of selected stand attributes is given in Table 1.

All stands were sampled using systematic grid sampling. Variable-radius (prism) plots were used in most stands to measure trees greater than 11.4 cm DBH. Fixed-radius plots were also used in some stands. DBH and species were recorded for each sampled tree, whereas total height and height to crown ratio were measured only on a subsample of trees. Using the subsampled trees, total heights were subsequently estimated using regression equations for trees not sampled for height and height to crown base (SLI 2008).

Fig. 1. Map of northwestern Oregon showing location and relative size of Clatsop State Forest.



LiDAR data

Area-based LiDAR metrics

The LiDAR data were collected in April 2007 using a Leica ALS50 II laser system. The sensor scan angle was ±14° from nadir, with a pulse rate designed to yield an average number of pulses of ≥8 points·m^2 over terrestrial surfaces. Classification of ground and vegetation points was performed by TerraScan v.7.012, as well as spatial interpolation of ground-classified points to create the digital terrain model (DTM). The data were collected using opposing parallel flight lines with a ≥50% overlap. All area-based LiDAR metrics used in this study were extracted from the raw point data using LiDAR FUSION (McGaughey 2008). A summary of the LiDAR metrics with corresponding descriptions is given in Table 2.

All LiDAR metrics were extracted using only first returns above a height of 3 m off the ground, with the exception of the cover metrics and canopy transparency, which used a variety of predetermined height thresholds, because a high number of first returns from the ground and low-lying vegetation may introduce confounding noise in the LiDAR metrics (Strunk 2008). The raw LiDAR intensities extracted in a particular scan were passed through a proprietary algorithm by the vendor to account for several variables such as localized trends in intensity values, scanning angle, and target distance resulting in intensity values per return that were calibrated to an eight-bit value with a range of 0 to 255.

Table 1. Summary of selected forest attributes in Clatsop State Forest (n = 190).

	Minimum	Maximum	Mean	SD
Density (trees/ha)	257.1	4201.3	945.3	678.7
Quadratic mean diameter (QDBH, cm)	9.8	52.8	29.5	9.2
Basal area (BA, m^2/ha)	19.7	86.3	50.2	12.9
Top height (Ht, m)	15.7	54.5	32.7	8.5
Cubic stem volume (CuVol, m^3/ha)	125.2	1318.4	568.7	244.7

Note: SD, standard deviation.

Grid-level LiDAR metrics

In addition to the aforementioned stand-level metrics, another set of LiDAR metrics was created to facilitate the synthetic prediction portion of this study. These metrics were created by dividing the entire area of sampled stands into 30 × 30 m grid cells simulating square-plot areas on the landscape. A separate LiDAR file was created using LiDAR FUSION for each individual grid cell. Finally, LiDAR FUSION was used once again to compute the relevant area-level LiDAR metrics for each grid cell.

Statistical analysis

Simulation

The primary assumption for SAE is that there is an insufficient sample size within each small area of interest to obtain precise direct estimates. The individual stands used in this study contain anywhere from 10 to 36 plots, depending on the stand area. It was necessary to simulate smaller sample sizes to assess the performance of the estimators as the sample size increased and to facilitate validation of the estimators using direct stand estimates of the attributes from the full samples as a surrogate for census information. To this end, only the 134 stands containing at least 20 plots were considered for the analysis. The simulation of reduced sample sizes was done by randomly selecting measured plots from each stand with replacement using 10%, 20%, 30%, 40%, and 50% sampling intensities for 500 iterations. Direct estimates (DE) in the form of stand-level means were calculated for each attribute at all sampling intensities, as well as for the full samples. With the exception of synthetic prediction (SP), which relied on external linear regression equations, all of the estimators were assessed for five sampling intensities separately. Stand-level DEs for each attribute based on the full sample (FS) were retained as validation data for each estimator in the absence of census information. All validation statistics including root mean squared error (RMSE), relative root mean squared error (RRMSE), bias, and relative bias (RB) were calculated using the full-sample DEs as the "observed" values for each stand.

Synthetic prediction (SP)

For the attributes of interest, SP were obtained by applying linear regression models that were previously developed by Goerndt et al. (2010) from plot-level ground-truthed data within in the same region but outside the population of interest. Past studies have indicated that relationships between for-

Table 2. Summary of LiDAR metrics computed using LiDAR FUSION.

Metric	Description
Height	Distribution of all first-return heights > 3 m
Percentile height, e.g., 5th, 10th, 20th, …, 95th	Height distribution by deciles of first returns > 3 m
Intensity	Distribution of all first-return intensities > 3 m
Percentile intensity	Intensity distribution by deciles of first returns > 3 m
Canopy cover (Cover_3, …, Cover_24)	Percentage (0–100) of first returns equal to or greater than a specified height (3, 6, …, 24 m) above the ground
Canopy transparencies	Percentage (0–100) of first returns above a specified height after the removal of returns below a lower specified height

est attributes and LiDAR metrics at the plot level can be potentially different at the area or stand level (Næsset 2002; Means et al. 2000). Therefore, the plot-level linear models from Goerndt et al. (2010) were applied to each 30 × 30 m grid cell to obtain estimates of the forest attributes of interest. Prior to computing stand-level attributes, each stand boundary cell (sliver cell) having an area of <675 m^2 was removed from the data set based on an assumption that cell size affected the relationship between field measurements and LiDAR metrics (Næsset 2002). The estimates of stand-level forest attributes were obtained by taking a weighted mean of the values from all of the remaining grid cells within each stand using cell areas as weights.

Imputation (MSN)

MSN imputation was assessed as an alternative form of indirect estimation of forest attributes in this study. The distance metric used for MSN imputation had the following form (LeMay and Temesgen 2005; Eskelson et al. 2009):

$$\mathbf{d}_{ij} = (\boldsymbol{X}_i - \boldsymbol{X}_j)'\boldsymbol{\Gamma}\boldsymbol{\Lambda}^2\boldsymbol{\Gamma}'(\boldsymbol{X}_i - \boldsymbol{X}_j)$$

where $\boldsymbol{X}_i$ is the vector of standardized auxiliary variables for the ith target stand, $\boldsymbol{X}_j$ is the vector of standardized auxiliary variables for the jth reference stand, $\boldsymbol{\Gamma}$ is a matrix of standardized canonical coefficients for the X variables, and $\boldsymbol{\Lambda}^2$ is a diagonal matrix of squared canonical correlations. The set of X variables for this analysis consisted of a matrix of the LiDAR-derived auxiliary variables described in Table 2, whereas the set of Y variables was a matrix of the DE for all five attributes of interest for every small area. All calculations for MSN were performed using the "yalmpute" tool for R (Moeur et al. 1999; R Development Core Team 2008; Crookston and Finley 2007).

Multiple linear regression (MLR)

MLR models were developed both as an indirect estimation component for CP and as an initial basis for creating the Fay–Herriot models for EBLUP derivation. These models were designed to estimate stand-level values of the attributes of interest using stand-level LiDAR metrics as the independent variables. Because of the impracticality of reassessing the MLR models for each of the 500 simulation runs, transformations and variable selection were assessed using the full-sample DEs as the response values for each attribute. After examination of residual plots, Shapiro–Wilk test results, and quantile–quantile (q–q) plots, it was determined that log transformation of QDBH was beneficial to correct for heteroskedasticity. Supersets of important explanatory variables were selected using a subset regression technique that identifies the explanatory variables that create the best-fitting linear regression models according to Bayesian information criteria (BIC). This was performed using the "regsubsets" tool available in the "leaps" package for R (R Development Core Team 2008; Lumley 2008). The resulting output contained the information for a total of 70 possible supersets ranked by BIC. Any superset with a resulting model that had a variance inflation factor (VIF) score greater than 9.5 was automatically dropped from the final list. Of the final list, one superset was chosen for estimating each attribute. The supersets were then used to fit separate MLR models for each of the 500 simulation runs.

Composite prediction (CP) — Fay–Herriot models and EBLUP

Fay and Herriot (1979) models are a special case of linear mixed models (Rao 2003, p. 115; Slud and Maiti 2006). The purpose of using a linear mixed model is to incorporate random effects dependent on the small area of interest to explain random variation between small areas that cannot be explained by the fixed effects of the model. The Fay–Herriot models used in this study had the following format (Fay and Herriot 1979; Rao 2003, p. 77; You and Chapman 2006):

$$[1] \qquad \widehat{\boldsymbol{\theta}}_i = \boldsymbol{z}_i^{\mathrm{T}}\boldsymbol{\beta} + \boldsymbol{b}_i\boldsymbol{v}_i + \boldsymbol{e}_i$$

where the individual error effects $\boldsymbol{e}_i$ are iid $N(0,\boldsymbol{\psi}_i)$, the random area effects $\boldsymbol{v}_i$ are iid $N(0,\sigma_v^2)$, $\boldsymbol{z}_i$ is a vector of fixed area-level covariates for small area i, $\boldsymbol{\beta}$ is a vector of regression coefficients for the fixed effects of the model, and $\boldsymbol{b}_i$ is a known positive constant typically assumed to be $\boldsymbol{b}_i = 1$. The Fay–Herriot model can be motivated as follows. Let $\boldsymbol{\theta}_i$ be a parameter of interest from small area i. Assume that it is related to a set of variables, $\boldsymbol{z}_i$, though the following linear model:

$$[2] \qquad \boldsymbol{\theta}_i = \boldsymbol{z}_i^{\mathrm{T}}\boldsymbol{\beta} + \boldsymbol{b}_i\boldsymbol{v}_i$$

under the assumptions described in eq. 1. We assume that there is a direct estimator, $\widehat{\boldsymbol{\theta}}_i$, available for $\boldsymbol{\theta}_i$, so that

$$[3] \qquad \widehat{\boldsymbol{\theta}}_i = \boldsymbol{\theta}_i + \boldsymbol{e}_i$$

Combining eqs. 2 and 3 yields eq. 1.

As previously stated, the Fay–Herriot model is a special case of a linear mixed model in that there is only one observation per small area. This means that the random effects cannot be calculated using standard ML or REML procedures for individual variables as is done with a standard nested-error model. After selecting the auxiliary variables to

include in the linear model, there are three steps that must be completed to calculate an EBLUP via the Fay–Herriot model: (*i*) estimation of sampling error variance ($\boldsymbol{\psi}_i$), (*ii*) estimation of random error variance (σ_v^2) using the estimated $\boldsymbol{\psi}_i$, and (*iii*) calculate the area-level EBLUP ($\widehat{\boldsymbol{\theta}}_i^H$).

Sampling error variance estimation

There are two sources of error that must be accounted for to create an EBLUP, random-error variance σ_v^2 and sampling variance $\boldsymbol{\psi}_i$. This partitioning of the error requires prior knowledge of one source of error to estimate the other. Although the individual sampling variances $\boldsymbol{s}_i^2$ calculated directly from the ground data can be used as estimates of $\boldsymbol{\psi}_i$, it is often not desirable because of the instability that it can introduce into the composite estimator (Datta et al. 2005; Rivest and Vandal 2003; Ybarra and Lohr 2008). Therefore, $\widehat{\boldsymbol{\psi}}_i$ is typically a smoothed estimator based on a constant mean variance (V_e) and the area-level sample sizes (n_i). Smoothed estimates ($\widehat{\boldsymbol{\psi}}_i$) were calculated using the following equation (Williams 2007):

$$[4] \qquad \widehat{\boldsymbol{\psi}}_i = \frac{V_e}{n_i}$$

where V_e is a constant mean variance from the population and n_i is the sample size in stand *i*.

The constant V_e was calculated using a generalized variance function (GVF), which is a mathematical model that describes the relationship between the variance of a survey estimator and its expectation (Wolter 1985, p. 201). GVFs can take many different forms and are often used to derive weighted estimates of variance for complex populations. The simplest approach for estimation of a mean variance for the population in this study would be to take the variance of all of the observations within the population. Two major problems with this approach are as follow: (*i*) it assumes that all areas of the population are weighted equally, which is invalid given that the individual stands within the Clatsop vary greatly in size, and (*ii*) it ignores the segmentation of the population with regard to individual stands, which would likely cause an inflated variance estimate. The ultimate goal was to calculate a smoothed estimator that accounts for the varying weight of attribute values throughout the population of interest. After assessing a number of approaches, we used a form of weighted mean variance based on the size (ha) of each small area relative to the total size of the population. Therefore, V_e was calculated using the following equation:

$$[5] \qquad V_e = \frac{\sum_i^m \boldsymbol{a}_i \boldsymbol{s}_i^2}{\sum_i^m \boldsymbol{a}_i}$$

where $\boldsymbol{a}_i$ is the size (ha) of small area *i*, and *m* is the total number of stands within the population.

Random error variance estimation

After deriving smoothed variance estimates for the sampling errors, the next step was to estimate σ_v^2. A number of estimation procedures exist for the error variance of the Fay–Herriot model including the actual Fay–Herriot method based on method of moments (Fay and Herriot 1979; Prasad and Rao 1990; Wang and Fuller 2003; Ybarra and Lohr 2008), maximum likelihood (ML) (Datta and Lahiri 2000; Wang and Fuller 2003; Slud and Maiti 2006), and restricted maximum likelihood (REML) (Das et al. 2004). In this study, σ_v^2 was estimated using both the Fay–Herriot method and the ML method. Both methods involve an iterative process that converges to an estimate for both σ_v^2 and $\boldsymbol{\beta}$. These estimates were obtained via the Fay–Herriot method using the following iterative solution (Fay and Herriot 1979; Rao 2003, p. 118):

$$[6] \qquad \sigma_v^{2(a+1)} = \sigma_v^{2(a)} + \frac{1}{h'_*(\sigma_v^{2(a)})}[m - p - h(\sigma_v^{2(a)})]$$

where $\sigma_v^{2(a+1)} \geq 0$,

$$[7] \qquad h(\sigma_v^2) = \sum_i (\widehat{\boldsymbol{\theta}}_i - \boldsymbol{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}})^2/(\boldsymbol{\psi}_i + \sigma_v^2 \boldsymbol{b}_i^2)$$

$$[8] \qquad h'_*(\sigma_v^2) = -\sum_i \boldsymbol{b}_i^2 (\widehat{\boldsymbol{\theta}}_i - \boldsymbol{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}})^2/(\boldsymbol{\psi}_i + \sigma_v^2 \boldsymbol{b}_i^2)^2$$

and

$$[9] \qquad \widetilde{\boldsymbol{\beta}} = \left[\sum_{i=1}^m \boldsymbol{z}_i \boldsymbol{z}_i^{\mathrm{T}}/(\boldsymbol{\psi}_i + \sigma_v^2 \boldsymbol{b}_i^2)\right]^{-1} \left[\sum_{i=1}^m \boldsymbol{z}_i \widehat{\boldsymbol{\theta}}_i/(\boldsymbol{\psi}_i + \sigma_v^2 \boldsymbol{b}_i^2)\right]$$

where *a* is the iteration number, $h'_*(\sigma_v^2))$ is an estimated derivative of $h(\sigma_v^2))$, and $\widetilde{\boldsymbol{\beta}}$ is adjusted for each iteration.

Note that because an adjusted estimate of $\boldsymbol{\beta}$ is obtained through the process of estimating σ_v^2, even if the algorithm converges to a negative value for $\widehat{\sigma}_v^2$ (in which case, it is set to zero), the resulting EPLUP will be different from that obtained from standard MLR. The same holds true for the ML method for which estimates of σ_v^2 were obtained using the following iterative process (Prasad and Rao 1990; Rao 2003, p. 119):

$$[10] \qquad \sigma_v^{2(a+1)} = \sigma_v^{2(a)} + [\mathrm{I}(\sigma_v^{2(a)})]^{-1} s(\widetilde{\boldsymbol{\beta}}^{(a)}, \sigma_v^{2(a)})$$

where

$$[11] \qquad \mathrm{I}(\sigma_v^{2(a)}) = \frac{1}{2}\sum_{i=1}^m \frac{\boldsymbol{b}_i^4}{(\sigma_v^{2(a)}\boldsymbol{b}_i^2 + \boldsymbol{\psi}_i)^2}$$

and

$$[12] \qquad s(\widetilde{\boldsymbol{\beta}}^{(a)}, \sigma_v^{2(a)}) = -\frac{1}{2}\sum_{i=1}^m \frac{\boldsymbol{b}_i^2}{(\sigma_v^{2(a)}\boldsymbol{b}_i^2 + \boldsymbol{\psi}_i)} + \frac{1}{2}\sum_{i=1}^m \boldsymbol{b}_i^2 \frac{(\widehat{\boldsymbol{\theta}}_i - \boldsymbol{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}})^2}{(\sigma_v^{2(a)}\boldsymbol{b}_i^2 + \boldsymbol{\psi}_i)^2}$$

Both iterative methods were initiated at $\sigma_v^2 = 0$ and required less than 10 iterations before convergence. It is possible to obtain a negative estimate of σ_v^2 through these methods, especially if estimates of $\boldsymbol{\psi}_i$ are very high relative to $\widehat{\boldsymbol{\theta}}_i - \boldsymbol{z}_i^{\mathrm{T}}\widetilde{\boldsymbol{\beta}}$. However, this typically did not occur in the analysis.

EBLUP estimation

The EBLUP $\widehat{\theta}_i^H$ was derived using the following formula (Slud and Maiti 2006):

$$[13] \qquad \widehat{\theta}_i^H = \gamma_i \widehat{\theta}_i + (1 - \gamma_i) z_i^{\mathrm{T}} \widetilde{\beta}$$

where

$$[14] \qquad \gamma_i = \frac{\widehat{\sigma}_v^2}{\widehat{\sigma}_v^2 + \widehat{\psi}_i}$$

Notice that the EBLUP is actually a special case of a composite estimator that uses both the adjusted fixed effects of the Fay–Herriot model and the DE for the area of interest derived from the inventory data (Slud and Maiti 2006; Ybarra and Lohr 2008; You and Chapman 2006). A positive weight γ_i for the EBLUP is dependent on having a positive value for $\widehat{\sigma}_v^2$. In the case of a zero value for γ_i, the EBLUP simply reverts back to a prediction based on the population-level, multiple linear regression, excluding local information from the stand.

The primary purpose of calculating an EBLUP is to reduce bias in the estimator from a regression-based model for individual stands, via a weighting function based on sampling variance and variance of model random effects as is seen in eq. 14. Because, the random error variance was estimated using two different methods, two EBLUPs were calculated for each attribute and sampling intensity combination. The first (EBLUP_A) was calculated using the Fay–Herriot method of moments and the second (EBLUP_B) was calculated using the ML method. Before back transformation from log scale, all EBLUP estimates for QDBH were bias corrected using a factor of 0.5 times the mean squared error as calculated in Rao (2003, p. 117):

$$[15] \qquad \mathrm{MSE}(\widehat{\theta}_i^H) = g_{1i}(\widehat{\sigma}_v^2) + g_{2i}(\widehat{\sigma}_v^2)$$

where

$$[16] \qquad g_{1i}(\widehat{\sigma}_v^2) = \widehat{\sigma}_v^2 b_i^2 \psi_i / (\psi_i + \widehat{\sigma}_v^2 b_i^2) = \gamma_i \psi_i$$

and

$$[17] \qquad g_{2i}(\widehat{\sigma}_v^2) = (1 - \gamma_i)^2 z_i^{\mathrm{T}} \left[\sum_{i=1}^{m} z_i z_i^{\mathrm{T}} / (\psi_i + \sigma_v^2 b_i^2) \right]^{-1} z_i$$

Composite prediction (CP) — MLR and SP

Aside from the EBLUPs, two other composite estimators were analyzed in this study. The first (CP_A) was a composite between SP and DE, and the second (CP_B) was a composite between MLR and DE. Both estimators were developed using the following equation:

$$[18] \qquad \widehat{Y}_i^C = \widehat{\phi}_i \widehat{Y}_{i2} + (1 - \widehat{\phi}_i) \widehat{Y}_{i1}$$

where $\widehat{Y}_{i1}$ is the DE of the attribute for the ith stand, $\widehat{Y}_{i2}$ is the regression-based estimator of the ith stand, and $\widehat{\phi}_i$ is the weight calculated for the ith stand. The weights were calculated using a variation of the James–Stein method similar to that presented by Rao (2003, p. 58):

$$[19] \qquad \widehat{\phi}_i = \frac{\widehat{\psi}_i}{(\widehat{Y}_{i2} - \widehat{Y}_{i1})^2}$$

As with the EBLUP, it is possible to estimate the weights for the composite estimator using s_i^2. However, although more conventional, the estimation method for the composite weights shown in eq. 18 is considered to be less stable than that for EBLUP as it constitutes estimation of weights on an area-by-area basis. Therefore, $\widehat{\psi}_i$ was used instead of s_i^2 to add stability to CP_A and CP_B. As with EBLUP, this method essentially uses the information regarding one source of error to estimate the other, namely the mean squared errors (MSE) of $\widehat{Y}_{i1}$ and $\widehat{Y}_{i2}$. As such, eq. 19 could also be expressed as

$$[20] \qquad \widehat{\phi}_i = \frac{\widehat{\psi}_i}{\widehat{\psi}_i + \mathrm{MSE}_{i2}}$$

where

$$[21] \qquad \mathrm{MSE}_{i2} = (\widehat{Y}_{i2} - \widehat{Y}_{i1})^2 - \widehat{\psi}_i$$

Equations 20 and 21 illustrate that the MSE of the regression-based estimator (MSE_{i2}) is estimated by using MSE of DE, which for the purposes of this study is considered to be $\widehat{\psi}_i$. Therefore, the estimated weight for a small area is completely dependent on the size of $\widehat{\psi}_i$ relative to the squared difference between the DE and regression-based estimator for that particular small area. The primary goal in using CP via regression-based estimators is to balance the potential bias of the regression-based estimator with the instability of the DE. For CP_B, the regression component of the estimators were bias corrected using a factor of 0.5 times the regression mean squared error prior to back transformation before eqs. 18–21 were applied.

Validation

As previously stated, all residuals used in this study were calculated using the FS direct estimate of each attribute and stand as the observed value or "truth". Because most of the estimators were evaluated over 500 subsamples, the performance statistics needed to be calculated accordingly. Precision was assessed using relative root mean squared error (RRMSE) as presented by Rao (2003, p. 62), averaged over all the stands:

$$[22] \qquad \mathrm{RRMSE} = \frac{1}{m} \sum_{i}^{m} (\mathrm{RMSE}_i / \widehat{Y}_{i0})$$

with

$$[23] \qquad \mathrm{RMSE}_i = \sqrt{\frac{1}{R} \sum_{r=1}^{R} (\widehat{Y}_{iP} - \widehat{Y}_{i0})^2}$$

where RMSE_i is the root mean squared error for stand i, m is the number of stands, R is the number of iterations, $\widehat{Y}_{iP}$ is the predicted value for stand i, and $\widehat{Y}_{i0}$ is the full sample direct estimate for stand i. Similarly, overall bias was assessed using relative bias calculated as

$$[24] \qquad \mathrm{RB} = \frac{1}{m}\sum_{i=1}^{m}(\mathrm{RB}_i)$$

with

$$[25] \qquad \mathrm{RB}_i = \frac{1}{R}\sum_{r=1}^{R}\left(\frac{\widehat{\boldsymbol{Y}}_{iP} - \widehat{\boldsymbol{Y}}_{i0}}{\widehat{\boldsymbol{Y}}_{i0}}\right)$$

where RB_i is the relative bias for stand i. Because SP was not resampled, the same statistics were calculated based on the one estimate for each stand. This was necessary to properly compare the performance of SP with the other estimators that were dependent on subsampling.

Results and discussion

The initial MLR models created through subsets regression using stand-level attributes and LiDAR metrics varied somewhat from the plot-level models presented in Goerndt et al. (2010). The five models for each forest attribute were fairly consistent in terms of the LiDAR metrics that were chosen in that each model used a similar number of LiDAR height and cover metrics. Although the MLR models performed reasonably well in terms of indirect estimation of the forest attributes, they had a tendency to underestimate for stands with high attribute values. Not unexpectedly, this feature was most prevalent for attributes that have lower correlation with LiDAR cloud metrics, e.g., tree density and BA. Although bias in regression analyses can be caused by many factors, in this study, it is most likely caused by a weakening of the relationship between LiDAR cloud metrics and forest attributes for stands with high vegetative density. This hypothesis is reinforced by the fact that underestimation usually occurred for the same group of stands regardless of the attribute being estimated. This denotes a drawback with using LiDAR data for forest attribute estimation, as many point-density LiDAR cloud metrics such as canopy cover and transparency can become less informative as forest density increases and the laser does not penetrate the forest canopy as well. As will be seen later, this was an important factor in assessing the performance of many of the estimators used in this study.

The computation of EBLUPs required the most information of any method in the analysis, necessitating estimation of the sampling variance of the direct estimator and the variance of the random effects. The estimates of $\widehat{\sigma}_v^2$ using the EBLUP_A and EBLUP_B methods were virtually identical and, therefore, so were the EBLUP weights and estimates of the variables of interest. Consequently, EBLUP_A and EBLUP_B will be referred to jointly as EBLUP for the remainder of this paper.

As shown in eq. 5, the weighting strategy for both CP_A and CP_B was different from that of EBLUP in that it accounted for the proportion of sampling variance to the squared difference between the synthetic-type estimator and the DE separately for each observation. Although not as stable as EBLUP, this weighting strategy provided more flexibility in composite estimation as the weight for each small area was not dependent on a constant value such as $\widehat{\sigma}_v^2$.

Because there were five different forest attributes included in the canonical correlation for MSN, there were five canonical covariates (axes) created for determining distance. The analysis for each method indicated that four of the axes were highly significant. CuVol, TpHt, and BA had the highest canonical correlation with the auxiliary data, whereas density and QDBH had the lowest. Performance statistics (RRMSE, RB) for each attribute of interest by estimation method and sampling intensity are given in Tables 3 and 4.

General comparison

Direct (DE) and indirect estimation

Direct estimators are design unbiased, whereas indirect or composite estimators can be biased due to the potential bias introduced by the model-based components. Consequently, comparisons between DE and other estimators are restricted to precision (RRMSE) and not bias (RB). Note that DE was considerably more precise than SP and MSN regardless of sample size (Table 3).

One noticeable characteristic is the poor overall performance of SP. SP yielded the highest RRMSE values for all of the attributes. These results show how sensitive models relating forest attributes and LiDAR metrics can be when applied to a different population, even within the same region. Of course, unlike the other estimators, SP did not change across the sampling intensities as it was not dependent on ground data from within the Clatsop. In terms of indirect estimation, the performance of MSN was far superior to that of SP in both precision and bias. Whether the superior performance was due to the method itself or to a better fit based on data from the study region cannot be determined from this study. The only attribute for which there was a similarity in precision of prediction between SP and MSN was QDBH, and that was only for 10% sampling intensity.

Composite prediction

The overall results indicate that EBLUP and CP_B are superior to CP_A in terms of RRMSE and RB. The results in Tables 3 and 4 also show how CP_B and EBLUP compare with DE based on the reduced sample sizes. Note that all CP methods were a substantial improvement over indirect estimation with regard to precision. EBLUP produced lower RB values than CP_B at low sampling intensities for all attributes except BA and Ht. Table 3 shows that with the exception of CuVol, CP_B and EBLUP yielded higher precision than DE at small sample sizes (10%–20%) and lower precision at higher sample sizes (30%–50%). The fact that EBLUP typically yielded higher precision than CP_B indicates that the weighting strategy used for EBLUP can usually provide greater stability for the estimator in terms of precision.

Note that MSN often achieved slightly lower bias than EBLUP, CP_A, and CP_B, especially at low sampling intensities. However, EBLUP, CP_B, and CP_A are far superior to SP and MSN in terms of precision (Table 3). One drawback to the CP methods for application is they all require a ground sample in the small area of interest. This is a characteristic not shared by SP and MSN.

Stand-level comparison (bias)

Although the performance statistics shown in Tables 3 and 4 are informative as to the overall performance of the estimators across the stands, it does not provide the whole picture with regard to individual stands. When using SAE, the main

Table 3. Estimated precision (RRMSE) from final validation of all estimation methods for each attribute of interest by sampling intensity (%). See Table 1 for attribute descriptions.

	10%	20%	30%	40%	50%
Density (trees/ha)					
DE	38.9	26.9	20.8	16.6	13.9
SP*	64.7	64.7	64.7	64.7	64.7
MSN	60.1	47.9	42.1	38.8	36.6
CP_A	40.8	30.2	24	19.7	16.7
CP_B	29.3	23.4	20	17.6	15.8
EBLUP	27.7	22.5	19.2	16.5	15.1
BA (m^2/ha)					
DE	24.1	16.7	12.8	10.2	8.6
SP*	58.7	58.7	58.7	58.7	58.7
MSN	38.8	31.2	27.6	25.4	24.4
CP_A	31.9	21.9	16.7	13.2	11
CP_B	18.2	15	13.3	11.9	11
EBLUP	17.3	14.3	12.3	10.7	9.4
CuVol (m^3/ha)					
DE	25.8	17.7	13.7	10.9	9.2
SP*	84.3	84.3	84.3	84.3	84.3
MSN	43.4	36	32.1	30.5	29.5
CP_A	42.1	27.8	20.8	16.2	13.2
CP_B	22.7	18.9	17	15.4	14.3
EBLUP	18.3	15	13	11.1	9.9
QDBH (cm)					
DE	19.6	13.6	10.4	8.4	7
SP*	27.8	27.8	27.8	27.8	27.8
MSN	30	25.4	23	21.8	20.7
CP_A	22.1	16.6	13.4	11.2	9.5
CP_B	16.3	12.6	10.5	9.1	8.1
EBLUP	14.9	11.9	9.6	8	6.8
Ht (m)					
DE	16.2	11.3	8.7	6.9	5.7
SP*	26.7	26.7	26.7	26.7	26.7
MSN	23.6	19.7	17.8	16.9	16.2
CP_A	19.8	14.2	11.1	9	7.5
CP_B	12.2	9.8	8.5	7.5	6.8
EBLUP	11.4	9.3	8	6.8	6

*No simulation

Table 4. Estimated RB from final validation of all estimation methods for each attribute of interest by sampling intensity (%). See Table 1 for attribute descriptions.

	10%	20%	30%	40%	50%
Density (trees/ha)					
SP*	41.2	41.2	41.2	41.2	41.2
MSN	–7.8	–6.7	–6.5	–6.3	–6.1
CP_A	–10.3	–5.4	–3.7	–2.6	–1.8
CP_B	–10.9	–8.1	–7.2	–6.5	–5.7
EBLUP	–8.7	–6	–5.1	–4.1	–3.5
BA (m^2/ha)					
SP*	57.5	57.5	57.5	57.5	57.5
MSN	–4.2	–3.4	–3.2	–2.9	–2.8
CP_A	–18.9	–11.7	–8.6	–6.6	–5.4
CP_B	–3.9	–3.2	–2.9	–2.5	–2.2
EBLUP	4.7	3.6	2.9	2.4	1.9
CuVol (m^3/ha)					
SP*	84.1	84.1	84.1	84.1	84.1
MSN	–5.5	–4.8	–4.3	–4	–3.9
CP_A	–28.2	–16.8	–12.1	–9.2	–7.4
CP_B	–5.9	–5	–4.9	–4.8	–4.6
EBLUP	–0.98	–0.48	–0.5	–0.36	–0.4
QDBH (cm)					
SP*	–20.4	–20.4	–20.4	–20.4	–20.4
MSN	–1.9	–1.6	–1.5	–1.3	–1.2
CP_A	7.3	5.7	4.7	3.9	3.3
CP_B	–4.2	–3.2	–2.5	–2.1	–1.9
EBLUP	3.4	1.7	1.3	0.93	0.69
Ht (m)					
SP*	26.7	26.7	26.7	26.7	26.7
MSN	–1.6	–1.4	–1.4	–1.3	–1.5
CP_A	–11.2	–7.5	–5.8	–4.8	–3.9
CP_B	–0.2	–0.34	–0.44	0.35	–0.28
EBLUP	1.5	0.9	0.72	0.66	0.6

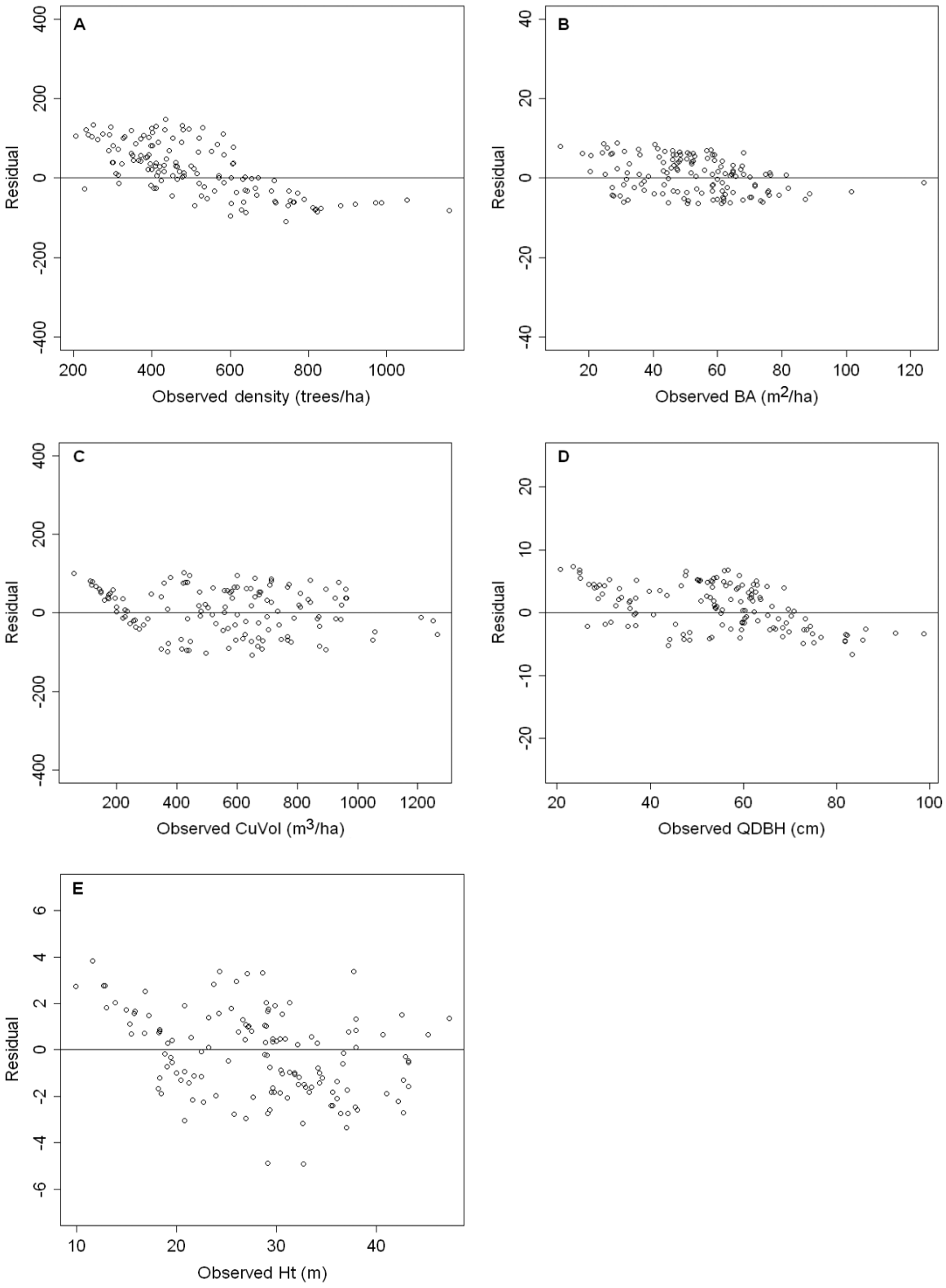
*No Simulation

objective is to obtain estimates for each of the small areas. Therefore, the best way to assess the performance of small-area estimators such as CP_B and EBLUP is to observe the bias for each small area. The stand-level mean residuals for the two best estimators (CP_B, EBLUP) were compared for a sampling intensity at which EBLUP and CP_B were more precise than DE (20%) using residual plots. As with all validation in this study, these residuals were calculated as the full sample estimate minus the predicted value from the estimator of interest. The values used for assessment were stand-level means of the residuals across the 500 subsamples. Figures 2 and 3 illustrate residual plots for CP_B and EBLUP for each attribute at 20% sampling intensity. The attributes are ordered by their degree of correlation with standard LiDAR cloud metrics, with density being the lowest and Ht the highest.

One of the most noticeable characteristics of the estimates shown in Figs. 2 and 3 is the poor performance of both methods for prediction of tree density. Both CP_B and EBLUP tend to overestimate for stands with high density values and underestimate for stands with low density values. This was not directly the result of a similar tendency in DE. This characteristic of EBLUP and CP_B was driven by the poor fit of the model relating tree density to LiDAR metrics. In an assessment of the MLR models developed for calculating CP_B and EBLUP, the models for density had the lowest adjusted R^2 value (20.9%) of any linear model in this study. Previous studies such as Goerndt et. al. (2010) have also shown that estimation of density using area-level LiDAR metrics was difficult because of the lack of correlation between density and canopy height characteristics. This problem seems less severe for CP_B because the weighting strategy described by eq. 19 resulted in CP_B relying heavily on DE. Ultimately, DE provided more precise estimates of density than either CP_B or EBLUP.

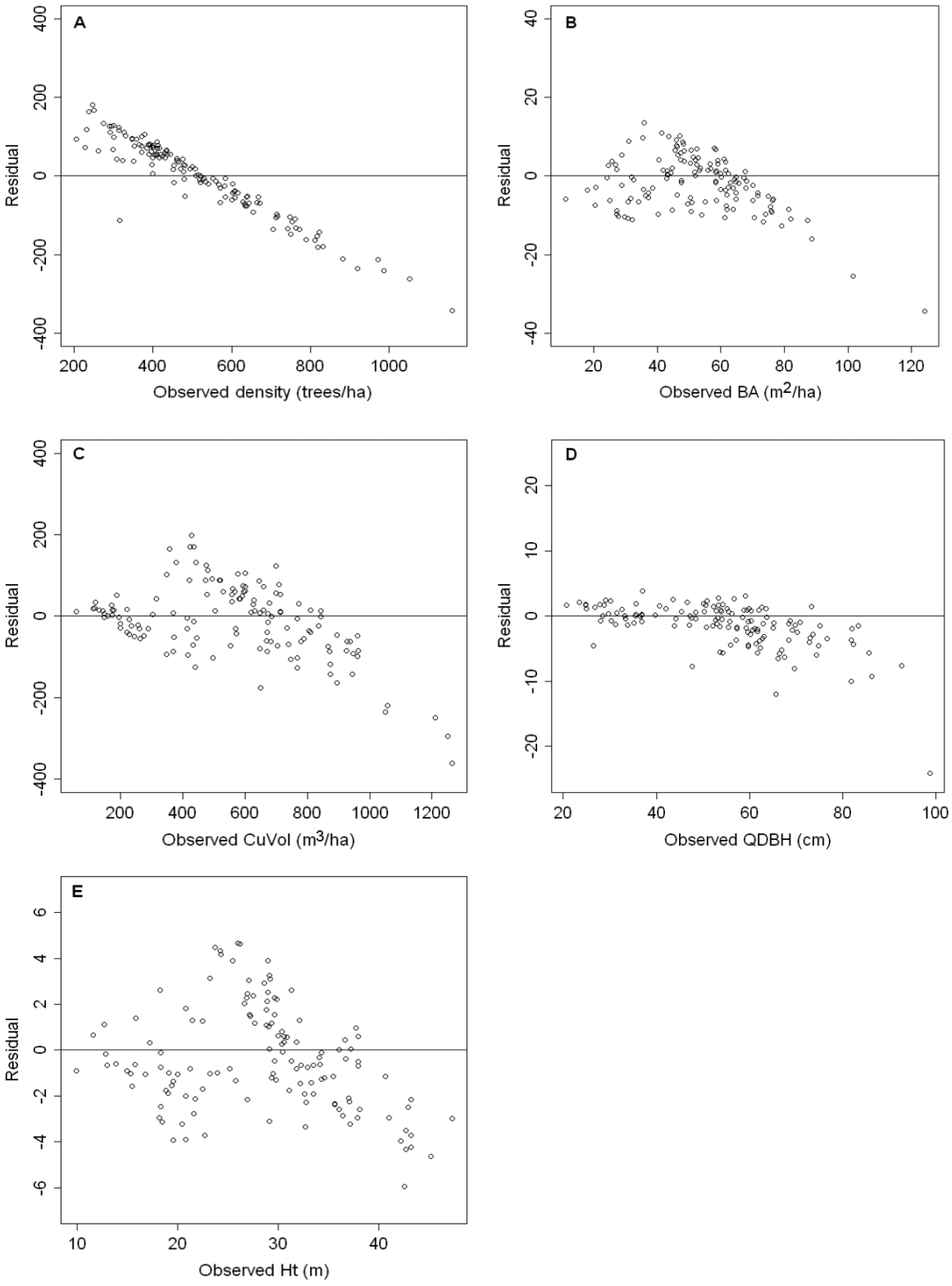
Although the difference in RB is very subtle for BA and Ht, it is still apparent from Figs. 2 and 3 that CP_B is superior to EBLUP in terms of bias. As seen with BA and CuVol, CP_B tended to rely on DE more than EBLUP. However, CP_B weighed the DE component more heavily for many observations; this was not true for all observations, because the weighting method for CP_B is not dependent on a constant value of $\hat{\sigma}_v^2$ and can therefore take on a wide range of

Fig. 2. Residual plots of CP_B at 20% sampling intensity for (A) density, (B) BA, (C) CuVol, (D) QDBH, and (E) Ht. See Table 1 for attribute descriptions.



values depending on the stand. Although more stable, EBLUP is more limited with regard to the range of weights that can be used for estimation. As long as $\widehat{\sigma}_v^2$ is a nonzero value, every stand will be assigned a weight for EBLUP whether it actually needs one or not. Through the versatility of its weighting method, CP_B was better able to achieve the proper balance between the regression synthetic component and DE with regard to bias (Tables 3 and 4).

Fig. 3. Residual plots of EBLUP at 20% sampling intensity for (A) density, (B) BA, (C) CuVol, (D) QDBH, and (E) Ht. See Table 1 for attribute descriptions.



Although CP_B tended to give higher weight to the DE component when compared with EBLUP, this was not the case for all subsamples. To assess this, two subsamples for 20% sampling intensity were chosen from the 500 representing a case in which RB was higher for CP_B than DE and one in which it was lower. RB was used simply as a way to identify cases in which the residuals were similar between CP_B and DE and cases in which they were not. As with Figs. 2–3, the residuals were calculated as the full sample estimate minus the predicted value from the estimator of inter-

Fig. 4. Residual plots of selected 20% sample with relative bias of (A) DE less than that of (B) CP_B. BA, basal area.

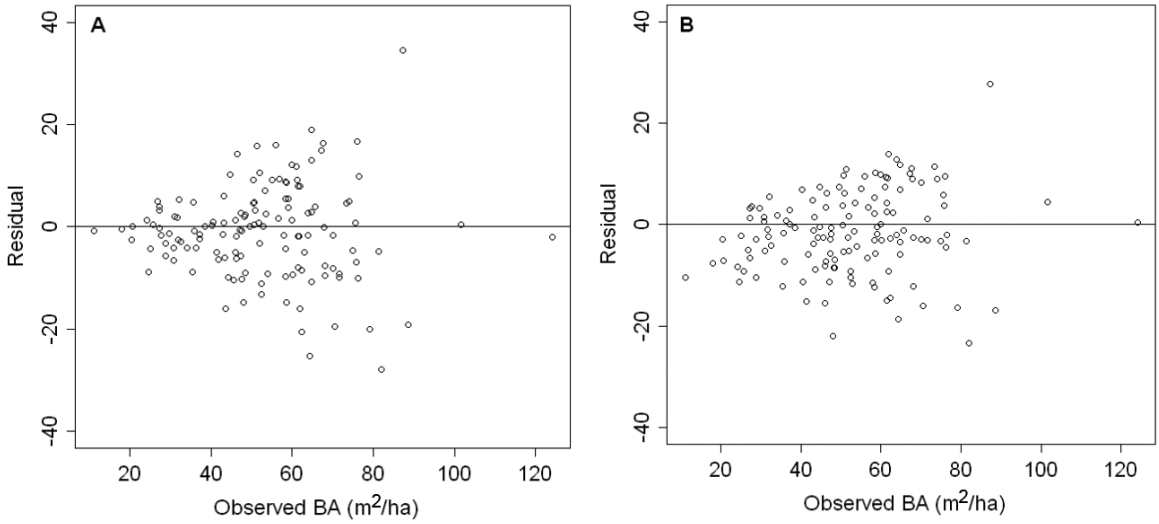
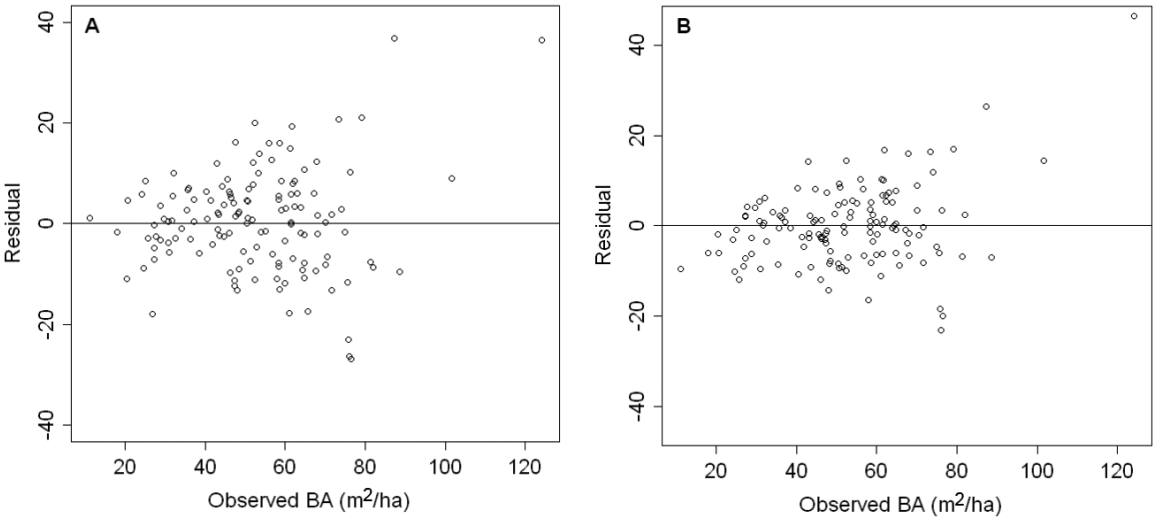


Fig. 5. Residual plots of selected 20% sample with relative bias of (A) DE greater than that of (B) CP_B. BA, basal area.



est. Figures 4 and 5 illustrate BA residual plots of CP_B and DE for the chosen subsamples at 20% sampling intensity with the full sample estimate set at zero.

In Fig. 4, CP_B has an obvious tendency for overestimation, denoting either a low dependency on the DE component, a weak correlation in the MLR component, or both. However, the subsample illustrated in Fig. 5 shows the high degree of similarity between the residuals for CP_B and DE, indicating that the DE component was probably weighted very heavily. The success of composite prediction depends on how well the LiDAR-based MLR component fits the particular small sample that is available and the weight that is put on the DE component.

Although each original MLR variable superset was developed using the full sample from the population of interest, there are obvious bias issues, particularly for stands with very high attribute values. More importantly, because the models did not change from one subsample to another in terms of auxiliary variables used, there were definitely subsamples for which the general model was less precise than for others.

Conclusion

Estimating forest attributes for small areas within larger populations of interest is a primary focus of forest researchers and practitioners. Estimation of forest attributes in forest stands using within-population information, whether it is through composite estimators, linear models, or imputation, has gained prominence over the last few decades. With rapid advances in the acquisition and processing of remotely sensed data such as LiDAR, these methods have reached new standards in terms of accuracy and precision of forest attribute prediction. This study has demonstrated how the integration of LiDAR metrics and SAE techniques can facilitate precise estimation of stand-level forest attributes by efficiently using available information from within the population of interest.

Of the composite estimators assessed in this study, EBLUP was superior to CP_B in terms of overall precision and bias (Tables 3 and 4). On a stand-by-stand basis, CP_B was less biased for stands with higher attribute values than EBLUP,

particularly for density. Although CP_B provided less biased estimates at the stand level than EBLUP, DE was still superior to all regression-based estimators for density. For all other attributes, EBLUP provided more precise estimation than CP_B and DE, though the advantage of using EBLUP was much greater for QDBH and CuVol than for BA and Ht. This study has shown that although SAE through composite prediction can often be beneficial, its use needs to be tailored specifically to the attribute of interest based on the auxiliary information being used. Ultimately, the performance of any form of SAE that relies on a regression component such as the ones in this study depends greatly on the strength of the relationship between the specific attributes being estimated and the auxiliary information that is used.

Acknowledgements

We are grateful to Emmor Nile and Dave Enck of the Oregon Department of Forestry for providing both LiDAR and ground data for use in this study. We thank Michael Wing for his expertise with regard to GIS. We also thank Greg Latta, Matt Gregory, and Emilie Grossmann for their superb advice regarding mapping, imputation, and many other aspects of this study. Funding for this project was provided by Forest Inventory and Analysis (FIA), U.S. Department of Agriculture Forest Service.

References

Breidenbach, J., Glaser, C., and Schmidt, M. 2008. Estimation of diameter distributions by means of airborne laser scanner data. Can. J. For. Res. **38**(6): 1611–1620. doi:10.1139/X07-237.

Chen, Q., Baldocchi, D., Gong, P., and Maggi, K. 2006. Isolating individual trees in a savanna woodland using small footprint LiDAR data. Photogramm. Eng. Remote Sensing, **72**(8): 923–932.

Costa, A., Sattora, A., and Ventura, E. 2003. An empirical evaluation of small area estimators. SORT, **27**: 113–135.

Costa, A., Sattora, A., and Ventura, E. 2004. Using composite estimators to improve both domain and total area estimation. SORT, **28**: 69–86.

Crookston, N.L., and Finley, A.O. 2007. yalmpute: an R Package for k-NN imputation. J. Stat. Softw. **23**(10): 1–16.

Das, K., Jiang, J., and Rao, J.N.K. 2004. Mean squared error of empirical predictor. Ann. Stat. **32**(2): 818–840. doi:10.1214/009053604000000201.

Datta, G., and Lahiri, P. 2000. A unified measure of uncertainty of estimated best linear predictors in small area estimation problems. Statist. Sinica, **10**: 613–627.

Datta, G.S., Rao, J.N.K., and Smith, D.D. 2005. On measuring the variability of small area estimators under a basic area level model. Biometrika, **92**(1): 183–196. doi:10.1093/biomet/92.1.183.

Eskelson, B.N.I., Temesgen, H., and Barrett, T. 2009. Estimating current forest attributes from paneled inventory data using plot-level imputation: a study from the Pacific Northwest. For. Sci. **55**(1): 64–71.

Fay, R.E., III, and Herriot, R.A. 1979. Estimates of income for small places: an application of James–Stein procedures to census data. J. Am. Stat. Assoc. **74**(366): 269–277. doi:10.2307/2286322.

Goerndt, M.E., Monleon, V., and Temesgen, H. 2010. Relating forest attributes with area-based and tree-based LiDAR metrics for western Oregon. West. J. Appl. For. **25**: 105–111.

Heady, P., Clarke, P., Brown, G., Ellis, K., Heasman, D., Hennell, S., Longhurst, J., and Mitchell, B. 2003. Small area estimation project report. Model-based Small Area Estimation Series, Office of National Statistics, London.

LeMay, V., and Temesgen, H. 2005. Comparison of nearest neighbor methods for estimating basal area and stems per hectare using aerial auxiliary variables. For. Sci. **51**(2): 109–119.

Lumley, T. 2008. Leaps: regression subset selection. R package version 2.7. Available at http://cran.r-project.org/web/packages/leaps/index.html.

McGaughey, R. 2008. FUSION/LDV: software for LiDAR data analysis and visualization. FUSION version 2.65. Available at http://forsys.cfr.washington.edu/fusion/FUSION_manual.pdf.

Means, J.E., Acker, S.A., Fitt, B.J., Renslow, M., Emerson, L., and Hendrix, C.J. 2000. Predicting forest stand characteristics with airborne scanning lidar. Photogramm. Eng. Remote Sensing, **66**(11): 1367–1371.

Moeur, M. 2000. Extending stand exam data with most similar neighbor inference. *In* Proceedings of the Society of American Foresters National Convention, 11–15 September 1999, Portland, Oregon. Society of American Foresters, Bethesda, Maryland.

Moeur, M., Crookston, N., and Renner, D. 1999. Most similar neighbor, release 1.0. User's manual. USDA Forest Service, Rock Mountain Research Station, Moscow, Idaho.

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.

Oregon Department of Forestry. 2008. Stand level inventory field guide. State Forests Management Program, Oregon Department of Forestry, Salem, Oregon.

Petrucci, A., Pratesi, M., and Salvati, N. 2005. Geographic information in small area estimation: small area models and spatially correlated random area effects. Statistics in Transition, **3**(7): 609–623.

Popescu, S.C. 2007. Estimating biomass of individual pine trees using airborne lidar. Biomass Bioenergy, **31**(9): 646–655. doi:10.1016/j.biombioe.2007.06.022.

Popescu, S.C., Wynne, R.H., and Nelson, R.E. 2003. Measuring individual tree crown diameter with lidar and assessing its influence on estimating forest volume and biomass. Can. J. Rem. Sens. **29**(5): 564–577.

Prasad, G.N., and Rao, J.N.K. 1990. The estimation of mean squared error of small area estimators. J. Am. Stat. Assoc. **85**(409): 163–171. doi:10.2307/2289539.

R Development Core Team. 2008. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rao, J.N.K. 2003. Small area estimation. Wiley Series in Survey Methodology, Hoboken, New Jersey.

Rivest, L.P., and Vandal, N. 2003. Mean squared error estimation for small areas when the small area variances are estimated. *In* Proceedings of the International Conference of Recent Advanced Survey Sampling. *Edited by* J.N.K. Rao. Laboratory for Research in Statistics and Probability in Canada, Ottawa, Ontario. pp. 197–206.

Rubin, D.B. 1976. Inference and missing data. Biometrika, **63**(3): 581–592. doi:10.1093/biomet/63.3.581.

Slud, E.V., and Maiti, T. 2006. Mean-squared error estimation in transformed Fay–Herriot models. J. R. Stat. Soc. Ser. B (Stat. Methodol.), **68**(2): 239–257. doi:10.1111/j.1467-9868.2006.00542.x.

Strunk, J.L. 2008. Two-stage forest inventory with lidar on the Fort Lewis military installation. M.S. thesis, University of Washington, Seattle, Washington.

Van Deusen, P.C. 1997. Annual forest inventory statistical concepts with emphasis on multiple imputation. Can. J. For. Res. **27**: 379–384. doi:10.1139/cjfr-27-3-379.

Wang, J., and Fuller, W.A. 2003. The mean square error of small area predictors constructed with estimated area variances. J. Am. Stat. Assoc. **98**(463): 716–723. doi:10.1198/016214503000000620.

Williams, A.N. 2007. Fay–Herriot small area estimation in the survey of business owners. M.A. thesis, University of Maryland, College Park, Maryland.

Wolter, K. 1985. Introduction to variance estimation. Springer Series in Statistics, New York.

Ybarra, L., and Lohr, S. 2008. Small area estimation when auxiliary information is measured with error. Biometrika, **95**(4): 919–931. doi:10.1093/biomet/asn048.

You, Y., and Chapman, B. 2006. Small area estimation using area level models and estimated sampling variances. Surv. Methodol. **32**: 97–103.