Computer Automation of a LiDAR Double-Sample Forest Inventory

by
Robert C. Parker

Forest and Wildlife Research Center
Mississippi State University

The Forest and Wildlife Research Center at Mississippi State University was established by the Mississippi Legislature with the passage of the renewable Natural Resources Research Act of 1994. The mission of the Center is to conduct research and technical assistance programs relevant to the efficient management and utilization of the forest, wildlife, and fisheries of the state and region, and the protection and enhancement of the natural environment associated with these resources. FWRC scientists conduct research in laboratories and forests administered by the University and cooperating agencies and industries throughout the country. Research results are made available to potential users through the University's educational program and through Center publications such as this, which are directed as appropriate to forest landowners and managers, manufacturers and users of forest products, leaders of government and industry, the scientific community, and the general public. Dr. George M. Hopper is director of the Forest and Wildlife Research Center

Author

Robert C. Parker is an associate professor in the Department of Forestry. His primary research interests are inventorying and monitoring of forest ecosystem components.

To Order Copies

Copies of this and other Forest and Wildlife Research Center publications are available from:

Publications Office
Forest and Wildlife Research Center
Box 9680
Mississippi State, MS 39762-9680

Please indicate author(s), title, and publication number if known.
Publications are also available at our website at www.cfr.msstate.edu.

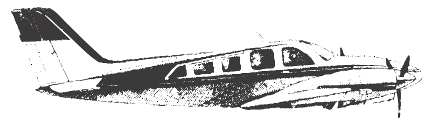
Citation

Parker, R.C. 2006. Computer automation of a LiDAR double-sample forest inventory. Forest and Wildlife Research Center, Bulletin FO275, Mississippi State University. 19 pp.

Research Bulletin FO275

FOREST AND WILDLIFE RESEARCH CENTER

Mississippi State University



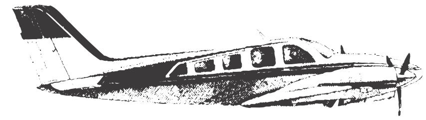
Computer Automation of a LiDAR Double-Sample Forest Inventory

by

Robert C. Parker

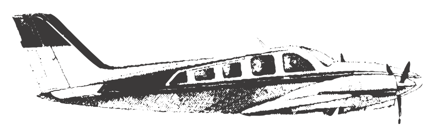
Forest and Wildlife Research Center
Mississippi State University

Abstract



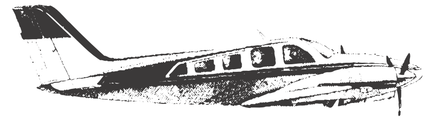
Mounted in aircraft, LiDAR (Light Detection And Ranging) technology uses pulses of light to collect data about the terrain below. LiDAR is capable of locating tree crowns with an accuracy of more than 90 percent for common spacings in pine plantations. A new system, the LiDAR Double-Sample automation system (LIDARDS), was developed to automate inventory and statistical computations for LiDAR tree height data and ground data in a stratified, double-sample inventory. LIDARDS, a Windows-based menu system, provides a set of data formats and computational procedures that facilitate the rapid computation of a stratified, LiDAR-based, double-sample forest inventory. Sample tree diameters and heights from ground plots are used to obtain prediction equations for height and dbh of target trees identified on LiDAR surfaces. LiDAR heights in the Phase 1 data are allocated in a Monte Carlo simulation to species-product classes on each matching Phase 2 ground plot on the basis of percent distribution by numbers of trees. Phase 1 LiDAR heights are randomly allocated to encountered species classes in each stratum and used to compute numbers of trees, basal area, and volume per acre. Phase 2 tree measures of dbh and height are used to compute LiDAR estimates of basal area (ft^2) and volume (ft^3) by using field derived dbh-height equations to predict dbh and volume. Comma delimited text files of Phase 1 and 2 estimates of trees, basal area, and volume on a per acre basis, double-sample regression estimates and associated precision and fit statistics, and partitioned volumes for each user-defined stratum are written to disk for subsequent use in spreadsheet or word processor software.

Introduction



Light detection and ranging (LiDAR) is a relatively new remote sensing tool that has the potential for use in the acquisition of measurement data for inventories of standing timber. LiDAR systems have been used in a variety of forestry applications for the quantification of biomass, basal area, and tree and stand height estimates (Nelson et al. 1988, Nilsson 1996, Magnussen and Boudewyn 1998, Lefsky et al. 1999, Means et al. 2000). Researchers in the Forest and Wildlife Research Center used small-footprint, multi-return LiDAR (0.25 shots per m^2) in a double-sample application with a ground-based forest inventory in central Idaho and achieved an unstratified sampling error of 11.5% on mean volume per acre at a 95% level of confidence (Parker and Evans 2004). Sampling error was defined as one-half the confidence interval on mean volume expressed as a percentage of mean volume. Scientists have also used small-footprint, multi-return LiDAR (4 shots per m^2, with a footprint size of 0.122 m and 1 shot per m^2 with a footprint size of 0.213 m) in southeast Louisiana to achieve stratified sampling errors of 9.5% and 7.6% (95% level of confidence) on mean volume per acre with the high- and low-density LiDAR, respectively (Parker and Glass 2004). The standard and sampling errors were not improved when the high- and low-density LiDAR surfaces were smoothed or when LiDAR heights were adjusted to ground values with a regression equation (Parker and Mitchel 2005). The double-sample models used for LiDAR-based inventories were adapted from the ground-based point sampling model by Avery and Burkhart (2002). The objective of this work was to develop a user-friendly, computer application of a double-sample forest inventory that allows the user to simultaneously analyze data from two LiDAR data sets and ground data for multiple species stands.

Inputs and Models



LiDAR and field inventory data

LiDAR data sets were surfaced to produce first-return canopy and last-return digital elevation models (DEM) with 0.2 m cell sizes using a linear interpolation technique. Tree locations and heights were determine with procedures developed by McCombs et al. (2003) that utilized a variable search window to identify tree peaks as points that were higher than 85% of the surrounding maxima. A spatial filtering technique derived from image analysis called smoothing was used to reduce tree location errors by minimizing the abrupt elevation changes in the initial canopy surface that could be erroneously interpreted as tree locations. A 1 m^2 filter moved across the LiDAR canopy surface, pixel-by-pixel, averaged the z-values within the window, and placed the result in the center pixel. Tree height was interpreted as the difference between canopy and ground DEM z-values at each identified tree peak location.

Inventory design for this double-sample application involved the use of a systematic grid of circular plots 0.05 ac in size on a 52.6 ft by 330 ft grid with every 10^{th} plot as a Phase 2 ground plot and all plots being Phase 1 LiDAR plots. Ground data on each Phase 2 plot included tree diameter at breast height (dbh) on all trees > 4.5 in. and total height, azimuth, and distance on 2 sample trees.

Phase 2 sample tree regression models

Sample tree diameters and heights from ground plots are used to obtain prediction equations for dbh and ground height of target trees identified on the LiDAR surfaces for each of the encountered species groups. The dbh-height models used are:

$$(1)\quad dbh = b_o + b_1 [Ln(H_{gr})]^{b2}$$
$$(2)\quad H_{gr} = b_o + b_1 (dbh)^{b2}$$
$$(3)\quad H_{gr} = b_o + b_1 (H_{Li})^{b2}$$

where: H_{gr} is measured ground height of trees identified on LiDAR plots,
H_{Li} is estimated height of the same trees from LiDAR surfaces, and
b_i are regression coefficients.

Cubic foot volume of single trees is estimated with the equation developed by Merrifield and Foil (1967) to predict Minor's cubic volume from:

$$(4)\quad ft^3 = b_o + b_1 (dbh) + b_2 (H_{gr}) + b_3 (dbh^2 H_{gr})$$

Double-sample, regression estimator models

Phase 2 tree measures of dbh and height are used to compute LiDAR estimates of basal area (ft^2) and volume (ft^3) by using field derived dbh-height equations to predict dbh and basal area, and using dbh and height in a standing tree volume equation to predict volume. Thus, double-sample models used in this computer application involve per acre mean estimates of LiDAR-derived basal area and volume for the double-sample models:

(5) $\overline{Y}_{lr}=\overline{y} + \beta(LiBA\text{-}liba)$

(6) $\overline{Y}_{lr}=\overline{y} + \beta(LiVOL\text{-}livol)$

where $\overline{Y}_{lr}$ = linear regression estimate of mean volume per acre from double-sample,
$\overline{y}$ = mean value of volume per acre (y_i) derived from Phase 2 plots,
LiBA = mean LiDAR derived basal area per acre from Phase 1 plots,
liba = mean LiDAR derived basal area per acre (x_i) from Phase 2 plots,
LiVOL = mean LiDAR derived volume per acre from Phase 1 plots,
livol = mean LiDAR derived volume per acre (x_i) from Phase 2 plots, and
β = linear regression slope coefficient for y_i as a function of x_i (volume or basal area).

Required Data Files

The computer application requires Phase 1 LiDAR tree heights, Phase 2 tree data including LiDAR heights of sample trees, Phase 2 regression coefficients for the dbh-height and volume models for each species, dbh file of minimum and maximum dbh limits for each species-product combination, and strata definition of plot numbers and tree age by stratum in comma delimited formats. Each analysis has a user-defined data set name which will be prefixed to all created or generated files and users may name all input data files.

Phase 1 LiDAR tree heights from each of up to two LiDAR data sets, where the file format is (plot#, LiDAR height) and the file names for up to two data sets (ds) are, for example, *datasetname_PH1ds1.csv* and *datasetname_PH1ds2.csv*. Each of the data sets can contain multiple tree heights per plot listed in any order.

Phase 2 data from ground and LiDAR plots, where the file format is (plot#, species code, product code, dbh, height, age, LiHds1, LiHds2) and the file name could be, for example, *datasetname_PH2Tree.csv.* Height is ground measured height if the tree was a sample tree, age is tree age, and LiHds1 and LiHds2 are LiDAR heights from data sets 1 and 2, respectively. Plot trees with a dbh and height were sorted and used to obtain the regression coefficients for equations 1–3. The required Phase 2 ground plot file is a summary of an original field data file of tree and plot data that had a file format (plot X-coordinate, plot Y-coordinate, plot#, species code, product code, azimuth, distance, dbh, height, age, sample tree x-coordinate, sample tree y-coordinate) where the x and y coordinates of the plot center were recorded with a Differential Global Positioning System (DGPS) and computed for sample trees. Calculated sample tree coordinates were used on the LiDAR surfaces to locate “trees” that match the ground sample tree locations.

Phase 2 regression coefficients previously computed by the user for each encountered species are listed in the comma delimited file in the equation sequence 1–4, where the file format is:

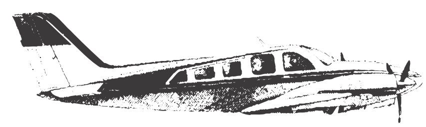
(b_0, b_1, $b_{2,}$, b_3) for equation 1 for each encountered species,

(b_0, b_1, b_2, b_3) for equation 2 for each encountered species,
(b_0, b_1, b_2, b_3) for equation 3 for each encountered species in LiDAR data set 1,
(b_0, b_1, b_2, b_3) for equation 3 for each encountered species in LiDAR data set 2,
(b_0, b_1, b_2, b_3) for equation 4 for each encountered species,
and the file name could, for example, be *datasetname_Coeffic.csv*. The regression coefficients are stacked in the file for each encountered species. Non-applicable coefficients are entered as 0 or 1 depending upon whether the coefficient exists and is used in the model. For example, if equations 1–3 have no intercept, the value of b_0 should be 0. If the exponent coefficient such as b_2 in equations 1–3 was not used, the value should be set to 1. If b_3 was not used in the model, the value should be set to 0.

DBH minimum and maximum values for each combination of species and product class are required where the file format is (species code, product code, minimum dbh, maximum dbh) and the file name could, for example, be *datasetname_DBH.csv*. An undefined species-product class would have a minimum and maximum dbh of 0.

Strata definition and average age lists the stratum number, beginning and ending plot numbers that define the stratum, and average age of the stratum, if used. The file format is (stratum#, beginning plot#, ending plot#, average age) and the file name could, for example, be *datasetname_ Strata.csv*. Each stratum can be defined by more than one sequence of beginning and ending plot numbers and average age can be set to zero if age is not an equation variable.

Process Flow and Outputs



Computer Application Menu

The computer application is a Windows®-based menu driven system (Figure 1). The user must complete the menu items in sequence from steps 1 through 9. Menu items 5–9 can be executed individually or menu item 10 will execute items 5–9 in sequential order if menu items 1 through 4 have been previously fulfilled. The system will not allow the user to execute a menu item unless the previous required items have been completed.

1. Data Set

The user enters a unique data set name that is used as a prefix to all system generated (intermediate and output) files. If the data set name already exists, the user is prompted to enter a new name or asked for permission to overwrite the previous data sets. During the installation process, the user defines the primary directory path where software program files are located and the system automatically creates a subdirectory named DATA FILES within the primary directory where user- and system-generated data are stored.

2. User Parameters

The user must define various parameters, conversion factors, and counts that are used by the system during the computations. The following is a list of parameters defined by the user:

PACF is the per acre conversion factor for expanding the tree tally on a phase 1 or 2 plot. PACF is the reciprocal of the plot size; e.g., a 0.05 acre plot has a PACF of 1/0.05 = 20.

maxSpecies is the maximum number of species codes in the data sets and number of lines of coefficients for equations 1-4 that will be listed in the coefficient file (*coeff.csv*).

maxProducts is the maximum number of product codes in the data sets and number of species-product combinations to be listed in the *dbh.csv* file. For example, if pulpwood, chip'n saw, and sawtimber are the defined products for a single species, then all species would have three products and each product would have a dbh definition line in the *dbh.csv* file. An undefined product would have a minimum and maximum dbh of 0.

maxIterations is the maximum number of iterations in the simulation phase where Phase 2 LiDAR heights are allocated to the species-product classes on a matching Phase 2 ground plot on the basis of percent distribution of number of stems on the Phase 2 ground plot. The optimum number of iterations is generally between 100 and 500.

maxPlots is the maximum plot number encountered in the data sets; not the number of plots. The maximum plot number is used to dimension array space during computations.

LiHgt Adjustment is Yes/No as to whether or not to adjust LiDAR heights to ground height with regression function (3) for ground height as a function of LiDAR height before computing dbh from height with equation 1.

Regression Coefficient Set is selected as 1 or 2. Currently, regression coefficients for models 1-4 are used in the computations; however, future versions of the software will allow other models to be used. The default value is currently set to 1.

3. Phase 2 Regression Coefficients

The regression coefficients in the *coeffic.csv* file are entered and edited with this option or the file can be created in a spread sheet, edited, and saved as a comma delimited text file in the data directory. All files developed with the create file option are prefixed with the data set name and named by the system. The coefficients are stacked in the file by equation number and species. See previous discussion on the *coeffic.csv* required data file.

4. Assign Plots and DBH Files

The user develops the *Strata.csv* file with the create or edit option or with a spreadsheet and imports the file into the data directory. The file is a comma delimited text file containing stratum number, beginning plot number and ending plot numbers to define the stratum, and average age. The user also develops the *DBH.csv* file with the create or edit option or with a spreadsheet. The file is a comma delimited text file containing species code, product code, minimum dbh, and maximum dbh for each species-product combination. All species-product combinations must have a defined minimum and maximum dbh line in the file, even if the values are set to 0.

5. Assign Heights and Volume to Phase 2 Trees

This is the first computation step in the system where plot totals of trees, basal area, and volumes on a per acre basis and percentages by species-product class are computed and a summary of Phase 2 ground estimates are obtained. Single tree volumes are computed for the trees in the *PH2Tree.csv* file using the volume coefficients for equation 4 in the *Coeffic.csv* file. If the tree height was measured on the ground, only the tree volume is computed. If the tree height was not measured on the ground plot, height is computed from dbh with equation (1) and volume is computed with equation (4) using the coefficients in the *Coeffic.csv* file. The file format of the new output file (*datasetname_PH2TreeV.csv*) is:

(plot, species, product, dbh, height, volume, LiDAR1 height, LiDAR2 height)

where the LiDAR1 and LiDAR2 height variables are the LiDAR heights from data sets 1 and 2, respectively. If there is only one data set, the value of the data set 2 height is set to 0. The format of the *PH2TreeV.csv* file is essentially the same as the *PH2Tree.csv* file with a volume column added after the height variable.

Plot totals and percent distribution of numbers of trees, basal area, and volume for each species-product class are computed during this option and results written to a comma delimited text file named *datasetname_ Ph2Plot.csv*. Plot totals for all species-product classes combined are written to text file named *datasetname_PH2PlotT.csv*.

A summary of species-product totals and percent distributions for the total data set is stored in a comma delimited text file named *datasetname_PH2Sum.csv*. The *PH2Sum.csv* file is used in menu option 9 to allocate volume estimates from the linear regression procedures to species-product classes in an unstratified and combined stratum.

6. Iteration

LiDAR heights in the Phase 1 data set are randomly allocated in a Monte Carlo simulation to species-product classes on each matching Phase 2 ground plot on the basis of percent distribution by numbers of trees on the ground plot. Percent distributions of trees/ac by species-product class are obtained for each Phase 2 ground plot from the *PH2Plot.csv* file and the probabilities of occurrence are computed and ordered (highest to lowest) for each species-product class. Phase 1 LiDAR heights for the same plot are obtained from the *PH1ds1.csv* or the *PH1ds2.csv* data sets. The Phase 1 heights are allocated to the species-product classes a total of “maxIteration” times and mean values were computed. The number of trees, basal area, and volume on a per acre basis are written to the output files *PH2Plotds1.csv* or *PH2Plotds2.csv* for LiDAR data sets 1 and 2, respectively.

7. Compute Phase 1 Species, N, BA and Volume

LiDAR heights from Phase 1 data files *PH1ds1.csv* and/or *PH1ds2.csv* are randomly assigned to species classes based on the percent distribution by species (in terms of number of trees per acre) in each stratum from the Phase 2 ground plot data. As each LiDAR height from the Phase 1 data file is read from the file, a species code is “assigned” based on a random assignment to the probability distribution for the stratum. The height, dbh, and volume function coefficients for the “assigned” species in the *Coeffic.csv* file are used to compute adjusted ground height, dbh, basal area, and tree volume for each LiDAR height in the Phase 1 plots. The resulting text files named *datasetname_Phase1ds1.csv* or *datasetname_Phase1ds2.csv* have a comma delimited format of:

(plot#, species, trees/ac, basal area/ac, volume/ac)

8. Combine Phase 2 ground and LiDAR estimates

The previous Phase 1 LiDAR and Phase 2 ground data files are combined and the results written to a comma delimited text file named *datasetname_Phase2.csv* of the format:

(plot#, GN, GBA, GVol, Li1N, Li1BA, Li1Vol, Li2N, Li2BA, Li2Vol)

where: GN, GBA, and GVol are ground plot estimates of trees/ac, basal area/ac, and volume/ac, respectively;

Li1N, Li1BA, Li1Vol are LiDAR data set 1 estimates of trees, basal area and volume, respectively; and Li2N, Li2BA, Li2Vol are LiDAR data set 2 estimates of trees, basal area and volume, respectively.

If only one LiDAR data set is used, the data set 2 values are omitted. This file is used to compute the regression relationships between ground volume (y_i= GVol) and LiDAR basal area (x_i= LiBA) or LiDAR volume (x_i=LiVOl).

A comma delimited summary file of Phase 1and Phase 2 estimates of trees, basal area, and volume on a per acre basis for each LiDAR data set is created with the name of *datasetname_VolSum.txt* (Table 1). This summary file can be manipulated with a spreadsheet or viewed and printed from menu item 11.

9. Compute and Allocate Double-Sample Estimates

Regression estimates are obtained for the ground volume as a linear function of LiDAR basal area and LiDAR volume for up to 2 LiDAR data sets. The double-sample, linear regression estimates of mean volume per acre are computed from the linear functions 5 and 6 for each stratum and combined strata (Table 2).

Regression estimates of mean volume per acre are obtained for nonstratified data (i.e all data combined), for each defined stratum (by stratum and plot number in *Strata.csv*), and combined strata estimates (with stratified random sample procedures) and printed to the text files *datasetname_Regress.txt* and *datasetname_StratVol.txt.* Regression slope, index of fit, linear regression volume estimate, standard error of the linear regression estimate, sampling error at the 95% level of confidence, Phase 1 and 2 means for the independent variables, numbers of plots used for Phase 1 (N1) and Phase 2 (N2) estimates, and the number of random samples (NRS) versus double-sampling (N1 and N2 plots) are printed to the *Regress.txt* file (Table 2). Index of fit is defined as the proportion of total sums of squares explained by regression or (1- SS_{error} /SS_{total}). The combined strata estimates for volume (equation 7) and standard error (equation 8) of each model and LiDAR data set are obtained by summing the weighted stratum estimates as:

(7) $$\bar{Y}_{lr,c} = \sum \frac{(n_{1i} + n_{2i})}{N} \left(\bar{Y}_{lr,i}\right)$$

(8) $$S_{\bar{Y}_{lr,c}} = \left[\sum \left(\frac{n_{1i} + n_{2i}}{N} \right)^2 \left(S^2_{\bar{Y}_{lr,i}} \right) \right]^{0.5}$$

where n_{1i} and n_{2i} are Phase 1 and Phase 2 sample sizes respectively in the i^{th} stratum and N = (n_{1i} and n_{2i}), i=1 to s strata.

Estimated samples sizes for Phase 1 and 2 of the double-sample are calculated with precision statistics from the current analysis and equations from Johnson:

(9) $$N_{rs} = \frac{t^2\, CV\%^2}{AE\%^2}$$

(10) $$n_1 = N_{rs}\left[(1-\rho^2)\sqrt{\left(\frac{c_2}{c_1}\right)\left(\frac{\rho^2}{1-\rho^2}\right)} + \rho^2\right]$$

(11) $$n_2 = N_{rs}\left[(1-\rho^2) + \rho^2\sqrt{\left(\frac{c_1}{c_2}\right)\left(\frac{1-\rho^2}{\rho^2}\right)}\right]$$

where: N_{rs} = sample size for a simple random sample from an infinite population,
CV% = coefficient of variation,
AE% = allowable error (absolute error as a percentage of the mean),
t = Student's t-value at n-1 df and α=0.05,
n_1 = sample size for Phase 1 with cost of c_1 per sample,
n_2 = sample size for Phase 2 with cost of c_2 per sample, and
ρ^2 = coefficient of determination.

The best regression estimate in terms of lowest sampling error is selected from each of the strata (nonstratified, user-defined stratum, and combined strata) estimates and used to partition the mean volume estimate to the species-product classes (Table 3). The mean volume is partitioned to the species-product classes on the basis of percent distribution of basal area and volume on the Phase 2 ground plots in each stratum. Each of the text files by the system can be manipulated with a spreadsheet or viewed and printed with menu item 11.

10. Do Steps 5 thru 9 Consecutively

This menu option executes menu items 5 through 9 in a sequential manner. If an error occurs in the sequence, users are advised to manually execute each step (5 to 9 individually) to determine the procedure where the error occurs. Menu items 5-9 are the consecutive steps necessary to compute the double-sample, after all input items such as regression coefficients and strata plots are defined. Items 1 through 4 must be defined by the user prior to electing menu items 5 through 9, or 10.

11. View Output Files

Summary files may be viewed or printed to paper with this menu option (Figure 2). The summary files available for viewing or printing are:

Volume Sum Text File (*VolSum.txt*)

The volume sum text file (Table 1) is a summary from Phase 1 LiDAR and Phase 2

LiDAR and ground computations of per acre trees, basal area, and volume by species-product class from menu item 8. It contains per acre estimates of number of trees (N), ft^2 basal area, and ft^3 volume and number of encountered plots by species-product class for each sampling phase. The number of Phase 2 plots may differ between LiDAR data sets because not all of the ground plots will have a matching LiDAR plot. If the number of Phase 2 plots differ between data sets, most likely the matching LiDAR plot was either not located or missed during the surfacing and height extraction processes. Failure to record on-the-ground coordinates of the plot center with a DGPS will result in "lost" plots because matching trees for the Phase 1 and Phase 2 plots can not be found. The inventory design for the study where these data were obtained prescribed a 10:1 ratio of LiDAR (Phase 1) to ground (Phase 2) plots, but the exact ratio was not attained because different numbers of plots were located in the two LiDAR data sets. The *VolSum.txt* text file (Table 1) permits the user to observe the differences and/or similarities between the Phase 1 and Phase 2 estimates of variables of interest.

Regression Text File (*Regress.txt*)

The regression text file (Table 2) contains summary data relative to the computation of linear regression, double-sample estimates of mean volume per acre and associated precision statistics for user-defined strata with double-sample models (5) and (6). The dependent variable is mean volume per acre (ft^3/ac) for each of two independent variables, mean LiDAR basal area per acre (ft^2/ac),and mean LiDAR volume per acre (ft^3/ac) for up to two LiDAR data sets. The results are presented in two tiers for each user-defined stratum, unstratified combined data, and stratified combined strata. In the first stratum tier, results are listed by a dependent-independent variable combination for each of two LiDAR data sets for models (5) and (6). The results include estimates of the slope coefficient (Beta), index of fit (IdxFt) for the linear regression equation, adjusted linear regression estimate of mean volume per acre (YbarLR), standard error of the regression estimate (S_ybarLR), correlation coefficient (Rho) between the dependent and independent variables, sampling error (SE%) at the 95% level of confidence, Phase 1 sample size (N1), and Phase 2 LiDAR (N2L) and ground (N2G) sample sizes. The second tier of results for a stratum lists Phase 1 and Phase 2 means of the independent variables LiDAR basal area and volume and the dependent variable ground volume, coefficient of variation (CV%) of the plot data, and sample size estimates for simple random sampling (N_rs) versus Phase 1 (N1) and Phase 2 (N2) sample sizes for a double-sample. Sample sizes are computed for a double sample assuming a 10:1 cost ratio between Phase 2 and Phase 1 plots and the calculated precision statistics for the stratum.

The last tier in the regression text file contains the combined strata estimates of the linear regression estimate of volume per acre and its associated standard error and sampling error (α =0.05). Combined strata estimates are obtained by summing the weighted stratum estimates with equations 7 and 8. The regression text file allows the user to determine sampling gains with stratification for up to two sets of LiDAR data.

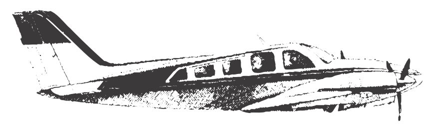
Strata Volume Text File (StratVol.txt)

The strata volume text file (Table 3) uses the "best" regression estimate in terms of lowest sampling error (SE%) from each stratum (in Table 2) at the 95% level of confidence and partitions the volume estimate into species-product classes using the average percent distribution of basal area (%BA Dis) and volume (%Vol Dis) within the stratum. In most situations, partitioning with percent volume distribution should produce the more realistic estimates of species-product volumes. If desired, the percent distribution of basal area and/or volume could be computed in a spreadsheet by species-product class from values in the percent distribution columns. The strata volume text file allows the user to see the best stratum estimate of mean volume per acre and its associated errors and the best estimates of species-product volumes.

Review Instructions

The brief instructions on data file contents and structures and output files are stored in a HTML help file and can be reviewed with menu item 12.

Discussion

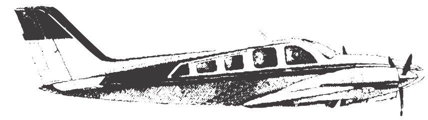


The LiDAR Double-Sample automation system automates the inventory and statistical computations for LiDAR data that have been previously processed to yield tree heights by plot and ground data that have been analyzed to yield regression coefficients for tree dbh and height relationships. Surfacing raw LiDAR data to produce a canopy and ground surface, interpreting the canopy surface for tree locations, and obtaining tree heights by plot location is an enormous task and this automation system does nothing to reduce the work load on the LiDAR data side of the process. It does, however, provide a set of data formats and computational procedures that facilitate the rapid computation of a LiDAR-based double-sample forest inventory. The LIDARDS system allows the user to set the number of iterations for the Monte Carlo simulation of species-product distribution on Phase 2 plots and whether to adjust LiDAR heights to ground heights with equation 3 before estimating dbh with equation 1. Scientists found the LiDAR height to ground height adjustment process for high- and low-density LiDAR on smoothed and unsmoothed surfaces increased the sampling error of the volume estimates. By turning the “height adjustment” procedure on and off, the user can determine the effects of adjusting LiDAR height to ground height for each data set.

The LiDAR Double-Sample automation system assumes the user can manipulate the raw data in spreadsheet software and save required data files in a comma delimited, text format. The LIDARDS system also assumes the user has an appropriate regression package that can produce the coefficients for the required tree dbh-height equations 1-3 from the ground data. An inherent disadvantage to the current system is the fixed models for the tree equations. Future versions of the LIDARDS system will offer other dbh-height models.

The LiDAR Double-Sample automation system runs well under the current Window environments and results are written to text files that can be accessed and manipulated with most word processor, text editor, and spreadsheet software.

Literature Cited



Avery, T.E., and H.E. Burkhart. 2002. Forest Measurements. 5th Edition. McGraw-Hill, New York, N.Y. 456 pp.

Johnson, E.W. 2000. Forest Sampling Desk Reference. CRC Press, Boca Raton, FL. 985 pp.

Lefsky, M.A., W.B. Cohen, S.A. Acker, G. Parker, T.A Spies, and D. Harding. 1999. LiDAR remote sensing of the canopy structure and biophysical properties of Douglas-fir western hemlock forests. Remote Sensing of Environ. 70:339-361.

Magnussen, S., and P. Boudewyn. 1998. Derivations of stand heights from airborne laser scanner data with canopy-based quantile estimators. Can. J. For. Res. 28:1016-1031.

McCombs, J.W., S.D. Roberts, and D.L. Evans. 2003. Influence of fusing lidar and multispectral imagery on remotely sensed estimates of stand density and mean tree height in a managed loblolly pine plantation. For. Sci. 49(3):457-466.

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

Merrifield, R.G. and R.R. Foil. 1967. Volume equations for southern pine pulpwood. Louisiana State Univ. North Louisiana Hill Farm Expt. Sta. Bul. Forestry 7. Homer, Louisiana.

Minor, C.O. 1950. Tables for estimating cubic volume of timber. Louisiana Agric. Expt. Sta. Bul. No. 445. 39 pp.

Nelson, R., W. Krabill, and J. Tonelli. 1988. Estimating forest biomass and volume using airborne laser data. Remote Sensing of Environ. 24:247-267.

Nilsson, M. 1996. Estimation of tree heights and stand volume using an airborne LIDAR system. Remote Sensing of Environ. 56:1-7.

Parker, R.C., and D.L. Evans. 2004. An application of LiDAR in a double-sample forest inventory. West. Jour. of Appl. For. 19(2):95-101.

Parker, R.C., and P.A. Glass. 2004. High versus low density LiDAR in a double-sample forest inventory. Southern Journal of Applied Forestry 28(4):205-210.

Parker, R.C., and A.L. Mitchel. 2005. Smoothed versus unsmoothed LiDar in a double-sample forest inventory. Southern Journal of Applied Forestry 29(1):40-47.

Figure 1. Computer screen for user input parameters for the LiDAR Double-Sample automation system.

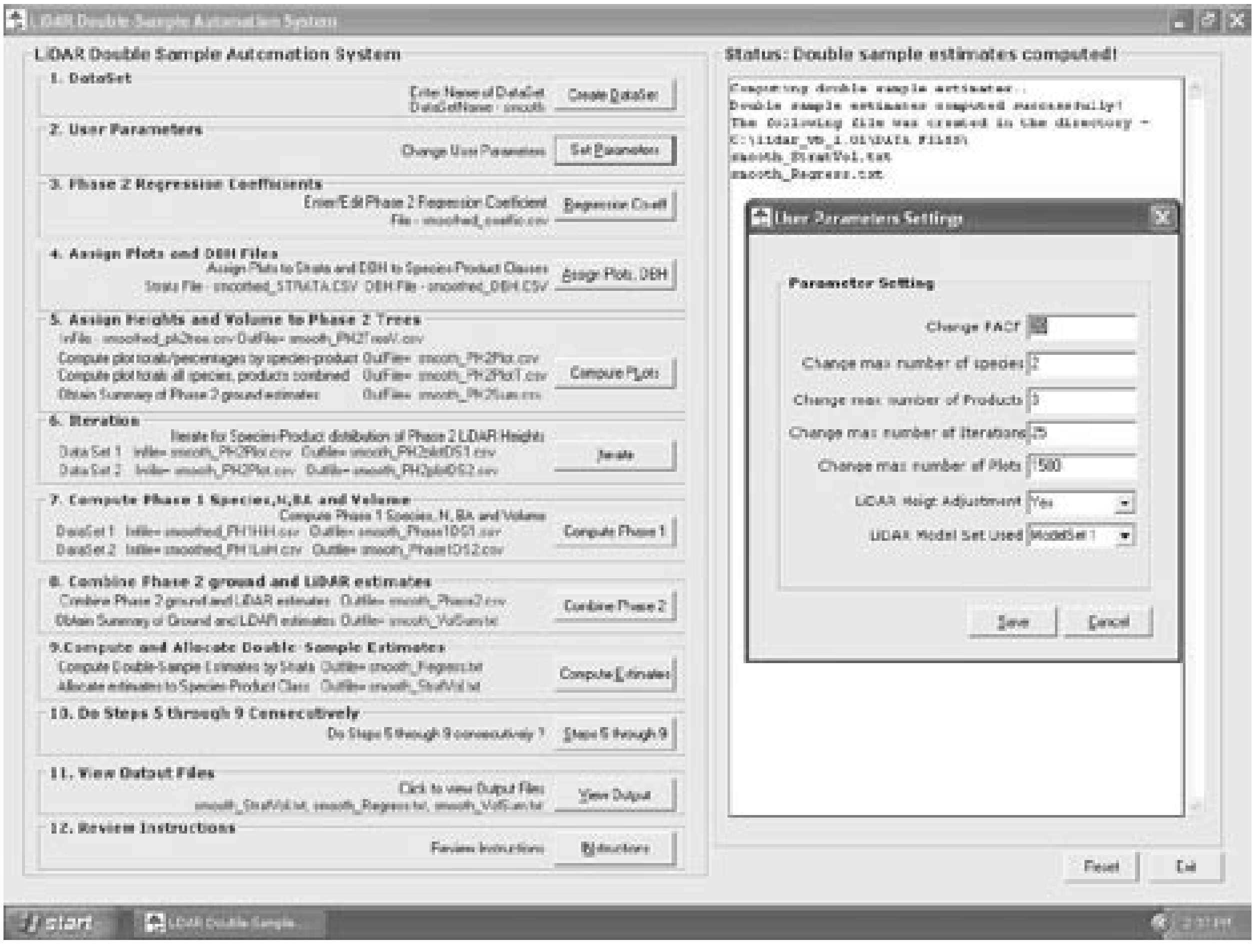


Figure 2. Computer screen for selecting output files to view from the LiDAR Double-Sample automation system.

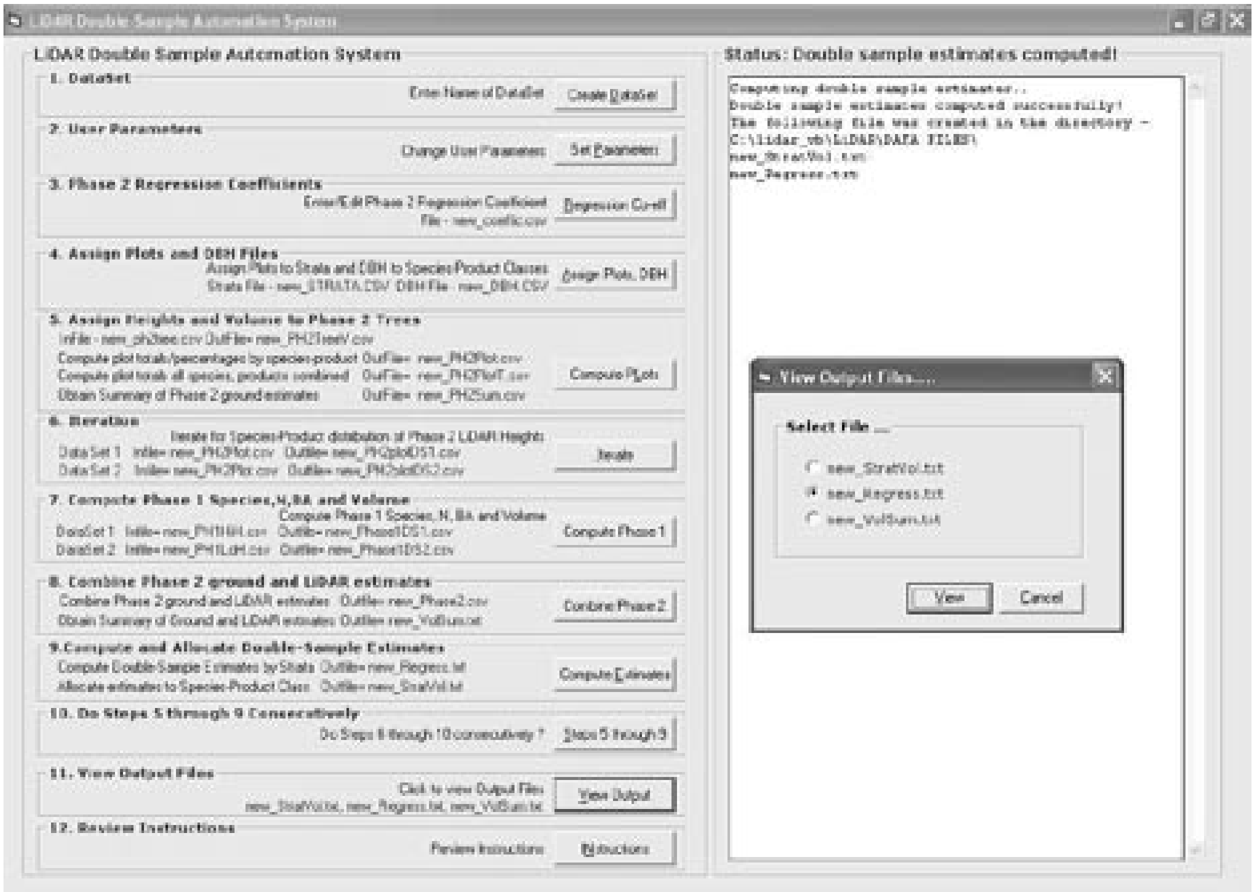


Table 1. Sample volume sum text file (VolSum.txt) of per acre estimates of numbers of trees (N), basal area (BA), and volume (Vol) and sample sizes (plots) for Phase 1 and 2 with two LiDAR data sets (ds) of double-sample inventory data in the LiDAR Double-Sample automation system.

Phase 2 Ground					
Species[a]	Product[b]	N	BA	Vol	plots
1	1	57.45	14.43	328.85	141
1	2	36.60	17.71	490.96	141
1	3	18.72	27.87	1043.41	141
2	1	2.95	7.84	154.57	141
2	3	5.25	7.33	210.51	141
Totals		144.97	105.95	2228.30	

Phase 2 LiDAR-ds1:					
Species[a]	Product[b]	N	BA	Vol	plots
1	1	71.23	24.07	601.64	132
1	2	19.77	7.49	188.79	132
1	3	31.47	38.81	1453.12	132
2	1	23.90	27.88	887.78	132
2	3	4.08	7.70	266.25	132
Totals		150.45	105.95	3397.58	

Phase 2 LiDAR-ds2:					
Species[a]	Product[b]	N	BA	Vol	plots
1	1	80.31	24.44	588.05	131
1	2	18.27	6.94	175.39	131
1	3	37.18	47.97	1799.17	131
2	1	26.47	30.73	974.33	131
2	3	5.11	9.87	343.76	131
Totals		167.34	119.94	3880.69	

Table 1. Sample volume sum text file (VolSum.txt) of per acre estimates of numbers of trees (N), basal area (BA), and volume (Vol) and sample sizes (plots) for Phase 1 and 2 with two LiDAR data sets (ds) of double-sample inventory data in the LiDAR Double-Sample automation system (continued).

Phase 1 LiDAR-ds1					
Species[a]	Product[b]	N	BA	Vol	plots
1	All	116.07	66.84	2035.99	1385
1	All	38.27	32.91	1100.37	1385
Totals		154.34	99.75	3136.36	

Phase 1 LiDAR-ds1					
Species[a]	Product[b]	N	BA	Vol	plots
1	All	119.07	72.12	2249.93	1349
1	All	41.85	37.67	1267.57	1349
Totals		150.92	109.79	3526.5	

[a] species 1 = pine, species 2 = hardwood
[b] product 1 = pulpwood, product 2 = chip'n saw, product 3 = sawtimber

Table 2. Sample regression text file[a] (Regress.txt) for unstratified (stratum 0), stratum 2, and combined strata (Co) double-sample inventory data for two LiDAR data sets (ds) from the LiDAR Double-Sample automation system.

St	Depend	Indpend	Beta	IdxFt	YbarLRS	S_ybarLR	Rho	SE%	N1	N2L	N2G
0	GrnVol	LiBA-ds1	12.0426	51.3	2153.5	94.52	0.7161	8.68	1385	132	141
0	GrnVol	LiVOL-ds1	0.3051	53.4	2148.5	92.53	0.7161	8.52	1385	132	141
0	GrnVol	LiBA-ds1	10.3343	62.6	2123.4	83.74	0.7161	7.80	1349	131	141
0	GrnVol	LiVOL-ds1	0.2640	62.8	2134.7	83.41	0.7926	7.73	1349	131	141

St	Variable	Phase1	Phase2	CV%	N1_rs	N1	N2
0	LiBA-Ds1 (sq.ft)	99.7	106.0	70.35	194	406	125
0	LiVOL-Ds1 (cu.ft)	3136.4	3397.6	70.35	194	409	121
0	LiBA-Ds2 (sq.ft)	109.8	119.9	70.46	194	419	103
0	LiVOL-Ds2 (cu.ft)	3526.5	3880.7	70.46	194	419	102
	Grn Volume-ds1		2281.3				
	Grn Volume-ds2		2285.6				
	Ground Volume		2228.2				

St	Depend	Indpend	Beta	IdxFt	YbarLRS	S_ybarLR	Rho	SE%	N1	N2L	N2G
2	GrnVol	LiBA-ds1	1.0705	57.0	2030.5	170.99	0.7548	17.00	471	44	52
2	GrnVol	LiVOL-ds1	0.2625	56.1	2019.3	172.90	0.7487	17.28	471	44	52
2	GrnVol	LiBA-ds2	10.7947	62.7	2094.5	162.30	0.7921	15.65	440	43	52
2	GrnVol	LiVOL-ds2	0.2604	60.6	2106.2	166.81	0.7783	16.00	440	43	52

St	Variable	Phase1	Phase2	CV%	N1_rs	N1	N2
2	LiBA-Ds1 (sq.ft)	94.4	104.8	77.83	247	527	145
2	LiVOL-Ds1 (cu.ft)	3073.4	3555.1	77.83	247	526	147
2	LiBA-Ds2 (sq.ft)	104.3	109.1	78.24	250	539	131
2	LiVOL-Ds2 (cu.ft)	3537.4	3689.0	78.24	250	537	137
	Grn Volume-ds1		2248.6				
	Grn Volume-ds2		2261.0				
	Ground Volume		2145.7				

Table 2. Sample regression text file[a] (Regress.txt) for unstratified (stratum 0), stratum 2, and combined strata (Co) double-sample inventory data for two LiDAR data sets (ds) from the LiDAR Double-Sample automation system (continued).

St	Depend	Indpend	YbarLRS	S_ybarLR	SE%
Co	GrnVol	LiBA-ds1	2180.9	91.12	8.20
Co	GrnVol	LiVOL-ds1	2168.0	90.24	8.17
Co	GrnVol	LiBA-ds2	2154.6	83.80	7.63
Co	GrnVol	LiVOL-ds2	2154.5	83.92	7.64

[a]St = stratum where 0 is unstratified and Co is combined strata estimate
Depend = dependent variable (y)
Independ = independent variable (x)
Beta = slope coefficient for ground volume (y) as a linear function of LiDAR basal area (x) or volume (x)
IdxFt = index of fit for the linear regression equation
YbarLR = adjusted linear regression estimate of mean volume per acre (ft^3)
S_ybarLR = standard error of the regression estimate
Rho = correlation coefficient between the dependent and independent variables
SE% = sampling error at the 95% level of confidence
N1 = Phase 1 sample size
N2L = Phase 2 LiDAR sample size
N2G = Phase 2 ground sample size
CV% = coefficient of variation
N_rs = sample size for simple random sample to achieve a 10% sampling error at $\alpha=0.05$
N1 = Phase 1 double-sample size to achieve a 10% sampling error assuming $c_1=1$ for Phase 1 plots
N2 = Phase 2 double-sample size to achieve a 10% sampling error assuming $c_2 = 10$ for Phase 2 plots

Table 3. Sample strata volume text file[a] (StratVol.txt) of double-sample inventory data for non-stratified (non-strat), stratum 2, and combined strata from the LiDAR Double-Sample automation system.

STRATA	YBarLR	S_yBarLR	SE%	Sp	Pr	%BADis	%VolDis	N	BA	Hgt	Vol
non-strat	2134.73	83.41	7.73	1	1	409.7	315.0	57	14.4	52.7	329
				1	2	502.9	470.3	37	17.7	61.7	491
				1	3	791.4	999.6	19	27.9	85.4	1043
				Sub		1704.0	1785.0	113	60.0	61.0	1863
				2	1	222.6	148.1	27	7.8	46.9	155
				2	3	208.1	201.7	5	7.3	71.6	211
				Sub		430.8	349.7	32	15.2	50.9	365
				Total		2134.7	2134.7	145	75.2	58.8	2228
non-strat	2094.46	162.30	15.65	1	1	257.3	189.3	45	10.6	51.3	234
				1	2	510.6	480.6	40	21.0	63.3	595
				1	3	714.4	893.6	20	29.4	84.3	1107
				Sub		1482.3	1563.4	106	61.0	62.3	1937
				2	1	236.8	158.7	34	9.9	47.6	197
				2	3	372.4	372.3	10	15.3	75.0	461
				Sub		612.2	531.0	44	25.2	53.7	658
				Total		2094.5	2094.5	149	86.2	59.8	2595
non-strat	2094.46	162.30	15.65	1	1	413.6	318.0	57	14.4	52.7	329
				1	2	507.6	474.7	37	17.7	61.7	491
				1	3	798.7	1008.9	19	27.9	85.4	1043
				Sub		1719.8	1801.6	113	60.0	61.0	1863
				2	1	224.7	149.5	27	7.8	46.9	155
				2	3	210.1	203.5	5	7.3	71.6	211
				Sub		434.8	353.0	32	15.2	50.9	365
				Total		2154.6	2154.6	145	75.2	58.8	2228

[a]YbarLR = adjusted linear regression estimate of mean volume per acre (ft^3)
S_ybarLR = standard error of the regression estimate (ft^3)
SE% = sampling error at the 95% level of confidence
Sp = species code (1=pine, 2=hardwood)
Pr = product code (1=pulpwood, 2=chip'n saw, 3=sawtimber)
%BADis = regression estimate allocated on the basis of percent basal area distribution
%VolDis = regression estimate allocated on the basis of percent volume distribution
N = number of trees per acre on Phase 2 ground plots
BA = basal area per acre (ft^2)on Phase 2 ground plots
Hgt = average tree height on Phase 2 ground plots
Vol = average volume per acre (ft^3) on Phase 2 ground plots

FWRC

Mississippi State
UNIVERSITY