Effects of Basal Area Factor and Plot Size on Precision and Accuracy of Forest Inventory Estimates

Peter Becker and Tom Nichols

ABSTRACT: We tested the effects of plot size (0.05-0.30 ac) and basal area factor (BAF) (5-30) on the accuracy and precision of per-acre estimates of tree number, basal area, biomass (all for trees ≥4.5 in. dbh), and sawtimber volume (for trees ≥11.6 in. dbh). Field sampling errors, such as missing in-trees, did not affect our tests. Virtual, variable- and fixed-radius plots were randomly located within an artificial matrix of 130 real plots in well stocked, upland hardwood forests of sawtimber-sized trees in the Missouri Ozarks. Inventory parameters were essentially independent of plot size and BAF, while their coefficients of variation decreased with plot size and increased with BAF. Thus, our results for random plots agreed with sampling theory, unlike a previous study using concentric virtual plots in West Virginia forests. A very concentrated zone of high tree density around some plot centers apparently caused the biased estimates by concentric plots. Compared with the entire composite forest, inventory means were accurately estimated (to within 5%) and size class distributions were well represented for plots ≥0.1 ac or ≤ 15 BAF. Our procedures provide a basis for selecting an efficient and cost-effective sampling design suited to forest characteristics and the inventory’s purpose.

Key Words: Missouri Ozarks, sampling bias, size class distribution, upland hardwood forest

Manuscript received September 25, 2010, accepted April 11, 2011.

Peter Becker (wongbeck@yahoo.com) Eastern Ozarks Forestry Council, P.O. Box 128, Bunker, MO 63629. Tom Nichols, Missouri Department of Conservation, Jefferson City, MO 65102. The Missouri Department of Conservation provided financial support to the Eastern Ozarks Forestry Council in aid of this research. Larry D. Vangilder developed the initial project providing the data for this study, and Randy Jensen oversaw plot establishment plus data collection and management. John R. Brooks, Vicente Monleon, Charles T. Scott, Harry Valentine, James A. Westfall, Harry V. Wiant, Jr., and anonymous reviewers commented helpfully on drafts of this paper. Gail Lucas greatly assisted with kernel smoothing of tree density.

Copyright © 2011 by the Society of American Foresters.
Brooks and McGill (2004) developed a computer-based, virtual sampling of real forest inventory plots in mature hardwood stands of West Virginia to assess how accuracy and precision of per-acre estimates of tree number, basal area, and volume varied with plot size and basal area factor (BAF). Contrary to expectations, fixed-radius and especially variable-radius plots showed substantial size-dependence trends of statistical means. A differently designed computerized sampling of full density plots in Alaskan old-growth spruce-hemlock also indicated overestimation of basal area as BAF increased (LaBau 1967). Although variable-radius (point or prism) plots sample trees with a probability proportional to their basal area, theoretically they should still provide unbiased estimates of mean, per-acre values (Clutter 1957, Grosenbaugh and Stover 1957).

The above computer samples agreed with previous field studies showing that variable-radius plots sometimes gave estimates substantially different from those for fixed-area plots (Husch 1955, Clutter 1957, Wiant et al. 1984; but cf. Grosenbaugh and Stover 1957 and references cited in LaBau 1967). Investigators attributed these possible biases to a failure to include all trees within their size-specific, limiting distance in plots with small BAF. This should not have occurred with virtual sampling, suggesting that the biased sample estimates had some other, unknown explanation (Brooks and McGill 2004).

Brooks and McGill (2004) found that inventory parameters declined sharply with plot size to 0.1 ac and more gradually thereafter. Inventory parameters were level for BAF 10 to 15 and then increased sharply. When we imitated their sampling design of placing concentric, virtual plots within real forest plots, we obtained qualitatively identical results. Sampling bias was eliminated when we randomly sampled a composite forest created from a matrix of real forest plots.

Here we report the results of the random sampling to provide guidance for inventory design in mature, well-stocked, upland oak-hickory and oak-pine sawtimber forests. We included pole-sized trees to suit the design of continuous forest inventories and baseline inventories to assess carbon stocks. We also analyzed sawtimber volume for trees of the same size as those in Brooks and McGill (2004).

Methods

The real forest plots in this study were initially established to estimate acorn production (Vangilder 1997). The Missouri Ozark Forest Ecosystem Project (MOFEP) containing those plots was located in areas free from manipulation for at least 40 yr (Brookshire et al. 1997), and comprised mainly upland oak-hickory forests with an occasional strong pine component. Kabrick et al. (1997) provide a general analysis of the woody vegetation and environment.

Species individually comprising >10% of the trees in the plots and 60% in the aggregate were white oak (Quercus alba Linn.), scarlet oak (Quercus coccinea Muenchh.), and black oak (Quercus velutina Lam.). Other species individually comprising >1% of trees in the plots and 34% in the aggregate were shortleaf pine (Pinus echinata Miller), post oak (Quercus stellata Wangenh.), three hickory species (Carya spp.), blackgum (Nyssa sylvatica Marshall), and flowering dogwood (Cornus florida L.).

Plots were randomly located and oriented within ecological land types roughly in proportion to the type’s relative area within each of nine sites in five non-contiguous areas comprising 9,300 ac in three counties (Brookshire et al. 1997, Vangilder 1997). Within 130 rectangular plots (138.2 ft x 176.6 ft), species, dbh, and horizontal distance from the plot center to the estimated tree center were recorded in 1995 for all trees ≥4.5 in. dbh. Distance was measured with a steel tape to the nearest 0.1 ft, breaking chain when necessary on steep slopes.

We created a virtual composite forest by randomly arranging the 130 real plots in a 13 x 10 matrix to form an approximate square (Figure 1). A toroidal correction...
Table 1. Inventory parameters for entire composite forest.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>QMD (in.)</td>
<td>10.0</td>
<td>7.0 – 13.8</td>
</tr>
<tr>
<td>Trees/ac</td>
<td>148</td>
<td>71 – 282</td>
</tr>
<tr>
<td>Basal area (ft²/ac)</td>
<td>80.8</td>
<td>25 – 121</td>
</tr>
<tr>
<td>Biomass (dry tons/ac)</td>
<td>60.6</td>
<td>15 – 102</td>
</tr>
<tr>
<td>Volume (mbf/ac)</td>
<td>6.03</td>
<td>0.6 – 11.6</td>
</tr>
</tbody>
</table>

a All values are for trees ≥4.5 in.dbh except volume (International ¼ in. rule), which is for trees ≥11.6 in.dbh.
b For real forest plots.
c Quadratic mean diameter (QMD) = \([\sum \text{dbh}^2/N]^{0.5}\), where \(N\) = number of trees.
d mbf, thousand board feet

was applied to the edges by replicating the outer columns, rows, and corner plots outside their opposite members. It was as if the forest was folded on itself until the sides met to permit sampling without edge effects.

Sampling of the composite forest was conducted by testing whether a tree’s center occurred within virtual plots. Fixed-radius virtual plots ranged in size from 0.05 to 0.30 ac in 0.05-ac intervals, and variable-radius plots ranged from 5 to 30 BAF in 5-BAF intervals. For each plot size and BAF, 200 virtual plot centers were randomly generated within the area of the 130 original, real plots. To determine whether trees of specified dbh (in.) were counted in variable-radius plots, the limiting distance (ft) was calculated as 8.696(dbh)(BAF)^0.5 (Husch et al. 2003). For each virtual plot, number of trees, basal area, and biomass for all trees ≥4.5 in.dbh were calculated on a per-acre basis.

Tree density (trees/ac) in fixed-radius plots was calculated as the product of tree count and the reciprocal of plot area in acres. Tree density in variable-radius plots was calculated as \(\sum \text{BAF}/0.005454(\text{dbh})^3\) for all in-trees, where dimensions were measured in inches for dbh and ft²/ac for BAF (Avery and Burkhart 2002). Basal area in fixed-radius plots was calculated directly from dbh, measured to the nearest 0.1 in., thus avoiding bias introduced by grouping into diameter classes (cf. Grosenbaugh and Stover 1957). Basal area per acre in variable-radius plots was the product of tree count and BAF. Total, oven-dry, above ground biomass was estimated from dbh and species group according to Jenkins et al. (2003, Table 4). Sawtimber volume (International ¼ in. rule) was estimated from dbh, site index (measured for at least three trees per plot), and species group for the Central States according to Hahn and Hansen (1991, Table 6). Sawtimber volume was estimated for trees ≥11.6 in.dbh to facilitate comparison with the results of Brooks and McGill (2004), but it was not possible to use their volume equations for lack of merchantable height data. In accordance with sampling theory, “true” population parameters were calculated for the trees in the 130 real plots plus the inner half of their replicates (Table 1 and Figure 1).

Results and Discussion

Fixed-Radius Plots

Per-acre values of basal area, biomass and sawtimber volume increased slightly but significantly (analysis of variance [ANOVA], \(P < 0.05\)) with plot size, whereas tree density remained statistically constant (Figure 2). All means of inventory parameters in plots ≥0.1 ac were within 5% of the population’s true value.
As expected, the number of in-trees increased linearly with plot size, with a regression slope equal ($P > 0.99$) to the population mean of 148 trees/acre (Figure 2 and Table 1).

All coefficients of variation declined monotonically with plot size, but changed little for plots >0.2 acre (Figure 2). The coefficients of variation for biomass (not shown) were much smaller than those for sawtimber volume. This was expected because the inclusion of small trees better approximates the total volume supported by the carrying capacity in a given area by offsetting random variation in the occurrence of large trees.

The size class distribution of biomass was stable and close to the true value across plot sizes $\geq 0.1$ acre (Figure 3). Although the size class distribution of tree diameters was the typical reverse-J (data not shown), biomass peaked at intermediate size classes (12-16 in. dbh), above the quadratic mean diameter of 10.0 in. (Table 1).

**Variable-Radius Plots**

Mean tree number, basal area, and biomass per acre, but not sawtimber volume (ANOVA, $P < 0.01$), were statistically invariant with BAF (Figure 4). Means of all inventory parameters agreed to within 5% of the population’s true value.

From Husch et al. (2003), the expected number of in-trees is proportional to the reciprocal of BAF, as we found (Figure 4). Also as expected, the $y$-intercept was not significantly different from zero ($P = 0.61$), and the constant of proportionality (79.5) was not significantly ($P = 0.56$) different from the mean basal area of 80.8 ft$^2$/acre (Figure 4 and Table 1). The number of in-trees was less than the recommended minimum of five (Avery and Burkhart 2002) for BAF >15.

The coefficients of variation increased monotonically with BAF (Figure 4), as expected (Avery and Burkhart 2002). For a given number of in-trees, the coefficients of variation were 36% smaller for variable-radius than for fixed-radius plots for biomass and 33% smaller for sawtimber volume (data not shown). This agrees with the observation by Sukwong et al. (1971) of a higher efficiency (precision) of point compared with plot sampling for basal area estimation in artificial forests, as confirmed analytically by Matern (1972). For continuous forest inventories, however, permanent, fixed-area plots minimize sampling error with a
repeated-measures (“paired” plot) design. The size class distribution of biomass was stable and close to the true value for BAFs <20 (Figure 3).

**Explanation of Bias**

Our randomly located virtual plots produced results in agreement with sampling theory in that inventory means were essentially invariant with plot size or BAF, whereas variance decreased with plot size and increased with BAF (Figures 2 and 4). Our concentrically located plots produced biased estimates of inventory means (data not shown), much like those of Brooks and McGill (2004). Why?

Estimates of basal area per acre in forest simulations were unaffected by varying BAF from 8.7 to 21.8 (equivalent) in homogeneous forests (Mackisack and Wood 1990). BAF and other sampling design effects were observed, however, in a forest with strata having a 1.75-fold variation in basal area. The authors concluded that in the presence of such heterogeneity, much depends on the fortuitous location of the sample points.

Such variation (1.67-fold) occurred in the plots of Brooks and McGill (2004, Table 1), and to an even greater degree in ours (Table 1). Trees in our composite forest were highly aggregated, based on a regression slope coefficient of -0.37 for ln(CV) on ln(plot size) for tree density (data not shown). That coefficient was significantly different ($P < 0.01$) from the expected value of 0.5 for a random tree distribution (Reich and Arvanitis 1992). This, as well as the qualitative agreement between our study and that of Brooks and McGill (2004) for concentric plots, suggests that the observed biases were caused by a certain form of nonrandom sampling.

Sampling theory indicates that variable-radius plots should provide an unbiased estimate of the mean of inventory parameters regardless of tree spatial distribution, provided that plots are randomly located. To see this, consider that each tree is surrounded by an imaginary tree zone whose area is proportional to the tree’s basal area, as specified by the selected BAF. The probability that a tree will be sampled by a randomly located (point) plot is equal to the area of this imaginary tree zone divided by the area occupied by the entire tree population (Grosenbaugh and Stover 1957, Avery and Burkhart 2002). This probability is therefore a constant that is independent of the tree’s location within the sample plot (Avery and Burkhart 2002). Population totals for any desired inventory parameter can be estimated by summing individual (tree) values of the desired parameter divided by the corresponding individual probabilities of being drawn (Grosenbaugh and Stover 1957). The expected value of the desired variable for an array of sample plots equals the sum of the plot values of the inventory parameter multiplied by their respective relative frequencies (Dwass 1970). These relative frequencies are independent of tree spatial distribution for randomly located plots.

Concentric plots are not randomly located, however, and their statistical parameters are not independent. Detailed analysis of the concentric plots initially sampled in our study indicated that tree density was significantly higher (14.8%) in the inner 0.01-ac circular plot than in the outer ring of comparable width (11.8 ft; paired $t$-test, $P = 0.043$). The difference disappeared when the inner 0.03-ac plot was compared with an outer ring of comparable width (20.4 ft; paired $t$-test, $P = 0.34$). Thus, as concentric plot size increased or BAF decreased, mean inventory parameters quickly leveled and approached the true values because the high tree density in the very center of the plots was diluted (data not shown).

Tree density could decrease centrifugally in the real plots forming our forest matrix if trees far from the center were more likely to be missed during the field inventory. This has happened in other ungridded inventories of Forest Inventory Analysis plots in very high density coniferous forests of the Pacific Northwest (Vicente Monleon, pers. comm., US Forest Service, June 15, 2009), but this does not appear to explain our results. When our plots, first inventoried in 1995, were re-inventoried in 1998 and 2005 and ingrowth was excluded, only 136 new trees (1.3% of original count) were added. Moreover, the density of newly found trees decreased with distance from the plot center (data not shown).

A contour plot of tree density revealed that high values did not occur at the center of all plots despite the statistical differentiation of center and periphery (Figure 1 and paired $t$-tests described above). We cannot explain the cause of this statistical differentiation, but our results do show that very subtle variation in tree distribution can profoundly affect estimation of inventory parameters when concentric sampling is used.

**Implications for Inventory Design**

**Generality of Results**

The MOFEP area including our real forest plots provided a comprehensive profile of upland oak forest composition and structure in the eastern Ozarks (Shifley et al. 2000). As noted in Methods, however, sites were selected to be free of human disturbance for at least 40 years, so the plots represent mature, well-stocked stands of sawtimber-sized trees (Table 1 and Figure 3).

Our methods are applicable to all forests. Our general findings regarding the accuracy of fixed- and variable-radius plots should stand for both systematic
Random versus Systematic Sampling

Importantly, the dependence of inventory parameters on plot size and BAF, observed both by Brooks and McGill (2004) and in our preliminary analyses (data not shown), was associated with concentric sampling, not the sort of systematic sampling commonly practiced by foresters for convenience and cost effectiveness. Systematic sampling gives unbiased estimates of the mean, provided that some form of random selection is applied in the sampling process; e.g., to the initial plot (Avery and Burkhart 2002, Husch et al. 2003). Indeed, the more variable the forest area inventoried, the more likely that a particular systematic sample will give a better estimate of the mean than a completely random sample (Husch et al. 2003).

There is, however, no valid method of estimating the sampling error of a systematic sample (Avery and Burkhart 2002, Husch et al. 2003). A systematic sample of equidistant sampling units can estimate the maximum sampling error (Husch et al. 2003), and therefore provide a conservative estimate of precision. This will lead to overestimation of required sample size and unduly conservative hypothesis testing, which will increase inventory expense. Alternatively, systematic samples with at least five random starts may provide a reasonable means of estimating the sampling error (Husch et al. 2003). This could be incorporated in a multistage sampling scheme whereby the estimation of required sample size is refined at each stage. For practical reasons and to ensure full coverage of the sample area, locating plot centers on a grid is preferred, and the resulting sample error is taken as equivalent to that from randomly locating plot centers (Ken Cormier, pers. comm., US For. Serv., For. Manag. Serv. Ctr., Feb. 28, 2011).

Plot Type, Size, and Number

For forests structurally resembling those studied here, fixed-radius plots ≥0.1 ac will accurately estimate key inventory parameters and precision will scarcely improve in plots >0.2 ac. Inventories of biomass and sawtimber volume will be accurate at BAF 5-30, but more precise at low BAF. Both types of plots will characterize size class distributions well for ≥0.1 ac or ≤15 BAF.

Because our samples were random, our estimates of sampling error were unbiased. However, our synthetic population occupied 73 ac (130 real plots x 0.56 ac/plot), but sampled 9,300 ac. The between-plot variation was greater than would be observed on a real tract of comparable size because the real tract would be spatially more homogeneous. Our results are therefore conservative in their implications for sample size requirements to achieve a specified precision in field inventories.

Given that both systematic (other than concentric) and random samples can provide unbiased estimates of mean forest inventory parameters, the main value of virtual sampling as reported here is to provide an indication of how the coefficient of variation (sampling error) varies with plot size or BAF under random (or systematic) sampling. To plan a forest inventory that is statistically and practically efficient, just enough sampling units of the selected size should be measured to obtain the desired precision (Avery and Burkhart 2002).

Knowing how the coefficient of variation varies with plot size and BAF, the methods of Gambill et al. (1985) may be applied to determine the plot size or BAF that minimizes total inventory time, and therefore cost, at a specified precision. These procedures take account of tract size, coefficient of variation, and plot measurement time and may be used to compare the efficiencies of sampling by variable- and fixed-radius plots, as well as the effects of plot size or BAF. Rather than assuming that one size fits all, it becomes possible to select the sampling methodology best suited to forest characteristics and the purpose of the inventory.
Literature Cited


