Forests **2012**, *3*, 179-189; doi:10.3390/f3020179

OPEN ACCESS

forests

ISSN 1999-4907

www.mdpi.com/journal/forests

Article

Improving Woody Biomass Estimation Efficiency Using Double Sampling

David L. Parrott [1], John M. Lhotka [1,*], Songlin Fei [2] and B. Scott Shouse [3]

[1] Department of Forestry, University of Kentucky, 210 Thomas Poe Cooper Building, Lexington, KY 40546, USA; E-Mail: david.parrott2@gmail.com

[2] Department of Forestry and Natural Resources, Purdue University, 715 West State Street, West Lafayette, IN 47907, USA; E-Mail: sfei@purdue.edu

[3] Mountain Association for Community Economic Development (MACED), 433 Chestnut Street, Berea, KY 40403, USA; E-Mail: bsshouse@maced.org

* Author to whom correspondence should be addressed; E-Mail: john.lhotka@uky.edu; Tel.: +1-859-257-9701; Fax: +1-859-323-1031.

Received: 20 March 2012; in revised form: 19 April 2012 / Accepted: 20 April 2012 / Published: 3 May 2012

Abstract: Although double sampling has been shown to be an effective method to estimate timber volume in forest inventories, only a limited body of research has tested the effectiveness of double sampling on forest biomass estimation. From forest biomass inventories collected over 9,683 ha using systematic point sampling, we examined how a double sampling scheme would have affected precision and efficiency in these biomass inventories. Our results indicated that double sample methods would have yielded biomass estimations with similar precision as systematic point sampling when the small sample was ≥20% of the large sample. When the small to large sample time ratio was 3:1, relative efficiency (a combined measure of time and precision) was highest when the small sample was a 30% subsample of the large sample. At a 30% double sample intensity, there was a <3% deviation from the original percent margin of error and almost half the required time. Results suggest that double sampling can be an efficient tool for natural resource managers to estimate forest biomass.

Keywords: forest biomass; inventory design; prism sampling; point double sample

---

1. Introduction

The introduction of carbon markets and the potential for future bioenergy markets have heightened interest in quantifying forest biomass. In order to measure forest biomass, a wide array of sampling designs is available. Among the possibilities, double sample designs have been used to increase inventory efficiency for timber volume estimation [1] and could prove effective for forest biomass estimation.

Double sampling is designed to lower inventory times by providing an estimate of a target variable by utilizing a highly correlated auxiliary variable that is easy to measure [2]. Double sampling requires the sampling of two sets of points: a small set of sample points where the target and auxiliary variables are measured and a large set of sample points where only the auxiliary variable is measured. These two sets of points can be separate, or the small sample can be a subset of the large sample. The latter arrangement can save time by reducing travel between points since the small sample points contribute to both the large and small sample populations. Once the relationship between the target and auxiliary variable is calculated from the small sample data, regression analysis or a ratio estimator can be used to estimate the target variable from the auxiliary variable collected in the large sample.

Forest inventory research evaluating double sampling has primarily focused on using point double sampling with timber volume as the target variable and basal area as the auxiliary variable. Basal area is a logical auxiliary variable since it does not require direct measurement of tree diameters when performing point (*i.e.*, prism) sampling and is highly correlated with forest stand volume [3]. Due to this high correlation, studies have shown that double sampling can be effective at estimating standing timber volume [1,3–5]. Double sampling research related to timber volume estimation has investigated the number of plots or points necessary, time saved, and associated error and has led to various application methodologies.

While studies have concentrated on the efficiency of double sampling for timber volume estimation, limited information exists regarding the effectiveness of double sampling to estimate forest biomass. Since tree biomass, like volume, is correlated to diameter, it would seem feasible that double sampling could also be applicable to forest biomass inventories. The objective of this study was to take a collection of biomass inventories completed using a systematic point sampling design and perform a retrospective analysis to determine if a double sampling design would have affected inventory precision and efficiency.

2. Methods

2.1. Description of Data

The Mountain Association for Community Economic Development (MACED) provided biomass inventory data collected from 2007 to 2010. Inventories were from properties located in eastern Kentucky, U.S. containing mixed broadleaf forest types typically associated with the Appalachian region of the United States. Approximately 10,750 ha are enrolled in MACED’s forestry program with the potential of 87,817 ha to be added. MACED’s framework requires enrolled properties to have forest biomass (and carbon) inventories completed every ten years. This makes MACED’s forestry program an example of a large scale organization that would benefit from the most efficient methodology to inventory forest biomass. Increased efficiency and decreased costs would reduce the resources required for monitoring this large assemblage of forestland.

Available biomass inventories were completed using a systematic point sample design that targeted a property-level basal area estimate with a ≤10% percent margin of error (*i.e.*, cruise precision) at a 95% confidence interval. For these point sample inventories, a BAF 10 prism was used to measure trees with dbh ≥ 19.1 cm. In-plot procedures for biomass calculations included the identification of tree species and dbh measurement. Aboveground dry forest biomass was estimated for trees at each sample point using allometric equations provided by Jenkins *et al.* [6] wherein tree biomass is derived as a function of dbh with regard to species groups:

$$bm = \text{Exp}\left(\beta_0 + \beta_1 \ln dbh\right) \tag{1}$$

where bm = total aboveground woody biomass (kg), dbh = diameter at breast height (cm), and β_0 and β_1 are parameters specific to species groups from Jenkins *et al.* [6]. Property inventories quantified trees <19.1 cm dbh using fixed radius plots. As a result, our point sampling analysis could not incorporate trees <19.1 cm. Property-level woody biomass estimates, including those trees sampled in the BAF 10 prism and fixed radius plots, showed that trees >19.1 cm accounted for more than two thirds of the total aboveground woody biomass among the properties.

For analysis, 40 property inventories that met the following two criteria were selected from the MACED database: (1) original point sample intensities were sufficient to achieve a basal area estimate with a percent margin of error ≤10% at a 95% confidence; and (2) Inventoried forests were classified as oak/hickory as defined by the USDA Forest Service Forest Inventory and Analysis unit [7]. Of these selected inventories, property sizes ranged from 31.2 to 1,155.4 ha (Table 1) and covered 9,682.9 ha in total. The total number of sample points varied among the 40 properties due to different property sizes and variance in stand structure. Among all the properties, upland hardwood site index estimates ranged from 18 to 30 m. The dominant tree species among properties were *Acer rubrum* L., *Acer saccharum* Marsh., *Betula* spp., *Carya* spp., *Fagus grandifolia* Ehrh., and *Quercus* spp. Descriptive statistics of the original point sample inventories are summarized in Table 1 and individual property characteristics are presented in Appendix 1.

Table 1. Descriptive statistics for original systematic point sample inventories of the 40 properties used in double sample analysis.

Variable	Mean	Min	Max	SD
Area (ha)	242.1	31.2	1155.4	227.3
Points sampled	104.0	47.0	226.0	49.0
Basal area ($m^2\ ha^{-1}$)	21.3	17.4	30.6	2.5
Average dbh (cm)	31.0	25.7	36.1	2.3
Biomass ($mt\ ha^{-1}$)	144.9	114.8	202.9	17.1
Biomass margin of error (%)	7.4	3.2	12.0	2.2

2.2. Analysis

To compare outcomes of the systematic point sample inventories to the outcomes of double sampling, we first determined the precision of the original biomass inventories completed on the

40 selected properties. Standard error and percent margin of error associated with aboveground dry biomass ha^{-1} were calculated using the following equations, respectively:

$$\text{SE} = \frac{s}{\sqrt{n}} \tag{2}$$

where s = standard deviation, and n = total number of points and

$$\text{Percent Margin of Error} = \frac{SE \times t_{\alpha,v}}{\bar{X}_{Biomass}} \times 100 \tag{3}$$

where SE = biomass standard error, $t_{\alpha,v}$ = t-statistic for the chosen confidence interval (95%) and appropriate degrees of freedom, and $\bar{X}$ = mean biomass (mt ha^{-1}).

To determine how a double sampling design would have affected the percent margin of error and time required for the biomass inventories on each property, we compared the outcomes of the original inventories and those that double sampling would have yielded. As opposed to re-inventorying these properties with this new approach, we performed a retrospective double sample inventory, within each property, using the original point sample data. In the investigated double sampling scheme, basal area served as the auxiliary variable used to estimate forest biomass. Following recommendations provided by Avery and Burkhart [2], a ratio estimator was used rather than regression analysis since there was a linear relationship between basal area and tree biomass that passed approximately through the origin and because the variance in tree biomass increased as basal area increased. Use of a ratio estimator can also simplify calculations for practicing foresters.

Oderwald and Jones [3] presented a methodology to design a point double sample inventory that can substitute for a point sample inventory while achieving the same precision; however, this design requires more measurement points in the double sample inventory than in the point sample inventory. Due to the retrospective nature of this study, the Oderwald and Jones [3] methodology could not be applied. Instead, we investigated a methodology similar to one presented in Dilworth and Bell [8] where the double sample inventories maintained the same number of total points as the original point sample inventories. The large sample in the double sample design included all the points in the original inventory to provide basal area data. A subsample of the original points comprised the small sample that provided basal area along with the species and dbh data necessary to estimate aboveground forest biomass using the Jenkins [6] equations. Data from the small sample points were also collected using a BAF 10 prism. The number of points subsampled to create the small sample was selected as a percentage of each property's large sample (subsampling intensity). Since the optimum small sample subsampling intensity was unknown, a range of percentages, 10 to 90% using 10% increments, was used to evaluate trends in efficiency and precision among different subsampling intensities. Within each property, the number of points associated with each subsampling's intensity was randomly selected without replacement to serve as the small sample. To reduce bias from this random selection, we performed 10 iterations of these selections and used the mean basal area and biomass estimations of these iterations for further calculations.

For each property, a ratio of means (R_m) was calculated from the small sample data as follows [2]:

$$R_m = \frac{\overline{bm_S}}{\overline{ba_S}} \tag{4}$$

where $\overline{bm_s}$ = mean small sample biomass and $\overline{ba}_s$ = mean small sample basal area. Property mean aboveground dry biomass $\left(\overline{bm}_{DS}\right)$ was then estimated using the ratio of means and the mean basal area of the large sample based on the following equation [2]:

$$\overline{bm}_{DS} = R_m \times \overline{ba}_L \tag{5}$$

where R = ratio of means and $\overline{ba}_L$ = large sample mean basal area. Standard errors for double sample inventories were calculated using the following equation [2]:

$$SE_{DS} = \sqrt{\left(1-\left(\frac{n_S}{n_L}\right)\right)\left(\frac{S^2_{bm_S} + R_m{}^2 S^2_{ba_S} - 2R_m C_S}{n_S}\right) + \left(\frac{S^2_{bm_S}}{n_L}\right)} \tag{6}$$

where n_s= number of small sample, n_L = number of points in the large sample, $S^2_{bm_S}$ = small sample biomass variance, $S^2_{ba_S}$ = small sample basal area variance, and C_S = small sample biomass and basal area covariance. Percent margin of error was then calculated using Equation 3. Departure from the original percent margin of error was simply determined by taking the absolute difference between the percent margin of errors obtained from the original inventory and the double sample inventories. The standard error, percent margin of error, and departure from the original percent margin of error were calculated for each property. Among all properties, the mean standard error, percent margin of error, and difference in percent margin of error was calculated for each double sample intensity.

The percent of time saved using double sampling was calculated as:

$$\text{Percent Time Saved} = \left[1-\left(I + \frac{T_L(1-I)}{T_S}\right)\right] \times 100 \tag{7}$$

where I = subsampling intensity (proportion), T_L = time required to perform a large sample point, T_S = time required to perform a small sample point. Based on operational observation of inventories across the 40 sampled properties, small sample points were estimated to take three times longer to complete than large sample points. This estimate was corroborated by Merten *et al.* [1] who found that, within the Appalachian hardwood stands, a BAF 10 prism basal area count averaged 1.86 minutes and points where basal area and volume were measured averaged 6.32 minutes. However, time requirements for small and large sample points likely vary within different forest structures, so we completed a sensitivity analysis using 2:1, 3:1, 4:1, and 6:1 small to large sample point time ratios. Travel time would be unaffected by the double sampling method used in this study since all points would be visited regardless of subsampling intensity.

Relative efficiency provided a comparison of the precision and time associated with the original point and the double sample inventories. Relative efficiency was computed using a variation of the equation presented in Merten *et al.* [1]:

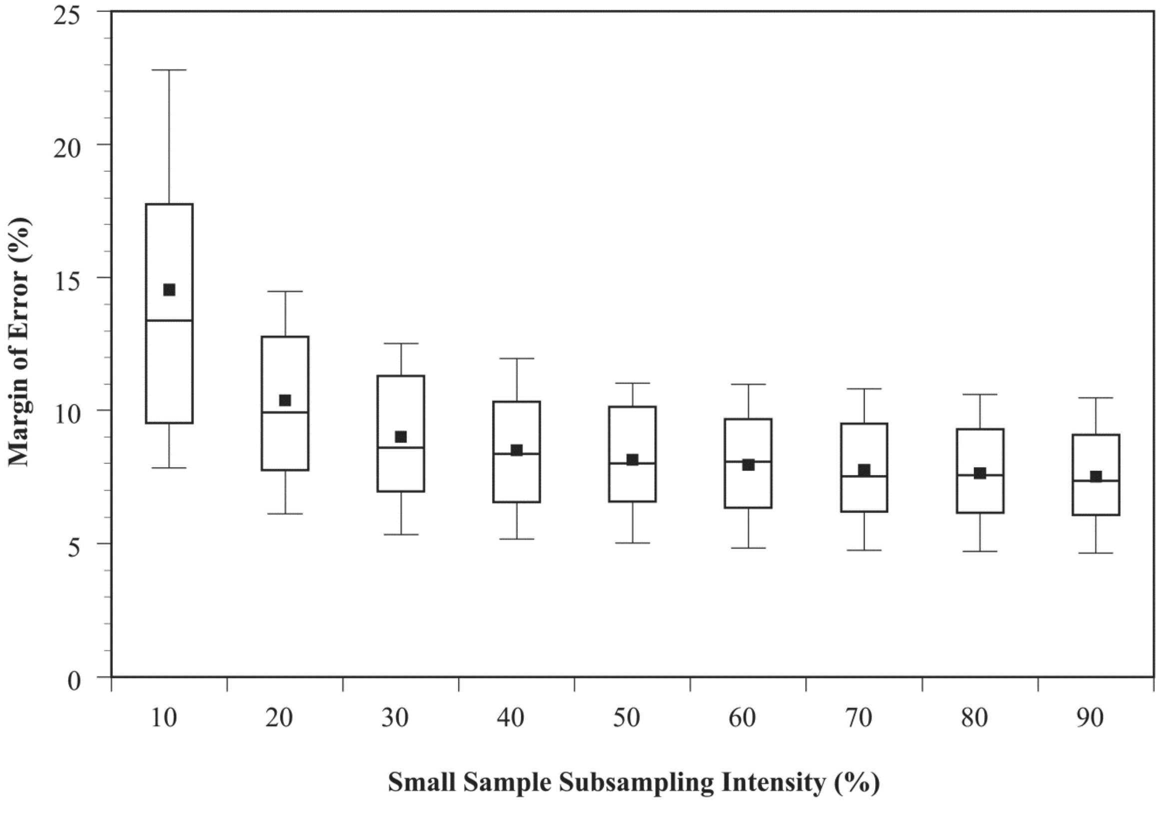
$$RE = \frac{SE^2_{SS} \times T_{SS}}{SE^2_{DS} \times T_{DS}} \times 100 \tag{8}$$

where SE_{SS} = biomass standard error of the original inventory, T_{SS} = time necessary for the original inventory, SE_{DS} = biomass standard error of the double sample inventory, and T_{DS} = time necessary for the double sample inventory. A relative efficiency >100% would be considered more efficient than the original inventory while a relative efficiency <100% would be considered less efficient. For these calculations, 2:1, 3:1, 4:1, and 6:1 small to large sample point time requirement ratios were again considered.

3. Results and Discussion

Original inventory analysis showed that properties averaged 131.4 mt ha^{-1} of aboveground dry biomass in trees >19.1 cm dbh; Standard error and percent margin of error for aboveground woody biomass was 4.57 mt ha^{-1} and 7.43%, respectively. Substitution of the systematic point sampling inventory with a double sampling methodology generally caused minimal departure from the original outcomes and improved efficiency when the small sample subsampling intensity was between 70% and 20% (Table 2). Although percent margin of error increased in the double sample inventories as the intensity of the small sample subsampling decreased, mean departure from the original percent margin of error was ≤3% for small sample point subsampling intensities as low as 20% (Table 2). Variability in the percent margin of error among properties also increased as the small sample subsampling intensity decreased (Figure 1).

Figure 1. Distribution of percent margin of error (aboveground tree biomass) among 40 properties using a double sample design with different intensities of small sample subsampling from the large sample.



In terms of time, every 10% reduction of the small sample subsampling intensity resulted in a 5 to 9% increase in time saved depending on the ratio of time required for small and large sample points (Table 2). When time and precision were combined to measure relative efficiency (Equation 8) and a 3:1 time requirement ratio was assumed, results indicated that efficiency was reduced when the small sample subsampling was >70% and <20% of the large sample. Relative efficiency peaked when the small sample subsampling intensity was 30% (Table 2). At this small sample subsampling intensity, percent margin of error was reduced by <3% and required about half the time of the original inventory. This intensity is slightly higher than the 20 to 25% intensity proposed by Avery and Burkhart [2] but was within the 20 to 40% range that Merten *et al.* [1] found most efficient to estimate timber volume when using a BAF 10 in Appalachian hardwoods stands.

Table 2. Margin of error, deviation from original inventory margin of error, time saved, and relative efficiency for different levels of subsampling and four different small to large sample time requirement ratios in double sample biomass inventories.

			Relative efficiency (%)				**Time saved (%)**			
Small sample intensity (%)	Margin of error (%)	Margin of error deviation (%)	2 to 1 time ratio	3 to 1 time ratio	4 to 1 time ratio	6 to 1 time ratio	2 to 1 time ratio	3 to 1 time ratio	4 to 1 time ratio	6 to 1 time ratio
100 *	7.43	0	100	100	100	100	0	0	0	0
90	7.52	0.1	92	93	94	95	5	7	8	8
80	7.64	0.22	94	98	100	102	10	13	15	17
70	7.76	0.34	97	103	107	110	15	20	23	25
60	7.94	0.52	99	108	113	119	20	27	30	33
50	8.15	0.72	102	115	122	131	25	33	38	42
40	8.5	1.08	102	119	130	143	30	40	45	50
30	9.01	1.58	102	124	139	158	35	47	53	58
20	10.37	2.95	90	116	135	162	40	53	60	67
10	14.54	7.12	68	94	115	150	45	60	68	75

* Small sample point intensity at 100% represents findings of the original point sample inventory on the selected properties.

Evaluation of other time requirement ratios indicated that little efficiency (~2%) could be gained from double sampling when small sample points only required twice the time as large sample points; however, as time requirement ratios increased, relative efficiency increased for all subsampling intensities (Table 2). For the 4:1 and 6:1 time requirement ratios, relative efficiency peaked at 30% and 20%, respectively. However, at the 20% small sample subsampling intensity, mean percent margin of error was >10%, which may fall outside of the acceptable range of error for some biomass inventories. This situation stresses the need for practitioners to evaluate the relative efficiency and margin of error together when deciding on the appropriate subsampling intensity to use when double sampling.

While our study evaluated the loss of precision due to double sampling, Oderwald and Jones [3] presented a methodology in which a double sample scheme can be developed that will yield the same variance as a specified point sample inventory. In their design, more points are included in the double sampling inventory than would be used in the specified standard point sample inventory. This method is only justified when the time necessary to perform small sample points is significantly larger than the time it takes to measure large sample points [3]. When evaluated by Coble and Grogan [9] using a BAF 10 prism in hardwood forests, the Oderwald and Jones [3] method did not result in faster inventory times. Coble and Grogan [9] also evaluated the double sample strategy described by Dilworth and Bell [8] that was similar to our 30% intensity that used a 3:1 ratio of basal area points (*i.e.*, large sample) to measurement points (*i.e.*, small sample). Coble and Grogan [9] found that a 3:1 large to small sample ratio increased efficiency when compared to other double sampling techniques due to reduced time. However, Coble and Grogan [9] found that more inventories fell outside of the acceptable margin of error using the Dilworth and Bell [8] methodology than systematic point sampling or the Oderwald and Jones [3] double sampling methodology. Future research should investigate the efficiency of double sampling biomass inventories using methodologies similar to those presented by Oderwald and Jones [3] where precision in double sampling and point sampling are equivalent.

Findings from this study are based on property inventories within oak/hickory forest types where mean basal area and dbh were 21.3 m^2 ha^{-1} and 31.0 cm, respectively. We cannot speculate whether results of the double sampling and analysis would have been comparable for stands with dissimilar species compositions or diameter distributions. Results presented in this paper were also based on inventories that omitted trees <19.1 cm dbh. In the available biomass inventories, small trees were inventoried using fixed radius plots; therefore, an analysis of point double sampling that included these small stems was not possible. Future research should investigate the use of double sampling in biomass inventories that include smaller diameter trees, use different BAFs, and incorporate nested plot designs with prism sampling and fixed radius plots as these factors may alter the precision and time required to complete forest biomass inventories.

In this study, tree biomass was estimated using the allometric equations presented in Jenkins *et al.* [6], which estimate biomass based on dbh. The use of biomass equations that incorporate additional predictors such as tree height may alter the relationship between basal area and biomass observed in this study. Thus, the use of other biomass equations may also modify the precision and efficiency trends seen in this study. Further research is warranted to explore the effect of allometric biomass equation form on double sampling efficiencies and whether a ratio estimator is appropriate when these alternative equations are employed.

4. Conclusions

Our results suggest that a 30% subsampling intensity for double sample biomass inventories incorporating the equations by Jenkins *et al.* [6] may be optimal within stand types resembling those evaluated in this study. Evaluation of different time requirement ratios indicate that double sampling provides little improvement in inventory efficiency when the time ratio between small and large sample points is 2:1 and that efficiency is maximized at 20% to 30% subsampling intensity when ratios are 4:1 and 6:1. These double sampling results were derived from systematic point sample inventories that were successful in estimating mean property basal area with a margin of error ≤10%. If an inventory has a larger percent margin of error for mean basal area, it is likely that there would be higher variability in forest biomass estimation derived from double sampling. Therefore, if a double sample inventory like the one outlined in this study is to be used, the number of points necessary for the large sample to achieve a suitable estimate of basal area (≤10% percent margin of error) should be established prior to making decisions on the intensity of the small sample. Once the size of the large sample has been determined, the proportion of points to subsample for the small sample can be selected based on the desired precision and inventory time investment.

Although we examined double sampling intensities as high as 90%, for practical use, double sampling should be used with small sample subsampling intensities of 50% or less since this is when the benefit of time saved becomes significant relative to the departure in accuracy. The objectives of the forester performing the biomass inventory should dictate whether it is more valuable to save time over precision and to what degree of precision loss is acceptable. While time requirements and precision associated with double sampling may vary among region, forest type, and stand structure, results from this study show that point double sampling may be useful at improving efficiency and reducing cost of quantifying biomass availability in forest stands.

Acknowledgments

The authors would thank the University of Kentucky, Department of Forestry and the Mountain Association for Community and Economic Development (MACED) for supporting this project.

Conflict of Interest

The authors declare no conflict of interest.

References and Notes

1. Merten, P.R.; Wiant, H.V.; Rennie, J.C. Double sampling saves time when cruising Appalachian hardwoods. *North. J. Appl. For.* **1996**, *13*, 116–118.
2. Avery, T.E.; Burkhart, H.E. *Forest Measurements*; 5th ed.; McGraw-Hill: New York, NY, USA, 2002; pp. 250–252.
3. Oderwald, R.G.; Jones, E. Sample sizes for point, double sampling. *Can. J. For. Res.* **1992**, *22*, 980–983.
4. Oderwald, R.G. Stock and stand tables for point, double sampling with a ratio of means estimator. *Can. J. For. Res.* **1994**, *24*, 2350–2352.
5. Oderwald, R.G. Augmenting inventories with basal area points to acheive desired precision. *Can. J. For. Res.* **2003**, *33*, 1208–1210.
6. Jenkins, J.C.; Chojnacky, D.C.; Heath, L.S.; Birdsey, R.A. National-scale biomass estimators for United States tree species. *For. Sci.* **2003**, *49*, 12–35.
7. Woudenberg, S.W.; Conkling, B.L.; O'Connell, B.M.; LaPoint, E.B.; Turner, J.A.; Waddell, K.L. *The Forest Inventory and Analysis Database: Database Description and Users Manual Version 4.0 for Phase 2*; U.S. Department of Agriculture, Forest Service, Mountain Research Station: Fort Collins, CO, USA, 2009.
8. Bell, J.F.; Dilworth, J.R. *Log Scaling and Timber Cruising*; Oregon State University Book Stores, Inc.: Corvallis, OR, USA, 2007; pp. 182–252.

9. Coble, D.W.; Grogan, J. Comparison of systematic line-point and double sampling designs for pine and hardwood forests in the western gulf. *South. J. Appl. For.* **2007**, *31*, 199–206.

Appendix 1.

Table A1. Property characteristics for large tree (dbh > 19.1 cm) inventories used in double sample analysis. Includes stand area, mean site index, mean dbh, basal area, and aboveground biomass.

Stand	Area (ha)	Site index (m)	Dbh (cm)	Basal area (m^2 ha^{-1})	Aboveground biomass (mt ha^{-1})
1	31.2	26	29.5	17.6	117.6
2	32.4	24	32.3	18.8	136.1
3	32.8	21	28.9	21.6	116.3
4	36.0	27	31.9	22.6	160.7
5	49.8	25	35.8	24.9	169.1
6	54.2	18	30.3	17.4	114.8
7	59.1	25	29.8	21.4	144.1
8	78.1	24	32.0	21.9	150.6
9	88.2	20	29.2	25.5	151.1
10	88.2	29	28.9	21.3	139.0
11	99.1	22	32.1	22.7	161.1
12	100.0	25	28.4	20.0	127.3
13	117.4	21	32.3	19.1	136.3
14	119.0	20	30.1	19.3	118.8
15	120.2	20	32.9	19.3	143.1
16	122.2	22	30.0	19.0	131.1
17	146.1	21	31.8	22.5	158.8
18	161.5	26	29.7	22.6	160.1
19	170.0	24	28.5	24.1	151.0
20	173.6	22	29.4	20.2	135.8
21	174.8	30	30.6	25.1	167.2
22	174.8	23	34.6	22.2	159.0
23	177.3	25	33.2	18.4	125.6
24	177.3	23	33.0	19.5	141.8
25	211.2	26	30.5	21.6	148.9
26	219.3	22	30.6	21.1	148.4
27	235.1	26	31.3	19.4	138.6
28	271.5	19	29.2	22.3	140.6
29	300.3	20	36.0	20.7	149.4
30	300.7	25	30.1	30.6	203.0
31	337.5	25	35.0	22.3	150.5
32	348.0	23	31.9	21.3	149.0
33	399.4	24	33.2	18.4	125.6
34	412.8	21	28.8	21.0	143.8
35	478.7	24	34.1	23.2	165.7
36	538.2	22	30.5	19.4	137.8
37	601.8	23	30.0	20.5	142.0

Table A1. *Cont.*

Stand	Area (ha)	Site index (m)	Dbh (cm)	Basal area (m^2 ha^{-1})	Aboveground biomass (mt ha^{-1})
38	625.2	23	25.7	22.1	149.2
39	664.5	21	26.3	19.5	130.5
40	1155.4	23	32.3	23.3	156.3