Edge-Tree Correction for Predicting Forest Inventory Attributes Using Area-Based Approach With Airborne Laser Scanning

Petteri Packalen, Jacob L. Strunk, Juho A. Pitkänen, Hailemariam Temesgen, and Matti Maltamo

Abstract—**We describe a novel method to improve the correspondence between field and airborne laser scanning (ALS) measurements in an area-based approach (ABA) forest inventory framework. An established practice in forest inventory is that trees with boles falling within a fixed border field measurement plot are considered "in" trees; yet their crowns may extend beyond the plot border. Likewise, a tree bole may fall outside of a plot, but its crown may extend into a plot. Typical ABA approaches do not recognize these discrepancies between the ALS data extracted for a given plot and the corresponding field measurements. In the proposed solution, enhanced ABA (EABA), predicted tree positions, and crown shapes are used to adjust plot and grid cell boundaries and how ALS metrics are computed. The idea is to append crowns of "in" trees to a plot and cut down "out" trees, then EABA continues in the traditional fashion as ABA. The EABA method requires higher density ALS data than ABA because improvement is obtained by means of detecting individual trees. When compared to typical ABA, the proposed EABA method decreased the error rate (RMSE) of stem volume prediction from 23.16% to 19.11% with 127 m^2 plots and from 19.08% to 16.95% with 254 m^2 plots. The greatest improvements were obtained for plots with the largest residuals.**

Index Terms—**Forestry, remote sensing.**

I. Introduction

DISCRETE-RETURN lidar data acquired by small footprint airborne laser scanning (ALS) have increased in its popularity for forest inventories [1]. Research has focused on two approaches to predict forest attributes with ALS: individual tree crown delineation (ITD) and area-based approach (ABA) methods. As the names imply, these methods operate at different scales. In ITD, the idea is to first delineate individual tree crowns and then to predict tree level attributes. Many algorithms have been proposed to delineate individual trees (see reviews by Vauhkonen *et al.* [2] and Kaartinen *et al.* [3]). Most algorithms use canopy height models (CHM) in which pixel values correspond to height at above ground level. Trees are assumed to be peaks in the CHMs and tree crowns are delineated by segmentation. Since ITD applications typically link field-measured tree attributes to ALS data, they require accurate tree level positioning. Many ITD research articles address tree height only, but in real world inventories other tree attributes are needed too. Takahashi *et al.* [4], for instance, predicted stem volume and Vauhkonen *et al.* [5] simultaneously imputed tree species, diameter, height, and volume.

In ABA, plot-level ALS metrics are related to plot-level field data. Plot-level field data are obtained by measuring individual trees from sample plots and summing (or in some cases, averaging) trees' values for each plot. Plot-level ALS metrics are statistics calculated for the ALS point heights above the predicted ground surface elevations. Most ALS metrics are statistics which reflect either height or density information. Common ALS metrics are, e.g., height percentiles (height) and the proportion of echoes above 2 m (density). Normally, metrics are computed by echo categories although "only" and "first of many" echoes contain most relevant height and density information. Plot attributes can be estimated one by one, e.g., using regression models [6], or all the attributes can be estimated at once, e.g., using nearest neighbor imputation [7]. ABA always requires accurate positioning of sample plots.

In both ITD and ABA, predictions are made wall-to-wall across the forest, but there are fundamental differences in the approaches. In ITD, the inventory is performed on a tree-object basis, and the result is a map of predicted tree locations with corresponding attribute predictions. Attributes from tree-objects are then aggregated to the desired spatial units, commonly stands. In ABA, predictions are performed for a grid where the cell area approximately corresponds to the sample plot area. As with ITD, ABA predictions are typically aggregated to larger units of area such as forest stands. Thus, in ABA the purpose is not to predict individual tree attributes, but to extract ALS points for trees within appropriate areal units (sample plots or grid cells) and use them to infer areal forest attributes (e.g., stem volume ha^{-1}).

An established practice in forest inventory is to designate a tree as falling within a sample plot if the center of the bole is inside the plot bounds [8]. The same approach is used with ABA to select ALS points for a plot or grid cell—points are designated as falling within a plot or grid cell if their coordinates

Manuscript received May 09, 2014; revised November 24, 2014; accepted January 30, 2015. Date of publication March 01, 2015; date of current version March 27, 2015.

P. Packalen and M. Maltamo are with the Faculty of Science and Forestry, School of Forest Sciences, University of Eastern Finland, 80101 Joensuu, Finland (e-mail: petteri.packalen@uef.fi; matti.maltamo@uef.fi).

J. L. Strunk is with Washington State Department of Natural Resources, Olympia, WA 98504-7000 USA (e-mail: jacob.strunk@dnr.wa.gov).

J. A. Pitkänen is with Natural Resources Institute Finland (Luke), 80100 Joensuu, Finland (e-mail: juho.pitkanen@luke.fi).

H. Temesgen is with the Department of Forest Engineering, Resources and Management, College of Forestry, Oregon State University, Corvallis, OR 97331 USA (e-mail: temesgen.hailemariam@oregonstate.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSTARS.2015.2402693

TABLE I
MEAN, MINIMUM, AND MAXIMUM OF PLOT ATTRIBUTES BY PLOT SIZE

	Plot size 127 m^2			Plot size 254 m^2		
	Mean	Min	Max	Mean	Min	Max
BWD (cm)	21.0	11.2	33.92	21.7	12.3	32.5
HDOM (m)	19.74	9.9	33.4	20.2	7.5	33.6
Stem number ($n\,ha^{-1}$)	1305	392	2981	1314	393	3262
Volume ($m^3\,ha^{-1}$)	223.2	72.0	486.6	226.1	69.5	535.6

BWD, basal area-weighted mean diameter; HDOM, dominant height.

fall within the plot or grid cell. However, this approach to selecting ALS points results in edge tree discrepancies with ABA because some of the crown from "out" trees extend into the plot ALS data (type 1 discrepancy), and crown from "in" trees extend outside of the plot ALS data (type 2 discrepancy). Both types of discrepancies decrease the observed relationship between ALS metrics and plot attributes. The magnitude of the errors is affected by plot size, as the proportion of near plot-edge trees increases on average as plot size decreases. Nelson *et al.* [9] showed that estimates of biomass and volume were affected by the type 1 discrepancies. The authors used thin 5 m wide sample plots, which resulted in a high relative edge to area ratio. Recently, Næsset *et al.* [10] also pointed out the issue with type 1 discrepancies when estimating change in forest biomass from multitemporal ALS datasets.

To the best of our knowledge, Mascaro *et al.* [11] is the only article where the edge-tree discrepancy has been addressed. They addressed the discrepancy by spatially mapping trees' carbon content using their crown footprints relative to the plot borders. The steps used to edge-correct carbon values depended only upon their field measurements and allometric models. First, they estimated crown radii from tree diameters using allometric models, and then the trees' predicted carbon content was distributed uniformly over the predicted crown areas. Finally, crown-distributed carbon predictions (small cells) were overlaid with plot borders. The amount of mapped carbon falling within a plot served as the edge-corrected carbon value for the plot. The remaining ABA steps were then applied normally with the edge-corrected carbon values used for model development. The study provides an indication of the magnitude of edge-tree errors, but in general it cannot be used in forest inventory because it does not follow the established definition of whether trees are "in" or "out" of the plot depending upon bole location.

In this study, we present a method to account for edge-tree discrepancy in ABA. We follow the traditional definition of whether trees are "in" or "out" of the plot depending upon bole location, but use methodologies from ITD to identify and correct type 1 and 2 errors. The aim is to improve the performance of ABA by considering the predicted crown attributes of edge-trees.

II. MATERIAL

A. Study Area and Field Data

The study area is a boreal managed forest area in eastern Finland (62°31′N, 30°10′E). The field measurements were carried out in the summer of 2010. A total of 79 field plots were placed purposively (nonrandom) to reflect species and size variation for the study area. The allocation of field plots was based on development class and dominant tree species. Scots pine (Pinus sylvestris L.) is the dominant tree species representing about 75% of volume and the remainder consists of Norway spruce (Picea abies [L.] Karst.) and a mixture of native deciduous species.

This study used field measurements from concentric 6.37 m (127 m^2 area) and 9 m (254 m^2 area) radius circular plots. These data were extracted from larger 20 × 20 and 30 × 30 m^2 stem-mapped plots. We will refer to these as 127 and 254 m^2 plots, respectively. Field measured trees were positioned relative to predicted ITD tree locations that were confirmed in the field. ITD detected trees not corresponding to field trees were excluded whereas trees not detected by the ITD algorithm were positioned using angle and distance measurements to nearby ITD detected trees. The coordinates for the undetected trees were estimated using the least squares adjustment method described by Korpela *et al.* [12]. A summary of plot-level attributes is provided in Table I. For all trees with either DBH exceeding 4 cm or height exceeding 4 m, the diameter at breast height (DBH), height, and tree species were recorded. A tree's DBH was estimated as the average of its maximum diameter, and the diameter perpendicular to its maximum diameter. The volumes of individual trees were calculated as a function of DBH and tree height using the species-specific models developed by Laasasenaho [13]).

B. ALS Data

ALS data were collected on June 26, 2009 using an Optech ALTM Gemini laser scanning system. The nominal pulse density was approximately 12 pulses/m^2. The test site was scanned from an altitude of approximately 600 m above ground level, with a field of view of 26° and side overlap between transects of 55%. Pulse repetition frequency was set to 125 kHz. Side overlap of 55% means that each location is covered by at least two flight lines. This configuration was used to maximize the probability that trees have ALS hits from two sides. This reduces the occurrence of shadow areas in which tree foliage obstructs the path of the laser from reaching behind trees.

A digital terrain model (DTM) was constructed by first classifying points as ground and nonground hits according to the approach described by Axelsson [14]. A raster DTM of 0.5 m spatial resolution was then obtained by interpolation using Delaunay triangulation. Heights above ground (dZ) for ALS

points were calculated by differencing their elevations above the ellipsoid from corresponding DTM elevations.

III. Methods

A. Overview

The basis for our strategy is to predict tree positions and their crown delineations, and then modify plot and cell areas using predicted tree positions and crowns. We will refer to this method as enhanced ABA (EABA). Modified plots are used in the modeling stage, and correspondingly modified cells are used when a model is applied for an entire area. The method requires that ALS data is suitable for individual tree delineation (i.e., point density must be higher than in normal ABA). EABA has five additional steps beyond typical ABA processing including: 1) individual trees are detected and their crowns are segmented; 2) edge trees (tree crown intersecting the plot boundary) are identified; 3) edge trees are labeled as "in" (tree apex is within the plot boundary) or "out" (tree apex is outside of the plot boundary); 4) edge trees which are "in" the plot are used to extend the plot boundary (type 2 discrepancy); and 5) the original plot and edge trees which are "out" are intersected to identify type 1 discrepancy regions. Following steps 1) to 5), ALS points are extracted for each modified plot or cell, and the remaining analyses are performed in the typical ABA fashion.

In EABA, predictions are still made for the original plot or cell areas, only predictor variables are computed using modified plots and cells. More detailed descriptions of key EABA steps are explained below. Finally, EABA performance is benchmarked against a typical ABA strategy.

B. Individual Tree Crown Delineation

A preliminary CHM was created with pixel values set either to the maximum ALS point height (dZ) within each pixel or NoData, if there were no ALS hits. Two passes of median filtering with a 3×3 neighborhood were then run on the NoData pixels only, in which a center pixel was replaced if there were at least 5 (first pass) or 3 (second pass) height values within the neighborhood. The remaining NoData pixels were set to zero. A third pass was run on pixels considered as holes. A pixel was designated as a hole if at least six of its neighbors were more than 5 m higher than the pixel itself. The value for the hole pixel was then replaced with the median of its neighbors. The pixel size of the CHM was 0.5 m.

The CHM was low-pass filtered with a height-based filtering [15], in which the scale of filtering increases relative to the height of the pixel being processed. The CHM heights were classified into eight height classes and a fixed Gaussian filter was run on the pixels of each class. The upper height limit of lowest class was set to 4 m and the Gaussian scale parameter (σ) to 0.3. The lower height limit of top-most class was set to 28.6 m and the scale parameter to 1.0. The remaining six height classes were spaced equally between the first and last classes with the σ parameters scaled linearly between 0.3 and 1.0.

Trees were located on the CHM using watershed segmentation. Watershed segments were delineated using a drainage direction-following algorithm [16], [17] on the negative image of the height-filtered CHM. To separate crown and background segments, pixels lower than two meters in height were masked from the original segments. Finally, small segments with less than four pixels were combined with a neighbor segment based on the smallest average gradient on the segment boundary between two segments.

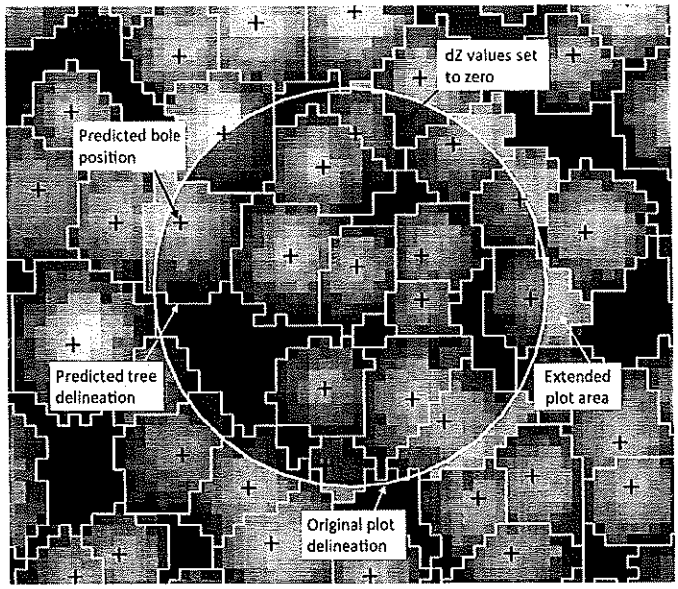

Fig. 1. Modified plot delineation and predicted trees overlaid on the canopy height model. Black crosses depict the predicted bole positions and green polygons are used to designate crown borders. Extended plot areas are highlighted in yellow, and areas with zeroed dZs are highlighted in red.

TABLE II
Mean, Standard Deviation (SD), Minimum, and Maximum for Modified Plot Areas With Respect to Plot Size

	Plot size 127 m^2		Plot size 254 m^2	
	Type 1	Type 2	Type 1	Type 2
Mean (%)	13.0	13.6	9.8	9.3
SD (%)	17.9	9.6	4.9	6.2
Min (%)	0.8	0.2	3.1	0.5
Max (%)	33.9	44.7	26.1	30.9

Type 1 and Type 2 refer to discrepancy type.

C. Modified Sample Plot

The dZs are set to zero for type 1 discrepancy regions, i.e., in areas where predicted trees are centered outside of the plot but the crowns intersect with the original plot. Fig. 1 provides an example of an actual plot that has edge discrepancies. There are eight trees with type 1 edge discrepancies. Areas with zeroed dZs are shown in red (horizontal lines in BW version). In preliminary tests we found that zeroing points for type 1 discrepancies was more effective than attempting to adjust the plot border to omit the points. In this case trees outside of the plot do not intrude into the plot and therefore ALS points hit the ground.

To correct for type 2 discrepancies, the bounds of the sample plot are extended to contain the full tree crowns. Extended plot

TABLE III
ESTIMATES OF COEFFICIENTS FOR ABA AND EABA USING 127 AND 254 m^2 SAMPLE PLOT SIZES

ABA 127 m^2	EABA 127 m^2	ABA 254 m^2	EABA 254 m^2
β_1 −3.2246 (0.0051)	β_4 −5.1479 (0.0004)	β_7 −3.7944 (0.0064)	β_{10} −4.6052 (0.0002)
β_2 0.6845 (4×10^{-16})	β_5 0.7638 ($<2\times10^{-16}$)	β_8 0.7119 ($<2\times10^{-16}$)	β_{11} 0.7252 ($<2\times10^{-16}$)
β_3 0.1222 (9×10^{-10})	β_6 0.3033 (6×10^{-16})	β_9 0.2472 ($<5\times10^{-12}$)	β_{12} 1.4497 ($<2\times10^{-16}$)

The p-values of the coefficients are given in parentheses.

areas are marked in yellow (vertical lines in BW version) in Fig. 1. By extending plot borders we improve the chance that ALS points for a plot correspond to field measured trees. Note that although the modified plot area is used when computing predictor variables, predictions are still made for the original plot or cell areas.

D. Predictor Variables

Predictor variables were computed separately with first and last echoes. First echoes contained original echo categories "first of many" and "only" and last echoes "last of many" and "only". Height (dZ) percentiles 5, 10, 20, ..., 80, 90, 95 were computed for each sample plot using ALS points with dZs above 2 m (*hx*-variables). For example, the percentile height metric *h90* for a plot is the 90th percentile height of dZs above 2 m. For each height percentile, corresponding density was computed as the proportion of echoes below the height percentile in relation to all points (*px*-variables). In addition, the mean (*havg*) and standard deviation (*hstd*) of above 2 m points, and proportion of dZs above 2 m in relation to all points were calculated (*p2m*). Similar sets of predictor variables have been used in many ABA studies. Predictor variables were computed for original (ABA) and modified sample plots (EABA) using 127 and 254 m^2 sample plots.

E. Modeling

Regression models for the plot volume (V) were constructed separately for ABA and EABA using 127 and 254 m^2 sample plot sizes. Scatter plots of V versus the ALS predictor variables indicated nonlinear relationship between V and the ALS predictor variables. After comparing models, we selected quadratic model (1) whereby the square of V appeared to provide linear relationship between plot volume and the ALS predictor variables (2)

$$V = (\beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p)^2 + \varepsilon \quad (1)$$

$$\sqrt{V} = \beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p + \varepsilon \quad (2)$$

where variance $var(\varepsilon) = \sigma^2$, V denotes plot volume, the subscript p indicates the number of predictor variables in the model, X indicates the various ALS predictor variables, and $\beta_o \ldots \beta_p$ indicate the regression coefficients. Prior to fitting the nonlinear model form, predictor variables were selected using a stepwise algorithm (*step* function in R environment; [18]) which relied upon AIC/BIC to rank potential models.

TABLE IV
ERROR STATISTICS FOR ABA AND EABA USING 127 AND 254 m^2 SAMPLE PLOT SIZES

	Plot size 127 m^2		Plot size 254 m^2	
	ABA	EABA	ABA	EABA
RMSE (%)	23.16	19.11	19.08	16.95
RSS (m^3 ha^{-1})	211 176	143 775	147 085	115 971

Residual sum of squares (*RSS*) is $\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$ and root mean square error (RMSE) is $100 \times \frac{1}{\bar{y}}\sqrt{\frac{1}{n}RSS}$, where n is the number of plots, y_i is the observed value for plot i, $\hat{y}_i$ is the predicted value for plot i, and $\bar{y}$ is the mean of y.

This particular algorithm accepts a user-supplied parameter k which can be used to increase the penalty for each additional term included in the model. The stepwise selection process was performed for the linear form of the regression model (2). Once the predictor variables were selected, model coefficients were estimated for the nonlinear model form (1) using *nls* function in R environment [18].

F. Model Assessment

Error rates were compared using RMSE values computed for training data. RMSE is a commonly used measure of performance in forest inventory and modeling-related analyses. Actual prediction errors for new data are commonly larger on average, but the difference does not affect comparisons between ABA and EABA because it can be assumed that the rates of over-fitting will be equivalent for ABA and EABA due to equal model complexities [19].

We performed a hypothesis test to determine the probability that the difference between ABA and EABA could arise from random chance. Due to the known difficulty in specifying the sampling distribution of RMSE values, we tested the hypothesis that RMSE(ABA) $\leq$ RMSE(EABA) using a bootstrap approach [20]. Bootstrap samples were taken from the original plots using sampling with replacement. The bootstrap samples were always of the same size as the original data, here 79. This was repeated 5000 times and for each of these bootstrap samples we fitted ABA and EABA models and computed RMSEs. After the bootstrap simulations, we used the empirical sampling distribution to compute the probability that the observed difference between RMSEs for ABA and EABA could arise through random chance.

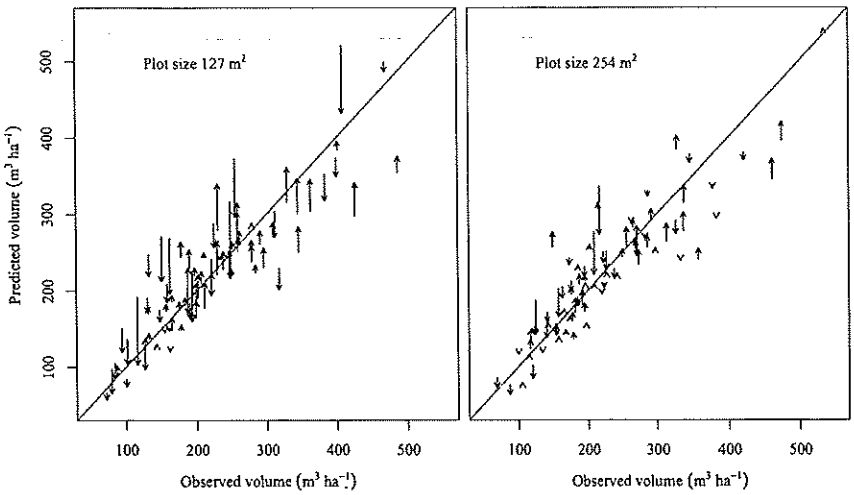

Fig. 2. Predicted volume by plot size for ABA and EABA methods. Arrowheads depict EABA predictions and tail ends depict ABA predictions. Observed volume is on the X axis, predicted volume is on the Y axis and the solid line is the 1-1 line.

IV. Results

Statistics for the areas of plots modified when performing EABA regions are presented in Table II. Statistics are presented in percent with respect to plot size. The average type 1 discrepancy area was 9.8% of the area of an original 254 m^2 precorrection sample plot and 13.0% of the area of an original 127 m^2 precorrection sample plot. The variation (standard deviation) among the areas corrected was high relative to the mean, especially with 127 m^2 sample plots. The average area of type 2 corrections was similar in area to the average area of type 1 corrections, although ranges were slightly greater.

In stepwise variable selection, a penalty term k was selected such that resulting models contained a small number of variables. Inclusion of extra variables may increase predictive performance, but also increases the risk of over-fitting. Selection of more complex models would also interfere with our primary objective, which is comparison of the ABA and EABA performances. We set the penalty term k to 9.5 which resulted in models with two predictor variables for both ABA and EABA. The models included one height and one density variable in each, which is a logical outcome because plot volume can be considered as a function of height and density. The stepwise variable selection procedure resulted in variables from first echoes only. Hence, all the predictor variables listed in (3)–(6) were computed from first echoes. The ABA model for volume using 127 m^2 plots was

$$V_{ABA,127} = (\beta_1 + \beta_2 havg + \beta_3 p2m)^2 + \varepsilon_1 \quad (3)$$

the EABA model for the volume using 127 m^2 plots was

$$V_{EABA,127} = (\beta_4 + \beta_5 havg + \beta_6 p50)^2 + \varepsilon_2 \quad (4)$$

the ABA model for volume using 254 m^2 plots was

$$V_{ABA,254} = (\beta_7 + \beta_8 h50 + \beta_9 p50)^2 + \varepsilon_3 \quad (5)$$

and the EABA model for the volume using 254 m^2 plots was

$$V_{EABA,254} = (\beta_{10} + \beta_{11} h50 + \beta_{12} p10)^2 + \varepsilon_4 \quad (6)$$

where $\beta_{1..12}$ are regression coefficients, and $\varepsilon_{1..4}$ are the residual vectors. The estimates of coefficients for all models are listed in Table III.

RSS and RMSE indicate a better fit for the EABA model than for the ABA model (Table IV). With respect to RMSE, the performance improved from 23.16% to 19.11% with 127 m^2 plots and from 19.08% to 16.95% with 254 m^2 plots, so the improvement was 17% with small and 11% with large plots. Although the variables selected in the two approaches differed, the choice of variables had minimal impact on our inference as model's RMSEs were nearly identical when we exchanged variables selected for ABA and EABA within the same plot size. The bias originating from nonlinear least squares fit was between 0.06% and 0.1%. The error rate for ABA is similar to what has been observed in other studies in Nordic countries.

In the event of a practical difference between EABA and ABA RMSE's, the statistical significance was tested with a bootstrap procedure. Following 5000 bootstraps with 127 m^2, plots RMSE (ABA) was less than RMSE (EABA) in 1.46% simulations. Corresponding figure for 254 m^2 plots was 0.62%. In other words, predictions are highly statistically significant with a "p-values" 0.0062 and 0.0146. These results indicate that the apparent improvement in RMSE for EABA is unlikely to arise due to random chance; we have a high degree of confidence that EABA provides lower error rate than ABA.

Fig. 2 shows predictions of ABA and EABA overlaid on the 1-1 line. Small sample plots are on the left side and large sample plots are on the right side. Observed volume is on the X axis and predicted volume is on the Y axis. The tip of the arrowhead depicts an EABA value and the base of the tail end depicts an ABA value. The direction and length of the arrow reflect the magnitude and direction of the differences between predictions from ABA and EABA. The changes reflect improvement (arrow toward 1-1 line) in most cases. Arrows are generally longer with the small sample plots which indicate that the magnitude of edge-tree correction was greater with small than large sample plots. This is a logical outcome since with small sample plots, larger portion of plot area was modified (see Table II). In the

cases where the ABA residuals are largest, the EABA approach appears to provide the most dramatic improvements.

V. Discussion and Conclusion

The EABA method we developed provided a lower error rate than a typical ABA approach. The greatest improvements were obtained for plots with the largest residuals. The improvement was achieved by detecting and delineating individual trees and including or excluding ALS points based on predicted crown attributes. This is conceptually the opposite of the protocol used by Mascaro *et al.* [11] in that they used predicted crown attributes to adjust the response values, while we used the predicted crown attributes to adjust predictor values. EABA follows more closely with typical field measurement protocols in that bole position exclusively defines whether a tree falls into the sample plot or not. Therefore, EABA enables prediction of new observations in the manner typically used in forest inventories.

The problem of sampling trees located near a stand edge is a well-known in forest inventory. Trees near the border of a region have different inclusion probabilities, which causes design bias if not taken into consideration [21]. In this study, we are concerned with model bias that arises due to measurement error in the predictor variables. The presence of noise in the response variable does not affect the consistency of the parameter estimates, however, the presence of noise in the predictor variables causes the fitted model coefficients to be biased toward zero ([22], p. 349). Since the ABA have type 1 and type 2 discrepancies, the EABA process reduces the source of bias.

Plot size was shown to affect edge errors and ABA performance. As plot size decreased, the proportion of edge-trees increases and consequently error due to edge-trees increased. This means that, on average, the improvements from edge correction methods should also increase for smaller plots. In this study, this effect was examined for two plot sizes, 127 and 254 m^2. The results agreed with our hypothesis: relative improvements were less for large plots (11% improvement) than for small plots (17% improvement). We achieved approximately the same error rate (RMSE) with smaller plots using EABA (19.11%) as with larger plots with typical ABA (19.08%). These results suggest that we could potentially measure 50% fewer trees with the smaller plots (50% smaller in area) using the EABA method and achieve the same performance as we would receive with ABA with the larger plots.

The proportion of edge-trees and magnitude of errors will also differ depending upon the type of forests, trees' dispersion patterns, tree size distributions, and crown shape. In savannah forests, for instance, there are typically few trees, and they have relatively wide crowns. Even for plots that are substantially larger than is typically used in boreal forests, edge-trees errors in savannah forests may have an even greater effect on error rates than in boreal forests.

The EABA method requires higher density ALS data than ABA because improvement is obtained by means of detecting individual trees. This is an important consideration because the cost of ALS data increases with density. However, some studies have indicated that the point density does not have to be very high in ITD. Kaartinen *et al.* [3], for instance, found only marginal improvements in tree delineation using point densities higher than 2 points m^{-2}. Given historical trends with respect to ALS technology, the advancement of ALS sensor technology means that point densities will most likely increase in future—increasing the number of instances where ITD methods can be used to support our EABA strategy. While EABA makes use of ITD methods, the approach does not require field measured tree positions—a significant advantage over typical ITD methods.

A research question that remains is what degree of accuracy of crown delineations is necessary to obtain improvements with EABA. One might also assume that in ABA edge-tree errors cancel out in adjacent cells, and therefore ABA would benefit more from aggregation than EABA. Although more studies are needed before we can make definitive conclusions about the relative performances of ABA and EABA methods, under the conditions encountered for this study, the proposed EABA method had a lower error rate than typical ABA.

References

[1] M. A. Wulder, C. W. Bater, N. C. Coops, T. Hilker, and J. C. White, "The role of LiDAR in sustainable forest management," *For. Chron.*, vol. 84, no. 6, pp. 807–826, 2008.

[2] J. Vauhkonen *et al.*, "Comparative testing of single-tree detection algorithms under different types of forest," *Forestry*, vol. 85, no. 1, pp. 27–40, 2011.

[3] H. Kaartinen *et al.*, "An international comparison of individual tree detection and extraction using airborne laser scanning," *Remote Sens.*, vol. 4, no. 4, pp. 950–974, 2012.

[4] T. Takahashi, K. Yamamoto, Y. Senda, and M. Tsuzuku, "Predicting individual stem volumes of sugi (Cryptomeria japonica D. Don) plantations in mountainous areas using small-footprint airborne LiDAR," *J. For. Res.*, vol. 10, no. 4, pp. 305–312, 2005.

[5] J. Vauhkonen, I. Korpela, M. Maltamo, and T. Tokola, "Imputation of single-tree attributes using airborne laser scanning-based height, intensity, and alpha shape metrics," *Remote Sens. Environ.*, vol. 114, no. 4, pp. 1263–1276, 2010.

[6] E. Næsset, "Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data," *Remote Sens. Environ.*, vol. 80, no. 1, pp. 88–99, 2002.

[7] P. Packalén and M. Maltamo, "The k-MSN method for the prediction of species-specific stand attributes using airborne laser scanning and aerial photographs," *Remote Sens. Environ.*, vol. 109, no. 3, pp. 328–341, 2007.

[8] F. Loetsch, F. Zoehrer, and K. E. Haller, *Forest Inventory*, vol. 2, München, Germany: BLV–Verlagsges, 1973, p. 469.

[9] R. Nelson *et al.*, "Technical note: Canopy height models and airborne lasers to estimate forest biomass: Two problems," *Int. J. Remote Sens.*, vol. 21, no. 11, pp. 2153–2162, 2000.

[10] E. Næsset, O.-M. Bollandsås, T. Gobakken, T. G. Gregoire, and G. Ståhl, "Model-assisted estimation of change in forest biomass over an 11 year period in a sample survey supported by airborne LiDAR: A case study with post-stratification to provide "activity data"," *Remote Sens. Environ.*, vol. 128, pp. 299–314, 2013.

[11] J. Mascaro, M. Detto, G. P. Asner, and H. C. Muller-Landau, "Evaluating uncertainty in mapping forest carbon with airborne LiDAR," *Remote Sens. Environ.*, vol. 115, no. 12, pp. 3770–3774, 2011.

[12] I. Korpela, T. Tuomola, and E. Välimäki, "Mapping forest plots: An efficient method combining photogrammetry and field triangulation," *Silva Fennica*, vol. 41, no. 3, pp. 457–469, 2007.

[13] J. Laasasenaho, "Taper curve and volume function for pine, spruce and birch," *Commun. Inst. For. Fenn.*, vol. 108, pp. 1–74, 1982.

[14] P. Axelsson, "DEM generation from laser scanner data using adaptive TIN models," in *Proc. Int. Arch. Photogramm. Remote Sens.*, Amsterdam, The Netherlands, Jul. 16–22, 2000, vol. XXXIII, part B4, pp. 110–117.

[15] J. Pitkänen, M. Maltamo, J. Hyyppä, and X. Yu, "Adaptive methods for individual tree detection on airborne laser based canopy height model," in *Proc. Int. Arch. Photogramm. Remote Sens.*, Freiburg, Germany, Oct. 3–6, 2004, vol. XXXVI, part 8/W2, pp. 187–191.

[16] J. M. Gauch, "Image segmentation and analysis via multiscale gradient watershed hierarchies," *IEEE Trans. Image Process.*, vol. 8, no. 1, pp. 69–79, Jan. 1999.
[17] P. M. Narendra and M. Goldberg, "Image segmentation with directed trees," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 2, no. 2, pp. 185–191, Mar. 1980.
[18] R Development Core Team, *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2011, ISBN 3-900051-07-0.
[19] T. Hastie, R. Tibshirani, and J. H. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. Berlin, Germany: Springer-Verlag, 2009, p. 745.
[20] B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap*. London, U.K.: Chapman & Hall, 1993, p. 456.
[21] T. G. Gregoire, "The unbiasedness of the mirage correction procedure for boundary overlap," *For. Sci.*, vol. 28, no. 3, pp. 504–508, 1982.
[22] J. Kmenta, *Elements of Econometrics*, 2nd ed. Ann Arbor, MI, USA: Univ. of Michigan Press, 1997, p. 800.

Petteri Packalen was born in Rauma, Finland, in 1973. He received the M.Sc, Lic.Sc., and D.Sc. degrees in Forestry from the University of Joensuu, Joensuu, Finland, in 2002, 2007, and 2009, respectively.

Currently, he is a University Researcher in forest mensuration with the School of Forest Sciences, Faculty of Science and Forestry, University of Eastern Finland, Joensuu, Finland. Previously, he has been an Assistant, Senior Assistant, and Professor with the Faculty of Forestry, University of Joensuu. From August 2011 to July 2012, he was a Visiting Research Scientist at the Oregon State University, Corvallis, OR, USA. He has authored over 70 peer-reviewed research articles. Recently, his focus has been on nearest neighbor imputation, combined use of ALS and spectral data in forest inventory, and the use of ALS in wildlife management. Since 2007, he has also been a Consultant for remote sensing-based forest inventory. His research interests include both practical and theoretical aspects of utilizing remote-sensing data in the monitoring and assessment of the forest environment.

Jacob L. Strunk received the B.Sc. and M.Sc. degrees in forest science from the University of Washington, Seattle, WA, USA, and the Ph.D. degree in forestry and M.Sc. degree in statistics from Oregon State University, Corvallis, OR, USA.

From 2012 to 2014, he worked part time with Aerometric Inc., Seattle, WA, USA and was a Post-Doc with Oregon State University. Since the beginning of 2014, he has been working with Washington State Department of Natural Resources, Olympia, WA, USA, co-implementing a remote-sensing augmented forest inventory and continuing his research on related topics. He also has a courtesy faculty appointment with Oregon State University.

Dr. Strunk is a member of the Society of American Foresters and the American Society of Photogrammetry and Remote Sensing.

Juho A. Pitkänen was born in Kuopio, Finland, in 1964. He received the M.S. degree in forestry from the University of Joensuu, Joensuu, Finland, in 1991.

From 1991 to 2005, he was mainly a Project Researcher and a Research Assistant with the University of Joensuu. Since 2005, he has been a Senior Researcher with the Finnish Forest Research Institute, Joensuu, Finland, which was merged into the Natural Resources Institute Finland, when it was established in 2015. His research interests include remote sensing in large area forest inventories and individual tree detection-based inventory methods on aerial image and ALS data.

Hailemariam Temesgen received the B.Sc. degree in plant sciences from Alemaya University of Agriculture, Alemaya, Ethiopia, in 1985, the M.Sc. degree in quantitative silviculture from Lakehead University, Thunder Bay, ON, Canada, in 1992, and the Ph.D. degree in forest biometrics from the University of British Columbia (UBC), Vancouver, BC, Canada, in 1999.

From 1999 to 2003, he was a Research Associate in forest biometrics and measurements with UBC and a Visiting Scientist at the Institute of Forest Management, and yield studies at the University of Gottingen, Göttingen, Germany. Since 2003, he has been Assistant, Associate, and Full Professor in forest biometrics and measurements with Oregon State University (OSU), Corvallis, OR, USA. His research interests include the development and applications of natural resource measurement and analysis techniques including developing efficient imputation, sampling and modeling techniques, and linking ground and remotely sensed data.

Prof. Temesgen was the recipient of the Emerging Scholar Award for Excellence at OSU in 2009 and Xi Sigma Pi Student Society's Mentor award in 2008. He serves as coordinator for the International Union of Research Organization's Research Group 4.01—Forest Mensuration and Modeling. He served as Associate Editor of the *Western Journal of Applied Forestry*, and currently serves as Associate Editor for *Forest Ecology and Management*.

Matti Maltamo was born in Jyväskylä, Finland, in 1965. He received the M.Sc., Lic.Sc., and D.Sc. degrees (Hons.) in forestry from the University of Joensuu, Joensuu, Finland, in 1988, 1992, and 1998, respectively.

He is a currently the Professor of Forest Mensuration Science with the Faculty of Science and Forestry, University of Eastern Finland, Joensuu, Finland. He has also worked as Visiting Professor at the Research Group of Professor Erik Naesset, Norwegian University of Life Sciences, Akershus, Norway. He was together with Naesset and Jari Vauhkonen the editor of the textbook "*Forestry Applications of Airborne Laser Scanning—Concepts and Case Studies*" (2014). He has authored more than 150 scientifically refereed papers. His research interests include forestry applications of ALS. He is an Associate Editor of the *Canadian Journal of Forest Research*.

Prof. Maltamo won together with professor J. Hyyppä the First Innovation prize of the Finnish Society of Forest Science in 2010 about "Bringing airborne laser scanning to Finland. Maltamo also obtained bronze A.K. Cajander medal of the Finnish Society of Forest Science, 2012.