Aggregating Pixel-Level Basal Area Predictions Derived from LiDAR Data to Industrial Forest Stands in North-Central Idaho

Andrew T. Hudak[1]
Jeffrey S. Evans[1]
Nicholas L. Crookston[1]
Michael J. Falkowski[2]
Brant K. Steigers[3]
Rob Taylor[3]
Halli Hemingway[4]

Abstract—Stand exams are the principal means by which timber companies monitor and manage their forested lands. Airborne LiDAR surveys sample forest stands at much finer spatial resolution and broader spatial extent than is practical on the ground. In this paper, we developed models that leverage spatially intensive and extensive LiDAR data and a stratified random sample of field plots across two mixed conifer forest landscapes in north-central Idaho. Our objective was to compare alternative models for producing unbiased maps of basal area per acre (BAA; ft2/acre), towards the greater goal of developing more accurate and efficient inventory techniques. We generated 60 topographic or stand structure metrics from LiDAR that were used as candidate predictor variables for modeling and mapping BAA at the scale of 30m pixels. Tree diameters were tallied in 1/10 and 1/5 acre fixed-radius plots (N = 165). Four models are presented, all based on 12 predictor variables. The first imputes BAA as an auxiliary variable from an imputation model that uses the machine learning algorithm randomForest in classification mode, and was developed in a prior study to map species-level basal areas of 11 conifer species; the second uses randomForest in regression mode to predict BAA as a single response variable from these same 12 predictor variables. The third is a linear regression model that predicts ln-transformed BAA using a best subset of 12 different predictor variables; the fourth again uses randomForest in regression mode, based on the same best subset of 12 variables selected for the linear regression model. We aggregated the pixel-level predictions within industrial forest stand boundaries, and then used equivalence plots to evaluate how well the aggregated predictions matched independent stand exams (having projected the tree growth in FVS and updated the stand tables to July 2003, the time of the LiDAR acquisition). All four models overpredicted BAA, but the bias was significant only in the case of the regression model. Predictions from the two randomForest models run in regression mode were very similar, despite using different predictor variables. We conclude that randomForest can be used to impute or predict canopy structure information from LiDAR-derived topographic and structural metrics with sufficient accuracy for operational management of conifer forests. In the future, tree lists could be imputed from LiDAR-derived canopy structure metrics empirically related to plot-level tree measurements. This will allow projections of tree growth at the pixel level across forested landscapes, instead of at the stand level as is the current norm.

Keywords: forest inventory; forest management; Forest Vegetation Simulator (FVS); imputation; modeling and mapping; randomForest; regression; remote sensing

In: Havis, Robert N.; Crookston, Nicholas L., comps. 2008. Third Forest Vegetation Simulator Conference; 2007 February 13–15; Fort Collins, CO. Proceedings RMRS-P-54. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station.

[1] Research Forester, Landscape Ecologist, Operations Research Analyst, respectively, USDA Forest Service, Rocky Mountain Research Station, Moscow, ID, e-mail: ahudak@fs.fed.us.

[2] PhD. Candidate, University of Idaho, Moscow, ID.

[3] GIS Supervisor, and Manager, respectively, Potlatch Forest Holdings, Inc., Lewiston, ID.

[4] GIS Administrator, Bennett Lumber Products, Inc., Princeton, ID.

Introduction

LiDAR Remote Sensing

An analysis comparing alternative remote sensing technologies demonstrated that LiDAR is more sensitive to forest canopy structure than passive optical imagery (Lefsky and others 2001). Canopy structure attributes characterized by LiDAR correlate well with stand structure attributes measured in field plots (Lefsky and others 2005a,b). This is true even in high biomass forests, where passive optical sensors become saturated. Correlations between LiDAR canopy structure metrics and stand structure metrics appear stronger in coniferous than in deciduous forests, due to the conical architecture of conifer trees allowing for greater penetration of LiDAR pulses into the canopy (Lefsky and others 2002).

Forest industries have taken note of the groundswell of promising research results regarding the utility of LiDAR data for forestry applications. LiDAR surveys are expensive but on a per acre basis can be competitive with the cost of traditional forest inventory, as labor costs have increased, especially in a market with stiff international competition. LiDAR costs are also counterweighted by the potential benefits of having highly detailed forest structure information mapped across the entire landscape surveyed. These factors have led to increased interest in operational use of LiDAR by forest industries.

Inventory Designs

The two industry partners in this project, Potlatch Forest Holdings, Inc. and Bennett Lumber Products, Inc., use similar stand-based inventory systems on their forest lands (Dennis Murphy, personal communication). The basic operating units are stands, which are delineated from aerial photographs. Trees within the delineated stand boundary are the population of interest and are sampled in randomly placed plots of variable radius (fig. 1a). The density of plots within a stand is based on a target sampling error for estimating volume. Plot design can vary between stands but generally is consistent within stands. Stand level parameters are generated from the plot level data using simple random sample estimators that vary depending on plot design. The stand-based inventory is updated using the Forest Vegetation Simulator (FVS) at six-month intervals (after the spring and fall growing seasons) by processing the original tree list at the plot level.

Scanning LiDAR systems provide spatially intensive and extensive canopy height measures that could facilitate forest inventory at the much finer scale of pixels rather than polygons (fig. 1). To estimate forest structure attributes of interest besides canopy height, the LiDAR height measures must be related to field measures of these attributes, measured in field plots randomly distributed across the full range of variation. This requires a preliminary stratification to distribute sample plot locations in an objective and representative manner within the forested landscape of interest. The sampling intensity of LiDAR could dramatically reduce the sampling intensity of inventory plots required, provided they are accurately geolocated. The reduced plot count in an inventory that uses LiDAR data (fig. 1b) argues for using fixed-radius plots for more accurate canopy structure characterization than with variable-radius plots. The field plots can then be used as training data, or reference observations, for predicting or imputing the forest structure attributes of interest to target pixels across the entire landscape. Neither the spatially explicit model inputs nor map outputs would rely on stand boundaries, which are subject to change, but could be aggregated to stand units if desired.

RandomForest

The randomForest (RF) method is so named because it uses random samples of data and variables through multiple model iterations, to generate a large group, or forest, of classification and regression trees (CART) (Breiman 2001). The classification output from RF represents the statistical mode of many decision tree classifications, hence achieving a more accurate and robust model than a CART. Randomly subsetting predictor variables allows RF to derive variable importance values and prevents problems associated with correlated variables and overfitting (Breiman 2001). The RF package in R (R Development Core Team 2004) includes two measures of variable importance (Liaw and Wiener 2002). The first measure quantifies each variable's effect on the mean squared error (MSE). Variables that markedly lower the MSE have higher importance

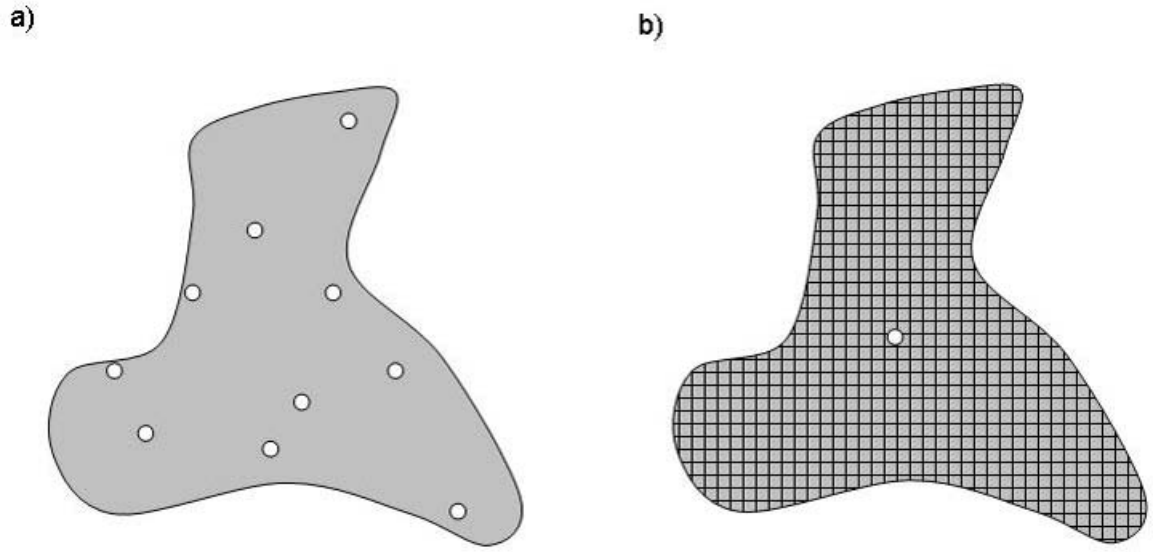

Figure 1—Conceptual diagrams illustrating a) current inventory design based on polygon units versus and b) proposed inventory design based on pixel units. The white dots represent field sample plots, which in a) have variable radius and are randomly located within the stand, but in b) have fixed radius and are randomly located within the landscape. The map unit in a) is the entire stand, while in b) each pixel is a map unit independent of the stand boundary.

values as compared to variables the have little effect on the MSE. The second measure, the Gini index, is a measure of node purity. Stronger predictor variables produce more consistent nodes across the forest of classification trees, thus having a higher Gini index. Because RF is nonparametric, the data may be rank-deficient, meaning the data may have more variables (columns) than observations (rows), have colinearities, or both. Skewed distributions in the response variables are also not a concern.

Crookston and Finley (2008) developed the "yaImpute" package in R, which includes a method based on RF classification along with several more traditional imputation algorithms methods. Imputation uses empirical relationships between attributes of interest (Y variables) measured on a sample of the observations (called reference observations) and predictor variables (X variables) available on all observations. These empirical relationships are calibrated using the reference observations. Observations that have no measured Y variables are termed target observations. A reference observation that is the nearest neighbor of a target in the multidimensional space is the source of values of Y variables that are imputed to the target. Nearness can be measured several different ways, including the most similar neighbor method introduced by Moeur and Stage (1995) and the gradiant nearest neighbor method introduced by Ohmann and Gregory (2002). The method introduced by Crookston and Finley (2008) that is based on the RF classification algorithm identifies a nearest neighbor by first concatenating the forest of classification trees across terminal nodes, and then finding the reference observation that most often shares terminal nodes with the target observation. In the case of either imputation or RF classification in yaImpute, the user defines the number of nearest neighbors to use, k, which can vary from 1 to n.

Independently of using RF in classification mode for imputation in yaImpute, the user can also run RF in regression mode to predict a single Y variable of interest (Liaw and Wiener 2002). In regression mode, the random vector takes on numerical values rather than class labels, as in classification mode (Breiman 2001). As in classification mode, out-of-bag estimation allows importance values to be assigned to the predictor variables, thus providing insight on the predictive ability of the model.

Objective

Our objective was to relate predictor variables derived from LiDAR to field data within field plots placed using a statistically rigorous sampling design to map BAA across mixed-conifer forest in north-central Idaho that is actively managed and predominantly

owned by forest industry. This objective was motivated by our access to stand exam data provided courtesy of our industry partners, to independently validate our pixel-level predictions, after aggregating them to the stand level.

This objective is an important step towards the greater goal of using FVS (Dixon 2002; Stage 1973) for imputing tree lists to spatial map units, be they cells (pixels) or stands (polygons), using LiDAR-derived predictor variables. This would enable forest managers to project growth, mortality, and the other processes already incorporated into FVS across entire landscapes.

Methods

Study Areas

The Moscow Mountain (80,789 acres) and St. Joe Woodlands (137,539 acres) study areas are situated in north-central Idaho (fig. 2). Conditions at Moscow Mountain are drier than at the St. Joe Woodlands, so forest canopies tend to have a more open structure on Moscow Mountain. Individual conifer species occur along a temperature/moisture gradient as has been described by Daubenmire (1966), beginning at the warm/dry end with *Pinus ponderosa* mostly on Moscow Mountain, to *Pseudotsuga menziesii*, *Larix*

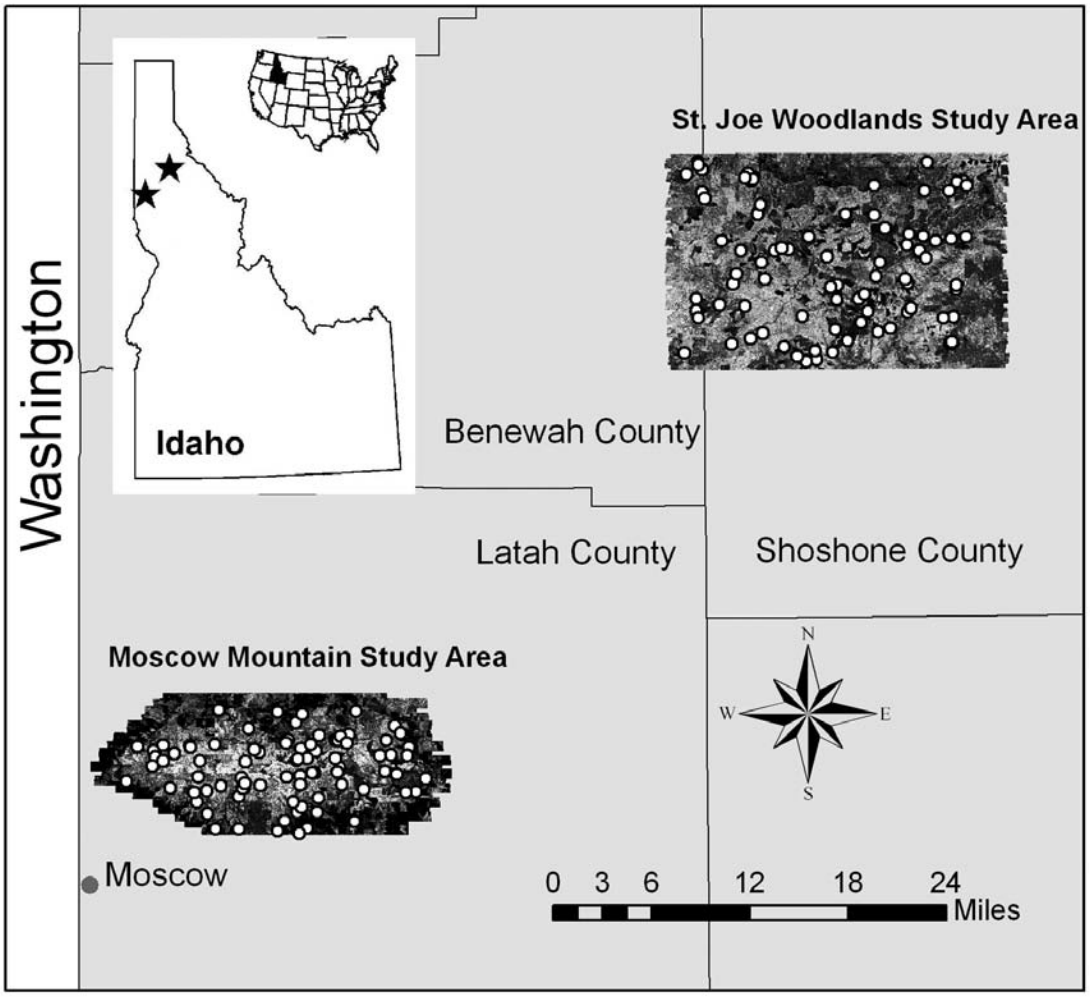

Figure 2—Study areas in north-central Idaho, with basal area per acre (BAA) predicted by Model 1 mapped in grey scale. Maps of BAA predicted by the other three models presented in this paper appear the same. The white dots represent field plot locations (N = 165).

occidentalis, *Pinus contorta*, *Abies grandis*, *Pinus monticola*, *Thuja plicata*, *Tsuga heterophylla*, *Picea engelmannii*, *Abies lasiocarpa*, and ending with *Pinus albicaulis* at the cool/wet end, or the highest elevations in the St. Joe Woodlands. More complex terrain at the St. Joe Woodlands (elevation range: 2,093–6,578 ft) than at Moscow Mountain (elevation range: 2,549–4,980 ft) produces longer and steeper gradients that drive more diverse species composition in the St. Joe Woodlands.

Field Sampling

Field sample plot locations were selected using a stratified random design. The stratification variables were elevation from a 30 m USGS digital elevation model (DEM), solar insolation (Fu and Rich 2000), and a mid-infrared corrected normalized difference vegetation index (NDVIc) (Nemani and others 1993) calculated from an Landsat ETM+ scene (18 August 2002). The NDVIc has been found to be superior to NDVI for estimating leaf area index in mixed-conifer forests of northern Idaho (Pocewicz and others 2004). Plots were geolocated with a Trimble Pro-XR Global Positioning System (GPS). A minimum of 150 points were recorded on the ground surface at plot center, and later differentially corrected and averaged for a final three-dimensional (3D) point position with ±2.6 ft horizontal and ±3.6 ft vertical accuracy (Trimble Pathfinder Office). Plots were 1/10 acre at Moscow Mountain and 1/5 acre at the St. Joe Woodlands and of fixed radius, with all trees ≥5 inches diameter at breast height (dbh) tallied. The sampling design failed to capture rare, late successional conditions, so two plots were randomly located within two old-growth stands (one in each study area) to capture the high end of the BAA gradient, for a total of 165 plots.

Measured tree diameters (N = 5240) were converted to tree basal areas, summed, and divided by the plot area to estimate BAA for modeling. Eleven plots at Moscow Mountain lacked trees ≥5 inches dbh but were assigned negligible values of 0.4356 ft^2/acre (0.1 m^2/ha), to enable their inclusion in the analysis when a natural logarithm (ln) transform was applied.

LiDAR Sampling

Horizons, Inc. (Rapid City, SD) flew the Light Distance And Range (LiDAR) survey at an altitude of 8,000 ft above mean terrain during the summer of 2003, using an ALS40 system operating at 1,064 nm and a pulse rate of 20 KHz. Data were delivered in the form of unclassified point data. Evans and Hudak (2007) developed a Multiscale Curvature Classification algorithm in ArcInfo Macro Language (AML) to classify the returns as either ground or non-ground. The classified ground returns were interpolated into a 2-m DEM, from which several topographic predictor variables were derived (table 1).

Subtracting the 2-m DEM from the unclassified LiDAR returns produced a canopy height layer normalized for topography. By definition, returns classified as ground returns equaled 0 m in height. Returns greater than 0 m in height were considered non-ground returns. Returns greater than 1 m in height were considered vegetation returns. Distributional statistics (min, max, mean, percentiles, variance, skewness, kurtosis, and so on) were calculated from the height and intensity values of the vegetation returns. Vegetation density was calculated as the percentage of total returns that were vegetation returns. The percentage of vegetation returns occurring within each of six defined canopy height strata was also calculated. All of these metrics were derived from the LiDAR data in 30-m bins. The Universal Transverse Mercator (UTM, Zone 11) Easting and Northing coordinates were added to produce 60 candidate variables for modeling (table 1).

Modeling

This analysis compares BAA predictions from four alternative models, as described below.

Model 1—For forestry applications, Y variables are typically measured in plots (for example, BAA in this study), while X variables are environmental variables typically measured by remote sensing (for example, LiDAR in this study). The field sampled plots have both X and Y variables and all map units have X variables. Hudak and others (2008) pruned the list of 60 candidate predictor variables (table 1) down to 12 for pre-

Table 1—Candidate predictor variables and those selected for the four models compared.

Variable	Description	Models 1 and 2	Models 3 and 4
EAST	UTM Easting (meters)		
NORTH	UTM Northing (meters)		
ELEV	Elevation (meters)	X	X
SLP	Slope (degrees)		
TSRAI	Topographic solar radiation aspect index (Roberts and Cooper 1989)	X	X
SCOSA	Percent slope*cos(aspect) transformation (Stage 1976)		
SSINA	Percent slope*sin(aspect) transformation (Stage 1976)		
INSOL	Solar insolation (HEMI 2000)		
CRR	Canopy relief ratio (Pike and Wilson 1971)		
HMIN	Heights minimum		
HMAX	Heights maximum		
HRANGE	Heights range	X	
HMEAN	Heights mean		
HAAD	Heights average absolute deviation	X	
HMAD	Heights median absolute deviation		
HSTD	Heights standard deviation		
HVAR	Heights variance		
HSKEW	Heights skewness		
HKURT	Heights kurtosis		
HCV	Heights coefficient of variation	X	X
H05PCT	Heights 5th percentile	X	
H10PCT	Heights 10th percentile		
H25PCT	Heights 25th percentile	X	
H50PCT	Heights 50th percentile (median)		
H75PCT	Heights 75th percentile		
H90PCT	Heights 90th percentile		
H95PCT	Heights 95th percentile		
HIQR	Heights interquartile range		
IMIN	Intensity minimum		
IMAX	Intensity maximum		
IRANGE	Intensity range		
IMEAN	Intensity mean		X
IAAD	Intensity average absolute deviation		
IMAD	Intensity median absolute deviation		
ISTD	Intensity standard deviation		
IVAR	Intensity variance	X	
ISKEW	Intensity skewness	X	
IKURT	Intensity kurtosis		X
ICV	Intensity coefficient of variation		
I05PCT	Intensity 5th percentile		
I10PCT	Intensity 10th percentile		
I25PCT	Intensity 25th percentile		
I50PCT	Intensity 50th percentile (median)		
I75PCT	Intensity 75th percentile		
I90PCT	Intensity 90th percentile		
I95PCT	Intensity 95th percentile		
IIQR	Intensity interquartile range		
DENSITY	Canopy density (vegetation returns/total returns * 100)	X	X
STRATUM0	Percentage of ground returns = 0 m		
STRATUM1	Percentage of non-ground returns > 0 m and <= 1 m in height	X	
STRATUM2	Percentage of vegetation returns > 1 m and <= 2.5 m in height		
STRATUM3	Percentage of vegetation returns > 2.5 m and <= 10 m in height		X
STRATUM4	Percentage of vegetation returns > 10 m and <= 20 m in height		X
STRATUM5	Percentage of vegetation returns > 20 m and <= 30 m in height		X
STRATUM6	Percentage of vegetation returns > 30 m in height		X
TEXTURE	Standard deviation of non-ground returns > 0 m and <= 1 m		
PCT1	Percentage 1st returns		
PCT2	Percentage 2nd returns	X	X
PCT3	Percentage 3rd returns		X
NOTFIRST	Percentage 2nd or 3rd returns		

dicting species-level basal areas and tree densities of 11 conifer species (in other words, 22 Y variables). In that analysis, the RF classification method produced more accurate results than seven more traditional imputation methods available in yaImpute. The RF classification model used to map species-level basal areas and tree densities was applied to map BAA as an auxiliary variable, using $k = 1$ nearest neighbor, and constitutes the first map considered in this analysis.

Model 2—The RF algorithm can also be employed as a regression tool, for predicting a single Y variable of interest. For this analysis, that variable of interest was BAA, across all tree species. The same 12 predictor variables used in Model 1 to impute a BAA map using RF in classification mode (Hudak and others, 2008) were also used to map BAA using RF in regression mode.

Model 3—A multiple linear regression model was developed for comparison because it is the predictive modeling technique most broadly used for relating field and remotely sensed data, and has been successfully applied for mapping BAA in this landscape (Hudak and others 2006). Regression is much more vulnerable to colinearity problems than nonparametric methods such as RF, so the maximum Pearson correlation allowed between predictor variables was 0.8. (The maximum Pearson correlation between predictor variables included in Model 1 was 0.9; Hudak and others, 2008.) Although RF is resistant to problems of colinearity and overfitting in either classification or regression mode, it is neither helpful nor instructive to include highly correlated predictor variables in the same model. The best subset of twelve predictor variables that satisfied this constraint was selected for predicting BAA, after ln transformation to correct the positive skew. This necessitated a bias correction (Baskerville 1976) to correct for the bias introduced by back-transforming predicted ln (BAA) to the natural scale, following Hudak and others (2006).

Model 4—The twelve predictor variables selected as the best subset for multiple linear regression were used in another RF model run in regression mode.

In summary, the output from four alternative models for predicting BAA were compared: (1) RF in classification mode based on 12 predictor variables used in an imputation model (yaImpute) from a prior analysis; (2) RF in regression mode based on the same 12 variables as in Model 1; (3) multiple linear regression based on a best subset of 12 new predictor variables; and (4) RF in regression mode based on these same 12 variables as in Model 3.

Mapping

Raster layers of the predictor variables selected by the models were generated for both study areas at a 30 m resolution using the fishnet command in ArcInfo. The intersect command was used to assign the corresponding cell ID to each LiDAR point. The LiDAR points then were exported from ArcInfo as a comma-delimited (csv) file containing six attributes: bin-ID, bin centroid X coordinate, bin centroid Y coordinate, height (Z coordinate), intensity, and return level. The csv file of LiDAR points was sorted on the Bin-Id using the DOS SORT command, then input into a Perl program developed to iteratively subset the LiDAR point data by bin-id and calculated the LiDAR metrics within each bin. Metrics were calculated within each bin and written to an output csv file. A batch file was written in R that looped through each output csv file, creating ArcInfo ASCII grids of the metrics selected as predictor variables in R.

The yaImpute package (Crookston and Finley 2008) also includes functions to assign values of the response variable(s) to target cells across the landscape, whether by imputation, regression or some other predictive model, wherever data for the predictor variables exist. A 30-m mapping resolution was used for this analysis.

Validation

Predictions of BAA at the 30 m pixel level were aggregated within stand boundaries delineated by industry partners Potlatch Forest Holdings, Inc., and Bennett Lumber Products, Inc., who also provided stand exam data for 1,024 and 177 stands, respectively. Tree growth in the stand exam data was projected forward from the time of inventory until July 2003, when the LiDAR survey was conducted. Thus, only the spring growing season was included in the 2003 projection. The updated stand projections were then

used to validate our predictions aggregated to the stand level using the ZONALMEAN function in ArcMap.

The regression-based method of equivalence tests (Robinson and Froese 2004; Robinson and others 2005) was used to validate predictions extracted from the four maps of BAA. Traditionally, models are validated under the null hypothesis of no difference between predictions and observations, or that the model is acceptable. However this approach is more likely to validate a model with low power (Robinson and Froese 2004). Equivalence tests begin with the null hypothesis that the model is unacceptable, thus shifting the burden of proof on to the model to demonstrate validity (Robinson and Froese 2004; Robinson and others 2005). The equivalence package in R regresses observations on to predictions, and uses bootstrapping to not only test between the similarity of means, but the similarity of individual predictions and observations, thus increasing statistical power for more robust model validation (Robinson and others 2005).

Results

The LiDAR predictor variables selected by Hudak and others (2008) for Model 1 included several height distributional metrics (for example, range, average absolute deviation, 5th and 25th percentiles) (table 1), for which Model 2 assigned importance values (fig. 3a). The best subset of LiDAR predictors selected for Model 3 included several upper canopy density metrics (canopy density in 3rd, 4th, 5th, and 6th strata) (table 1), for which Model 4 assigned importance values (figure 3b). The influential topographic predictors of elevation and topographic solar radiation aspect index were selected in both cases, as was density, coefficient of variation in height, and percentage of second returns (table 1; fig. 3). Two predictor variables derived from the intensity values were included in the subsets of both the imputation predictors (variance and skewness) and regression predictors (mean and kurtosis) (table 1; fig. 3).

All four models tended towards overprediction of BAA relative to the independent stand exams, although prediction residuals indicate BAA could be overpredicted greatly in some stands and underpredicted greatly in some others (fig. 4). The mean BAA from the imputation model (Model 1) was the least biased, while the mean predicted by multiple linear regression (Model 3) was most biased. The interquartile range of values predicted by the imputation and regression models was unrealistically larger than the interquartile range predicted by the two RF models in regression mode (Models 2 and 4), which closely matches the interquartile range of the stand exams (fig. 4).

Equivalence tests were used to validate the pixel-level predictions aggregated to the stand level with the stand-based inventories, updated with FVS to the July 2003 time of LiDAR acquisition. Each equivalence test regresses observations (stand exams) on predictions (aggregated pixels), then bootstraps the data to test the significance of both the intercept and slope terms of these simple linear regression models. Results are indicated graphically (fig. 5). The intercept test (gray error bars plotted around the mean value, but largely hidden by the black error bars) in each of the equivalence regressions does not significantly differ from its expected range (gray shaded region) for any of the models, as would be the case if the gray error bars were located outside the shaded region (fig. 5). The slope test (black error bars plotted around the solid diagonal line) in each of the equivalence regressions does significantly differ from its expected range (dotted diagonal lines) in the case of the imputation and regression models (Models 1 and 3), but not the two RF models run in regression mode (Models 2 and 4).

Discussion

Breiman (2001) found that RF did not incorporate randomness into the model quite as effectively in regression mode as in classification mode. One could associate non-randomness with bias, which might help explain why BAA imputed as an auxiliary variable by RF in classification mode (Model 1) was the least biased of the four models presented in this study (figs. 4 and 5). However, BAA predictions from the two RF models run in regression mode (Models 2 and 4) were not significantly biased either. In fact, the bias was only significant in the case of the multiple linear regression model (Model 3) (fig. 5). Although not presented in this paper, Hudak and others (2006) also overpredicted BAA using a multiple linear regression model, to an even larger degree than Model 3 in this study. The consistently positive and significant bias of these multiple linear regression

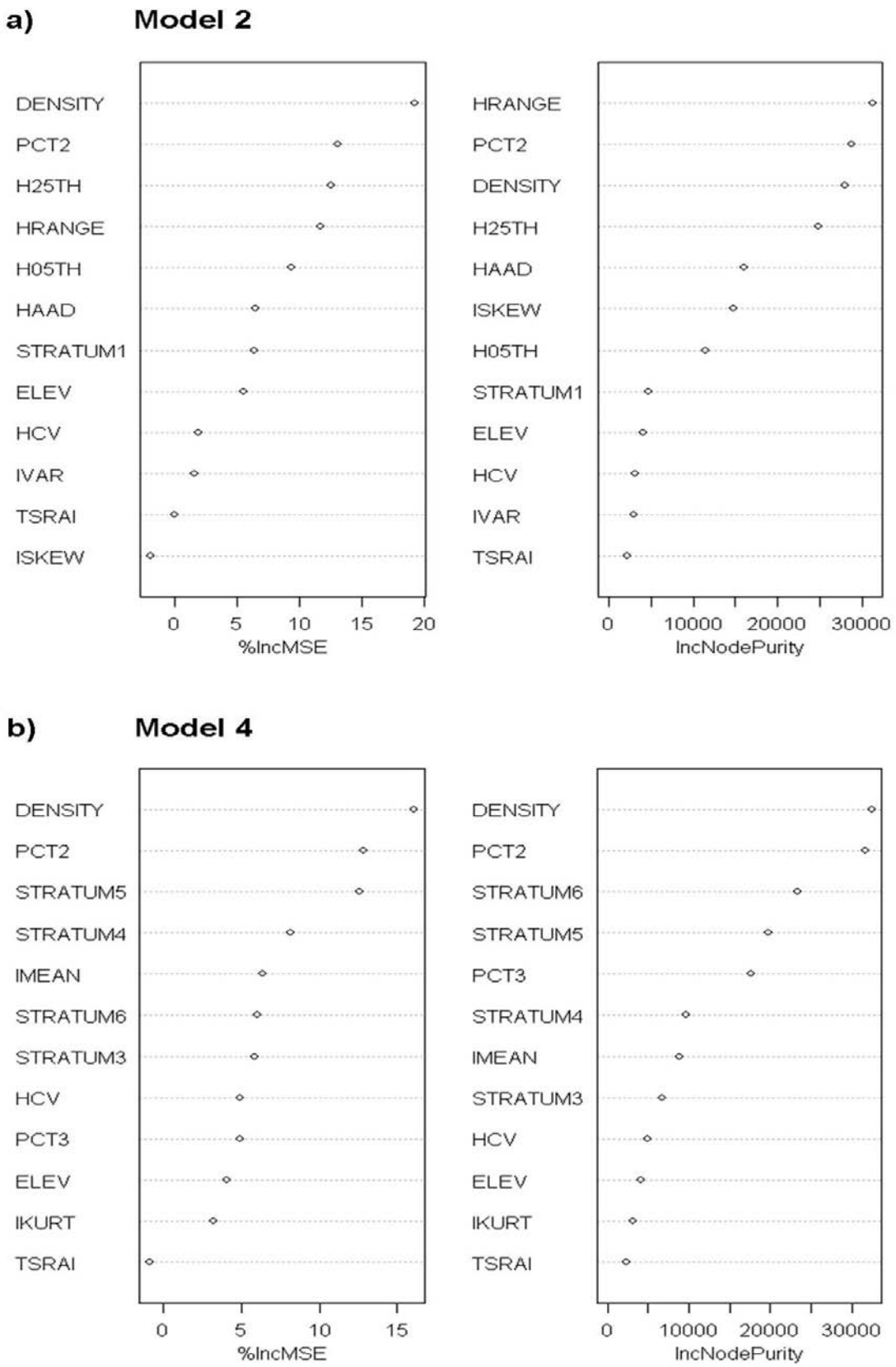

Figure 3—Importance plots from the randomForest models run in regression mode (Models 2 and 4) using a) 12 predictor variables selected for imputation (Model 1), and b) 12 predictor variables selected for multiple linear regression (Model 3). Two measures of relative importance are indicated in each case: influence on the mean squared error, and influence on node purity. Variables are plotted from the top with importance values in descending order.

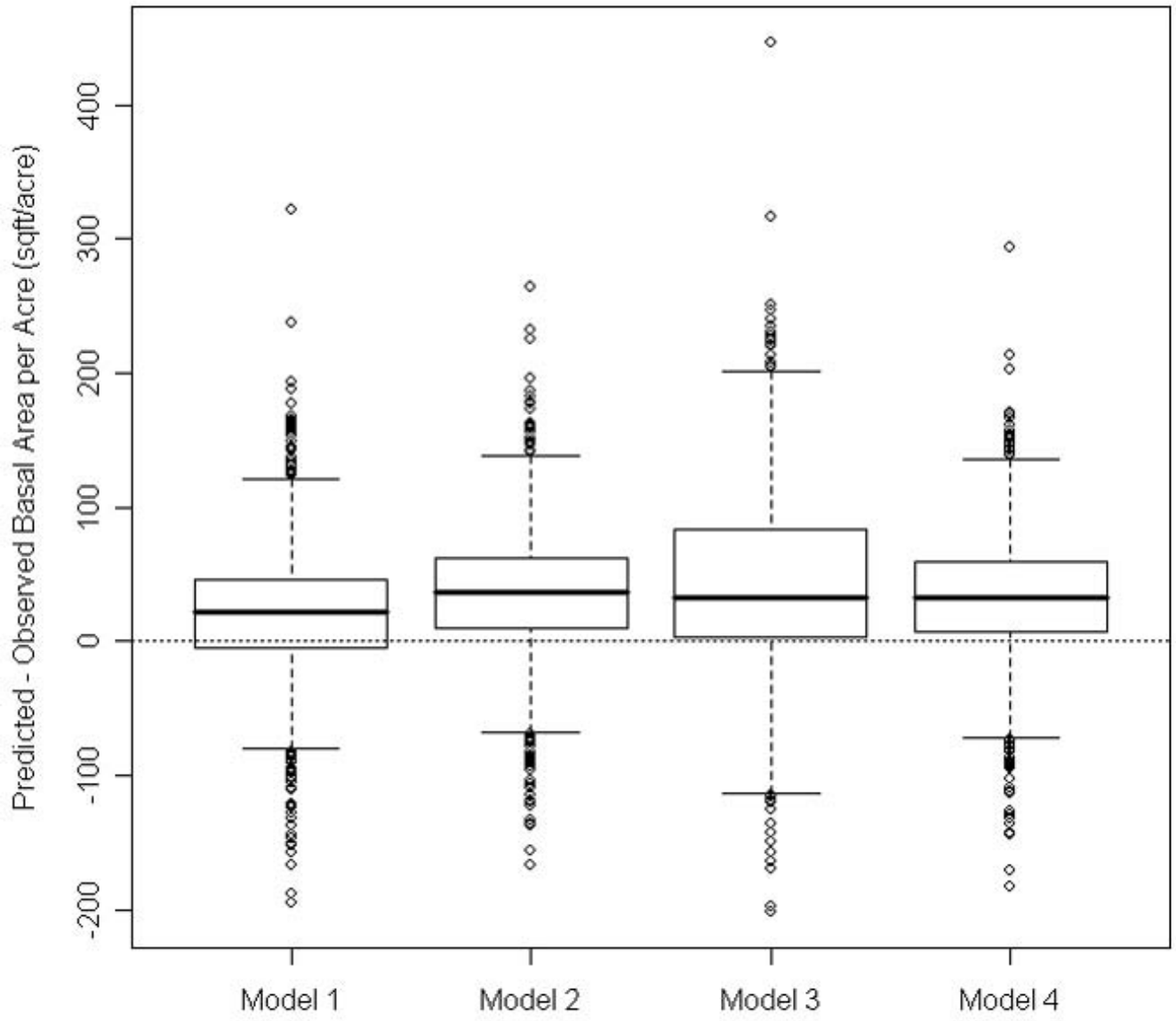

Figure 4—Boxplots comparing paired residuals, quantifying the difference between predicted pixel-level BAA aggregated within 1201 industry stands and observed stand-level BAA, for the four predictive models. Thick horizontal lines mark the medians, box ends represent lower and upper quartiles, line ends indicate the 5th and 95th percentiles, and dots show stands beyond the 5th and 95th percentiles. The dotted horizontal line indicates where model bias equals zero.

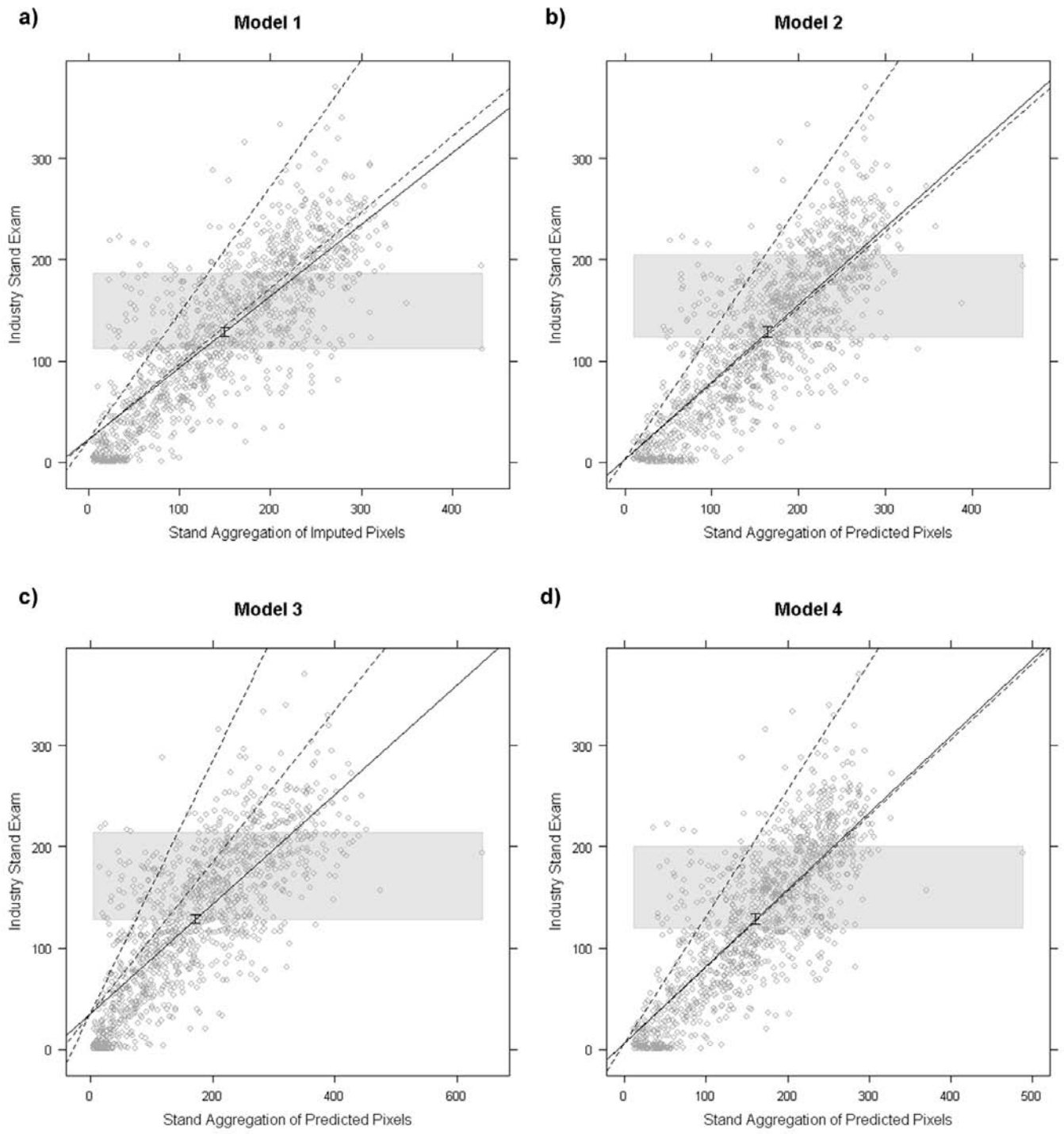

Figure 5—Equivalence plots that graphically indicate whether pixel-level BAA predictions aggregated to the stand level significantly differ from stand exams, based on the four models as labeled. The equivalence test regresses observations (stand exams) on predictions (aggregated pixels) in a simple linear regression, while bootstrapping the data. If the gray error bar (largely hidden by the black error bar) falls within the shaded gray region, then the intercept of the linear model does not significantly differ from its range of expected values. If the black error bar falls between the dotted lines, then the slope term of the linear model does not significantly differ from its range of expected values

models may be an artifact of the ln transformation and subsequent back-transformations, which warrants consideration of other methods for adjusting retransformation bias (e.g., Duan 2002). More likely, the stand-based inventories themselves are not an unbiased representation of the variation in BAA that formed the basis of our sampling design across our two study landscapes. BAA may be higher on the non-industrial forest lands within our study areas, which would be represented in the sample plots but not the stand exam data. We would need to limit the plots used to develop our models to only those occurring in industrial forest stands for which we have inventory data, to determine to what degree the models may be biased.

The three RF models also more satisfactorily reproduced the range of variability in the stand exams than the regression model (figs. 4 and 5). The random element of RF causes output to vary slightly between separate model runs, while multiple linear regression output is invariant. However, several runs of the RF models, each consisting of 500 classification trees, were found to produce very consistent results, so these differences were too negligible to alter our results to a degree that would change our interpretation or conclusions.

Hudak and others (2006) also considered the ten Advanced Land Imager (ALI) reflectance bands as candidate variables to predict basal area and tree density using multiple linear regression, but these spectral variables contributed little to the model. Similarly, Hudak and others (2008) found that these spectral variables, along with three simple vegetation indices, contributed only negligibly to imputation of basal area and tree density of 11 individual conifer species. Therefore, we did not pursue using spectral imagery in this analysis. Multispectral or hyperspectral imagery could aid discrimination between coniferous and deciduous species, or habitat types with variable phenologies. However, in the mixed conifer forest within our two study areas, LiDAR data alone appears to be sufficient for modeling structural attributes at the species level (Hudak and others, 2008). This is important because species can change timber value by a factor of 4 or 5 (Dennis Murphy, personal communication). Elevation and aspect play an obviously important role in determining vegetation structure and composition in these topographically complex landscapes, and can be mapped with unprecedented detail with LiDAR surveys. This paper further demonstrates that useful canopy metrics can be obtained from LiDAR intensity and density metrics, not just height metrics. In particular, vegetation density, or the proportion of total returns that are vegetation returns, is consistently a powerful predictor variable (fig. 3).

It is interesting that such similar predictions were obtained from the two RF models run in regression mode (Models 2 and 4; figs. 4 and 5). The predictor variables selected for imputation consisted of several height distribution metrics, some from the lower canopy, while the predictor variables selected for multiple linear regression were mostly density metrics from the upper canopy. We conclude that the canopy structure variation in these coniferous forests can be characterized equally well by different variable combinations. The 60 candidate variables considered for our models are likely many more than are necessary to develop a satisfactory model. Future research will test this suite of candidate variables in other forest types, to evaluate empirically which structural metrics have the greatest general utility.

LiDAR may prove essential in future forest inventory design. LiDAR data can provide the detailed height data that correlate well with tree diameter, basal area, and volume. Significantly fewer field plots may be required to build the empirical relationships necessary for predicting these and other attributes of interest to forest managers. Imputation of diameter distributions (i.e., tree lists) from independent LiDAR height, density, and intensity distributional metrics could provide spatially gridded inputs into FVS. This could change inventory designs from being based on stand (polygon) units to being based on cell (pixel) units (fig. 1). LiDAR-derived canopy structure layers could provide a more objective data source than aerial photos for delineating stands, for the purpose of aggregating gridded map outputs (e.g., BAA) to stand units for managers. Further research is needed to quantify the degree to which this may impact overall accuracy, efficiency, and cost of managing forested landscapes.

Conclusion

We found that applying randomForest in either classification or regression mode can consistently predict BAA from LiDAR-derived predictor variables. Mean BAA of pixels aggregated to the stand level did not significantly differ from independent stand based inventories. We recommend that forest industry invest in LiDAR and associated field plot surveys for improved forest management. These results move us another step closer to our goal of predicting tree lists from LiDAR structure metrics, so that FVS projections might be generated within map units across forested landscapes, for more precise forest inventory and management.

Acknowledgments

This research was funded through the Sustainable Forestry component of Agenda 2020, a joint effort of USDA Forest Service Research & Development and the American Forest and Paper Association. We thank David Hall for his programming assistance. We thank technical reviewers Dennis Ferguson and Dennis Murphy, and statistician Rudy King, for their helpful comments.

References

Breiman, L. 2001. Random forests. Machine Learning. 45:5–32.

Crookston, N.L.; Finley, A. 2008. yaImpute: an R package for k-NN imputation. Journal of Statistical Software. 28(10). Online: http://www.fs.fed.us/rm/pubs_other/rmrs_2008_crookston_n001.html. Package URL: http://cran.r-project.org/src/contrib/Descriptions/yaImpute.html.

Daubenmire, R. 1966. Vegetation: identification of typal communities. Science. 151:291–298.

Dixon, G.E., comp. 2002. Essential FVS: A user's guide to the Forest Vegetation Simulator. Internal Rep. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Forest Management Service Center. 193 p.

Duan, N. 2002. Smearing estimate: a nonparametric retransformation method. Journal of the American Statistical Association. 78:605–610.

Evans, J.S.; Hudak, A.T. 2007. A multiscale curvature algorithm for classifying discrete return lidar in forested environments. IEEE Transactions on Geoscience and Remote Sensing. 45:1029–1038.

Fu, P.; Rich, P.M. 2000. The Solar Analyst 1.0 User Manual. Lawrence, KS: Helios Environmental Modeling Institute (HEMI), LLC.

Hudak, A.T.; Crookston, N.L.; Evans, J.S.; Falkowski, M.J.; Smith, A.M.S.; Gessler, P.; Morgan, P. 2006. Regression modeling and mapping of coniferous forest basal area and tree density from discrete-return lidar and multispectral satellite data. Canadian Journal of Remote Sensing. 32:126–138.

Hudak, A.T.; Crookston, N.L.; Evans, J.S.; Hall D.E.; Falkowski, M.J. 2008. Nearest neighbor imputation of species-level, plot-scale forest structure attributes from LiDAR data. Remote Sensing of Environment. 112(5): 2232–2245.

Lefsky, M.A.; Cohen, W.B.; Spies, T.A. 2001. An evaluation of alternate remote sensing products for forest inventory, monitoring, and mapping of Douglas-fir forests in western Oregon. Canadian Journal of Forest Research. 31:78–87.

Lefsky, M.A.; Cohen, W.B.; Parker, G.G.; Harding, D.J. 2002. Lidar remote sensing for ecosystem studies. BioScience. 52:19–30.

Lefsky, M.A.; Hudak, A.T.; Cohen, W.B.; Acker, S.A. 2005a. Geographic variability in lidar predictions of forest stand structure in the Pacific Northwest. Remote Sensing of Environment. 95:532–548.

Lefsky, M.A.; Hudak, A.T.; Cohen, W.B.; Acker, S.A. 2005b. Patterns of covariance between forest stand and canopy structure in the Pacific Northwest. Remote Sensing of Environment. 95:517–531.

Liaw, A.; Wiener, M. 2002. Classification and regression by randomForest. R News. 2:18–22.

Moeur, M.; Stage, A.R. 1995. Most similar neighbor: an improved sampling inference procedure for natural resource planning. Forest Science. 41:337–359.

Ohmann, J.L.; Gregory, M.J. 2002. Predictive mapping of forest composition and structure with direct gradient analysis and nearest neighbor imputation in coastal Oregon, U.S.A. Canadian Journal of Forest Research. 32:725–741.

Pike, R.J.; Wilson, S.E. 1971. Elevation relief ratio, hypsometric integral, and geomorphic area altitude analysis. Geological Society of America Bulletin. 82:1079–1084.

R Development Core Team. 2004. R: A language and environment for statistical computing. ISBN 3-900051-07-0. Vienna, Austria: R Foundation for Statistical Computing. Online: http://www.R-project.org.

Roberts, D.W.; Cooper, S.V. 1989. Concepts and techniques of vegetation mapping. In: Land classifications based on vegetation: Applications for resource management. Gen. Tech. Rep. GTR-INT-257. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Research Station: 90–96.

Robinson, A.P.; Froese, R.E. 2004. Model validation using equivalence tests. Ecological Modelling. 176:349–358.

Robinson, A.P.; Duursma, R.A.; Marshall, J.D. 2005. A regression-based equivalence test for model validation: shifting the burden of proof. Tree Physiology. 25:903–913.

Stage, A.R. 1973. Prognosis model for stand development. Res. Pap. INT-137. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station.

Stage, A.R. 1976. An expression of the effects of aspect, slope, and habitat type on tree growth. Forest Science. 22:457–460.